

**MANAGING HETEROGENEOUS TRAFFIC ON RAIL FREIGHT NETWORKS
INCORPORATING THE LOGISTICS NEEDS OF MARKET SEGMENTS**

by

Oh Kyoung Kwon

B.A., Yeungnam University, 1983
M.C.P., Seoul National University, 1986

Submitted to the Department of Civil and Environmental Engineering
in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy
in Transportation Systems

at the

Massachusetts Institute of Technology

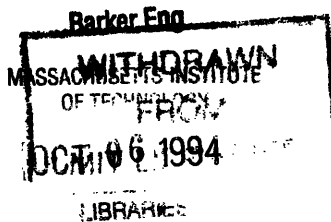
July 1994

© 1994 Massachusetts Institute of Technology
All rights reserved

Signature of Author
Department of Civil and Environmental Engineering
July 5, 1994

Certified by
Joseph M. Sussman
JR East Professor
Professor of Civil and Environmental Engineering
Thesis Supervisor

Accepted by
Joseph M. Sussman
Chairman, Departmental Committee on Graduate Studies
Department of Civil and Environmental Engineering



MANAGING HETEROGENEOUS TRAFFIC ON RAIL FREIGHT NETWORKS
INCORPORATING THE LOGISTICS NEEDS OF MARKET SEGMENTS

by

Oh Kyoung Kwon

Submitted to the Department of Civil and Environmental Engineering
on July 5, 1994 in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Transportation Systems

ABSTRACT

Recent studies have shown that the freight market can be divided into a number of market segments that have different preferences on elements of service quality and willingness to pay for additional service improvement. These studies suggest that a railroad must target markets in which to compete and must properly position its services to be competitive and profitable in those markets. The ability of a railroad to differentiate its services can provide a strategic advantage in the competitive transportation environment.

The research in this thesis aimed to develop both empirical and theoretical insights on how service differentiation strategy helps a railroad in gaining market share and in improving service quality and profitability.

Three areas were studied. First, the research examined whether and how railroads are currently differentiating services among different groups of traffic. The empirical analysis on trip time and reliability for different groups of traffic showed that railroads are currently differentiating services for major classes of traffic. Second, the research developed insights into the effects of service differentiation on the service provided to segmented markets and the costs to railroad of providing such differentiated services. Two simulation analyses showed that service differentiation strategies enable a railroad to provide market-sensitive services, to utilize service capacity more efficiently, and to potentially enhance profit. Third, the research examined how to further improve the ability of a railroad to differentiate services by designing an operating plan that fully considers the heterogeneity of shippers and their traffic. A dynamic freight car routing and scheduling model was developed as a practical decision support tool for managing heterogeneous traffic on rail freight networks. The results of the case study showed that the dynamic car routing and scheduling model can effectively allow a railroad to plan clearly differentiated services for heterogeneous shippers and their traffic, and thereby improve the ability of a railroad to differentiate services.

Thesis supervisor : Dr. Joseph M. Sussman

Title : JR East Professor

Professor of Civil and Environmental Engineering

ACKNOWLEDGMENTS

I would like to express my great appreciation for my supervisor, Professor Joseph Sussman. His guidance, support, and involvement during the entire process of this study have been invaluable and made the time a very wonderful learning experience.

I deeply thank Carl Martland who has served as a committee member and has provided me with valuable suggestions and insightful guidance throughout the whole thesis development.

I also want to thank my committee members, Professor David Bernstein and Professor Henry Marcus for their many helpful comments and suggestions. Professor Cynthia Barnhart allowed me to use her IBM RS 6000/370 and the OSL that were essential for the computational experiments to develop an important part of the thesis.

I thank many people in the Association of American Railroads. Mr. Tom Warfield collected the car cycle data for the empirical analysis of service differentiation and provided essential help in interpreting and analyzing the data. Mr. Peter French and Mr. Jim Lundgren offered numerous insights into rail freight transportation.

I am grateful to Dr. Patrick Little who has provided many good suggestions throughout the thesis development. I am also grateful to Rajesh Shenoj for his computer help.

I thank Dr. Shangyao Yan, Dr. Mark Hickman, Dr. Yusin Lee, Haiping Xu, Dinesh Gopinath, Jaikue Park and Daeki Kim for their support in my years in MIT. They are my good friends. I also thank Ms. Denise Brehm for her precise editorial comments.

I would like to acknowledge my parent for their prayers, affection and encouragement. Finally, I would like to express my heartfelt thanks to my lovely wife Nam Sook for her everlasting love, care, patience and support through the years in Boston.

The research is supported by a grant from the Association of American Railroads to the Center for Transportation Studies of MIT.

TABLE OF CONTENTS

Abstract	2
Acknowledgments	3
Table of Contents	4
List of Figures	7
List of Tables	9
1 Introduction	11
1.1 Background	11
1.2 Research statement	13
1.3 Structure of the thesis	14
2 Freight Market Segmentation and Service Differentiation	18
2.1 Introduction	18
2.2 Changes in shipper's motivations for transportation service choice	19
2.3 Relative importance of transit time and reliability for different shippers : total logistics cost analysis	21
2.4 Freight market segmentation and service differentiation	39
3 Empirical Analysis of Service Differentiation in Rail Freight Transportation	42
3.1 Introduction	42
3.1.1 Data source	43
3.1.2 Selection of O-D pairs	45
3.1.3 Trip time and reliability measures	47

3.2	Trip time and reliability of box car traffic	49
3.3	Trip time and reliability of covered hopper car traffic	56
3.4	Trip time and reliability of double-stack car traffic	61
3.5	Summary and conclusions	64
3.5.1	Comparison among different train services	64
3.5.2	Conclusions	66
4	Insights on Effects of Service Differentiation	68
4.1	Introduction	68
4.2	Case I : Train service between two terminals	69
4.2.1	Case description	69
4.2.2	Train make-up policies	71
4.2.3	Performance measures	74
4.2.4	Simulation analysis	76
4.2.5	Findings	94
4.3	Case II : Train service among multiple terminals	96
4.3.1	Case description	98
4.3.2	Train make-up policies	100
4.3.3	Performance measures	104
4.3.4	Simulation analysis	105
4.3.5	Findings	117
4.4	Conclusions	118
5	Hierarchical Structure of Decisions in Rail Operations and Models for Operating Plans : A Review	121
5.1	Introduction	121
5.2	Service design process incorporating heterogeneity of traffic	122
5.3	Hierarchical structure of decisions in rail operation	125
5.4	Review of modeling efforts in developing operating plans	128
5.4.1	Models for tactical decisions	129
5.4.2	Models for operational decisions	134

5.4.3	Conclusions	139
6	Freight Car Routing and Scheduling for Managing Heterogeneous Traffic on Rail Freight Networks	144
6.1	Introduction	144
6.2	Current freight car scheduling practice and areas for future improvement ..	145
6.3	Dynamic freight car routing and scheduling model	149
6.3.1	Incorporating service standards and traffic variability	149
6.3.2	Time-space network representation	151
6.3.3	Problem formulation : multicommodity flow problem	161
6.3.4	Solution approach : column generation solution approach	169
6.3.5	Inputs and outputs of the model	172
6.4	Case study	174
6.4.1	Case description	174
6.4.2	Analysis results	176
6.5	Conclusions	185
7	Conclusions and Future Research	188
7.1	Summary and conclusions	188
7.2	Contributions	199
7.3	Suggestions for future research	200
7.4	A final comment	201
	References	203
Appendix A	Rail Network Simulation Model	210
Appendix B	Input Data for Dynamic Freight Car Routing and Scheduling Model	222
Appendix C	Output of Dynamic Freight Car Routing and Scheduling Model : Train Capacity Utilization during the Planning Period	231
Appendix D	Output of Dynamic Freight Car Routing and Scheduling Model : Car Schedule by O-D, Departure Day and Departure Time	233

List of Figures

1.1	Structure of the thesis	17
2.1	Related logistics functions in the supply chain	20
2.2	Inventory level over time under conditions of demand and lead time uncertainty - continuous review inventory policy	23
2.3	Per unit logistics cost under different transit time distribution	29
2.4	Optimal reorder quantity under different transit time distribution	30
2.5	Optimal safety stock level under different transit time distribution	30
2.6	Average inventory per cycle with (Q^*, s^*) under different transit time distribution	31
2.7	Expected number of shortages per cycle with (Q^*, s^*) under different transit time distribution	32
2.8	Per unit inventory carrying cost under different transit time distribution	33
2.9	Per unit stock-out cost under different transit time distribution	33
2.10	Per unit ordering cost under different transit time distribution	34
2.11	Per unit in-transit inventory cost under different transit time distribution	34
2.12	Transit time and logistics cost change with different commodity value	36
2.13	Reliability and logistics cost change with different commodity value	36
2.14	Transit time and logistics cost change with different stock-out cost	38
2.15	Reliability and logistics cost change with different stock-out cost	38
3.1	Components of the car cycle	44
3.2	Example of O-D trip time distribution	48
3.3	Relation between mean trip time and reliability: box car service	54
3.4	Relation between time spent in other activities and reliability	54
3.5	Distribution of O-D pairs among different ranges of reliability performance : box car service	55
3.6	Relation between mean trip time and reliability: covered hopper car service	60
3.7	Distribution of O-D pairs among different ranges of reliability performance : covered hopper car service	60
4.1	Case network (I)	70
4.2	Effects of train capacity on service and cost	77
4.3	Effects of initial empty car on service and cost	78
4.4	Trip time distribution of all traffic under different train make-up rules	79
4.5	Trip time distribution by priority class under train make-up rule 1	83
4.6	Trip time distribution by priority class under train make-up rule 2	83
4.7	Trip time distribution by priority class under train make-up rule 3	84
4.8	Effect of train capacity on average trip time by priority	86
4.9	Effect of train capacity on variance of trip time by priority	86
4.10	Sum of operating cost and logistics cost by train capacity and initial car inventory (train make-up rule 1)	88
4.11	Sum of operating cost and logistics cost by train capacity and initial car inventory (train make-up rule 2)	89
4.12	Sum of operating cost and logistics cost by train capacity and initial car inventory (train make-up rule 3)	89

4.13 Case network (II)	98
4.14 Train service network (II)	99
4.15 Example of additional switching	103
4.16 Transit time by priority under different traffic mixes	113
4.17 Difference in per car logistics cost between static and dynamic policies: under different traffic mixes	117
4.18 Difference in per car logistics cost between static and dynamic policies: under different ratio of shipment value between high and low priority	117
5.1 Service design incorporating market segmentation	124
5.2 Hierarchical decisions in rail operations	126
6.1 Current car scheduling system	146
6.2 Example of car schedule	146
6.3 Dynamic car scheduling system	148
6.4 Representation of train schedule	154
6.5 Representation of blocks	156
6.6 Representation of terminal processing	158
6.7 Representation of delivery time windows	160
6.8 Representation of car-to-block and block-to-train assignments	164
6.9 Case network (III)	175
6.10 Train service network (III)	175
6.11 Train capacity utilization with shortest-time path based car schedule	179
6.12 Train capacity utilization with optimal car schedule	180
6.13 Train capacity bottlenecks	181
6.14 Mean lateness to service standards of different traffic classes for case problems	184
6.15 Standard deviation of lateness to service standards of different traffic classes	184
A.1 Structure of simulation model	211
A.2 Car movement through terminal and related operating plans	215

List of Tables

2.1 Characteristics of different shippers examined	27
2.2 Per unit logistics cost under different transit time distribution (value of commodity = \$50,000 per car-load)	28
2.3 Per unit logistics cost under different transit time distribution (value of commodity = \$100,000 per car-load)	35
2.4 Per unit logistics cost under different transit time distribution (value of commodity = \$50,000 per car-load, stock-out cost = 24%)	37
2.5 Estimation of demand elasticities to service quality	40
3.1 Data source	45
3.2 Selection of O-D pairs	47
3.3 Car cycle time : box car service	49
3.4 Car cycle time of local and interline moves : box car service	50
3.5 Aggregate trip time and reliability performance of selected O-D pairs: box car service	51
3.6 Correlation coefficients between variables : box car service	52
3.7 Trip time and reliability of large and small O-D pairs	53
3.8 Distribution of O-D and commodity groups among different ranges of reliability performance	56
3.9 Car cycle time: covered hopper car service	57
3.10 Car cycle time of local and interline moves : covered hopper car service	57
3.11 Aggregate trip time and reliability performance : covered hopper car service	58
3.12 Correlation coefficients between variables : covered hopper car service	59
3.13 Car cycle time : double-stack car service	61
3.14 Car cycle time of local and interline moves : double-stack car service	62
3.15 Aggregate trip time and reliability performance : double-stack car service	62
3.16 Trip time and reliability performance by direction and distance : double-stack car service	63
3.17 Trip time and reliability of different carriers : double-stack car service	64
3.18 Car cycle time for different train services	65
3.19 Trip time and reliability performance of different train services	66
4.1 Demand distribution by priority	71
4.2 Trip time performance under different train make-up rules	81
4.3 Testing differences in trip time performance among traffic classes	82
4.4 Testing differences in trip time performance among train make-up rules	82
4.5 Per car logistics cost under different train make-up rules	85
4.6 Testing differences in per car logistics costs among traffic classes	85
4.7 Per car operating and logistics costs change as train capacity increases	87
4.8 Train cost allocation process	92
4.9 Train cost per car by priority for different train make-up rules	92
4.10 Operating cost per car by priority for different train make-up rules (one simulation run)	93

4.11	Operating cost per car by priority for different train make-up rules	93
4.12	Testing differences in per car operating cost among traffic classes	94
4.13	Example of differentiated service products	94
4.14	Effects of traffic variability on service of all traffic : static make-up policy	107
4.15	Effects of train transit time variability on service of all traffic	108
4.16	Effects of traffic variability on per car operating cost	108
4.17	Effects of train transit time variability on per car operating cost	109
4.18	Effects of traffic variability on service by priority	110
4.19	Effects of train transit time variability on service by priority	110
4.20	Operating cost by priority under different traffic variability	111
4.21	Operating cost by priority under different train transit time variability	111
4.22	Effects of traffic mix on service by priority : static make-up policy	113
4.23	Effects of traffic mix on service by priority : dynamic make-up policy	114
5.1	Models for tactical decisions	142
5.2	Models for operational decisions	143
6.1	Train schedule : example	152
6.2	Block definition and train make-up plan : example	153
6.3	Summary of the inputs to the model	173
6.4	Summary of the outputs from the model	173
6.5	Test cases with different traffic mix	177
6.6	Problem size and computational performance	177
6.7	Scheduled transit time for different day of week : high priority	182
6.8	Scheduled transit time for different day of week : low priority	182
6.9	Optimal solutions for case problems	183
7.1	Trip time and reliability performance of different train services	190
7.2	Trip time performance under different train make-up rules	192
7.3	Cost per car by priority for different train make-up rules	193
7.4	Service level by priority : sensitivity on traffic mixes	197

CHAPTER 1

Introduction

1.1 Background

Improving service quality has become a more important issue to the railroad industry in this era of deregulation, initiated by the Staggers Act in 1980.¹ Shippers have become more sensitive to logistics-related costs and they now require transportation services that can improve their logistics processes throughout the supply chain. They seek to reduce inventories of raw materials and in-progress and finished products, and they need freight transportation services that can be closely coordinated with their procurement, production and delivery systems.

Freight transportation service can be measured by a number of factors such as price, transit time, reliability and other customer services. Surveys of shippers have frequently cited both the importance of service reliability in mode and carrier selections and the railroad's inability to achieve the high standards for reliability established by the trucking industry (Mercer [1991], Intermodal Index [1992, 1993]). There have been continuous efforts to improve the overall service

¹ This statute, commonly referred as the Staggers Act, was enacted on October 14, 1980 (Pub. L. 96-448, 94 Stat. 1895). It removed much of the ICC's authority over rail rates, reduced the importance of rate bureaus and authorized railroad contracts with shippers. The economic performance in the deregulated rail industry has been discussed by Frielaender [1991] and Frielaender, et al. [1992].

reliability in the rail industry to be more service competitive in the market (Martland, et al. [1992]).

An important question is whether such service improvement efforts are justified in all markets.

Recent studies have shown that the freight market can be divided into a number of market segments that have different preferences on elements of service quality and willingness to pay for additional service improvement (McGinnis [1978], Vieira [1992]). These studies concluded that it is not possible to implement any single marketing and service management program that will satisfy all current and potential shippers. A railroad must target markets in which to compete and must properly position its services to be competitive and profitable in those markets. The ability of a railroad to differentiate its services can provide a strategic advantage in the competitive transportation environment.

Opportunities for market segmentation and service differentiation can be helpful in both in gaining market share and in improving service quality. A railroad can better match its operating capability with the different needs of shippers in current and potential market segments. The underlying concept is that the marketing department can identify different classes of shippers, who have different requirements with respect to cost, transit time, reliability and other service characteristics. The operating department then adjusts its operations to provide the desired level of service to each class of shippers.

In order to match its operating capability with the needs of shippers in different market segments, a railroad needs to put its efforts into identifying homogeneous segments of shippers, and improving its ability to design and deliver differentiated service to these segmented markets.

1.2 Research statement

The research described here studies service differentiation in the rail transportation context. The objective of the research is to develop both empirical and theoretical insights on how service differentiation helps a railroad in gaining market share and in improving service quality. Three major areas of research were developed to accomplish the research objective.

- Development of empirical analysis on current service differentiation practice in the rail industry
- Develop insights on effects of service differentiation on service and costs of rail operations
- Develop better operating plans to improve the ability of railroads to differentiate services

First, the research examines how railroads are currently differentiating services among different groups of freight traffic. Railroads have practiced some level of service differentiation for broadly classified freight traffic, dividing it into three major types: general merchandise train service², unit train service and intermodal train service. An empirical analysis was performed of transit times and the reliability of car movements for these three major train services. Car cycle information for three car types was collected for this purpose: box car data for general merchandise train service, covered hopper car data for unit train service, and double-stack car data for intermodal train service. Transit times and various reliability measures were evaluated and compared for different train services.

Second, the research develops insights into the effects of service differentiation 1) on the service provided to segmented markets and 2) on costs to railroad. Railroads have been able to

² Some railroads run dedicated train service for the automobile (e.g., CSXT).

provide different levels of service to some classes of customers (for example, intermodal, auto) by establishing different operating plans for different train services, by giving priorities to certain trains and blocks, and by taking special action to make sure that certain train connections are made and that local deliveries take place on time. It is not clear, however, how much service can be differentiated among different classes of traffic, how much it costs overall to provide premium service to some classes of traffic, and how much benefit can be gained in terms of service levels to shippers and revenue potential to a railroad. Based on the analyses of example networks, it was found that a service differentiation strategy enables a railroad to provide fast and highly reliable service to service-sensitive shipper and slow and less reliable (but less costly) service to cost-sensitive shippers, without wasting its resources in improving overall service performance for all shippers. It was also found that varying the way that heterogeneity of traffic is incorporated into an operating plan results in different service and cost levels to different classes of traffic.

In this context, the research reviews hierarchical decisions and models for rail operations and finds that an operating plan that fully considers the heterogeneity of traffic has not yet been established. Current freight car routing and scheduling practice was identified as one important area to be improved for the purpose of service differentiation. The research developed a dynamic car routing and scheduling model to support the management of heterogeneous traffic on rail freight networks.

1.3 Structure of the thesis

The organization of the thesis is summarized in Figure 1.1.

Chapter 2 discusses the motivations for service differentiation from the perspectives of both gaining market share and improving service quality, based on findings of recent demand studies and total logistics cost analysis.

Chapter 3 discusses results from empirical analysis of the current practice of service differentiation in rail industry. Transit time and various measures of reliability were evaluated for three different train services: general merchandise train service, unit train service and intermodal train service.

To develop insights into the effects of service differentiation on services to different classes of shippers and costs to railroad, two types of train service were modeled, simulated and analyzed in Chapter 4.

To improve their ability to differentiate services, railroads need to re-evaluate and re-design hierarchical decisions for operations so as to incorporate the service requirements of different classes of shippers into their operations. Chapter 5 reviews the hierarchical decisions and the current practically- and theoretically-developed models for rail operations to examine the extent to which existing operating plans fully consider the heterogeneity of traffic.

Chapter 6 develops an optimization based decision support model for routing and scheduling heterogeneous traffic on rail freight networks. A case study is done for a rail network based on the data of a major railroad. Implementation issues are discussed.

Chapter 7 summarizes all important findings from the previous chapters and draws conclusions. It also discusses future research areas.

These entire research are fit together in the following manner. In Chapter 2, the heterogeneity of shippers and their traffic in the freight market is discussed. The ability of a railroad to differentiate its services can provide a strategic advantage in the competitive transportation environment. Based on the motivation resulting from the analysis of Chapter 2,

Chapter 3 empirically studied whether and how railroads are currently differentiating services. It is found that railroads are currently differentiating services for major classes of freight traffic. To further develop insights on the effects of service differentiation in the rail transportation context, Chapter 4 examined the effects of service differentiation on service levels for different classes of traffic and costs to railroad, based on the simulation analyses of idealized example networks. One of the important findings is that varying the way to incorporate heterogeneity of traffic into an operating plan results in different service levels to different classes of traffic and costs to railroad. Chapter 5 reviewed practically- and theoretically-developed models for hierarchical decisions in rail operations and found that the existing operating plan does not fully consider the heterogeneity of traffic. Freight car scheduling practice is identified as one of the important areas to be improved for the purpose of accomplishing service differentiation. Car scheduling is becoming a more important aspect of the operating plan, as railroads pursue more carefully scheduled and planned operations, and as more shippers demand the car schedule information for planning their procurement, production and distribution processes. Based on this finding, Chapter 6 proceeded to develop a dynamic car routing and scheduling model to support the management of heterogeneous traffic on rail freight networks. It is shown that the dynamic car routing and scheduling model effectively supports routing and scheduling heterogeneous traffic on rail freight networks, and thereby improving the ability of railroads to clearly differentiate services. Chapter 7 summarizes the important findings of these research works and draws several conclusions and makes suggestions for additional research.

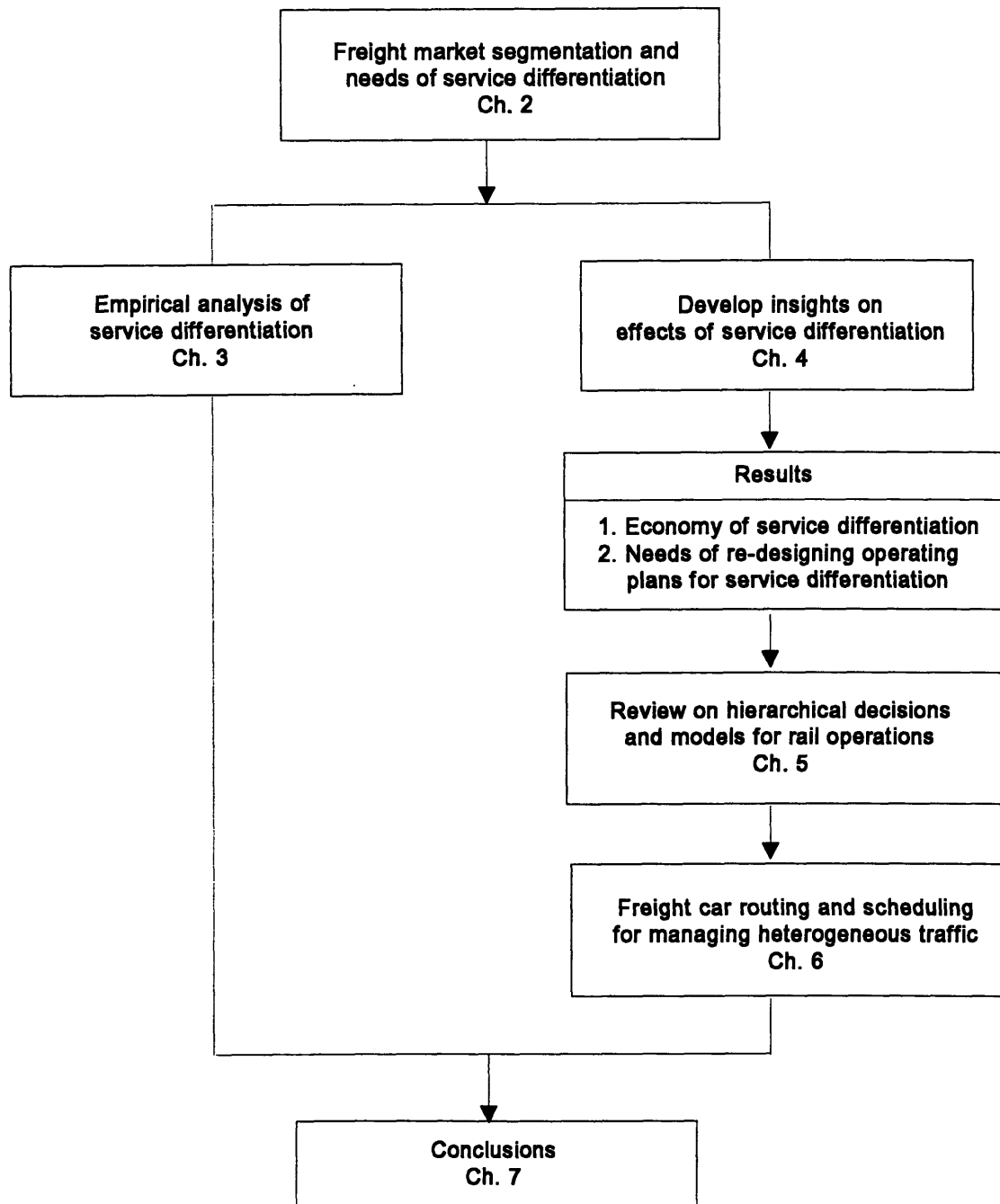


Figure 1.1 : Structure of the thesis

CHAPTER 2

Freight Market Segmentation and Service Differentiation

2.1 Introduction

In this chapter, the demand-side characteristics of the freight market are discussed. The changes in shipper's motivation for transportation mode and carrier choice are explored. The importance of transit time and reliability are emphasized based on the empirical freight demand studies and recent industry surveys. The importance of transit time and reliability and their relative importance for shippers with different characteristics are discussed based on the total logistics cost analysis. The importance of incorporating the heterogeneity of shippers and their traffic to the development of a carrier's service strategy is discussed taking recent studies on freight market segmentation into account, and the concept of service differentiation is introduced.

2.2 Changes in Shipper's Motivations for Transportation Service Choice

In the current competitive business environment, more shippers now consider the improvement of their logistics process to be critical to their competitiveness. An increase in the short product life cycle and the proliferation of new products and product families are dramatically affecting the management and deployment of inventory in the supply chain to maintain competitive customer levels. Customers are demanding higher quality at lower cost, rapid response, and immediate availability at the time of procurement and usage. Suppliers are increasingly being measured and evaluated, not only on the basis of production capability and quality, but also on the basis of ability to deliver just-in-time in small lots at greater frequencies to point of use.

Shippers now focus on the improvement of the logistics process for the entire supply chain, rather than on the improvement of individual logistics functions. Figure 2.1 shows related logistics functions of the supply chain. Streamlining the flow of goods from the supplier to the customer requires a transportation system that is part of a logistics process that connects suppliers, carriers and buyers. The growth of competition in the goods market has fostered shippers' interest in transportation service as an important link in the supply chain.

Understanding both the shippers' expectations of service and the ability to provide the desired service are therefore critical in developing a carrier's service strategy. There have been substantial studies to understand shipper's transportation mode and carrier choice behavior. Reviews on freight demand models can be found in Chiang, et al. [1981], McGinnis [1989] and Shen [1992]. These studies consistently identified the importance of transit time and reliability as critical decision factors in shippers' choice of mode and carrier.

Supply			Customer	
Sourcing	Inbound logistics	Manufacturing	Outbound logistics	After-sales service
Procurement planning	Traffic planning	Production planning and scheduling	Traffic planning	Spares planning and scheduling
Supplier evaluation and management	Carrier evaluation and management	Manufacturing	Carrier evaluation and management	Service center management
Component engineering	In-house fleet management	Transportation between plants	In-house fleet management	Return/service transportation
	Inbound delivery planning and monitoring		Outbound delivery planning and monitoring	
	Raw material/component warehousing		Plant to DC Inter-DC DC-customer	
			Finished goods warehousing and warehouse management	
			Inventory planning and deployment	

DC = Distribution Center

Source : Gopal and Cypress [1993], p. 3

Figure 2.1 : Related logistics functions of the supply chain

McGinnis [1989] reviewed 11 empirical freight demand studies to identify important determinants in shippers' transportation mode and carrier choices. This review identified one transportation cost variable (freight rates) and six non-cost transportation variables that affect freight transportation choice (reliability; transit time; loss and damage, and claims processing; shipper market considerations; carrier considerations; and product characteristics). While the relative importance of these variables varied from study to study, it concluded that reliability seemed to be consistently more important than all other variables, including transit time and freight rates. It also suggested that the variables that affect modal choice vary with the shipper.

Such conclusions are consistent with recent shipper surveys that show the importance of reliability in mode and carrier selection. Industry surveys also frequently cite the rail industry's inability to achieve the high standards for reliability established by the trucking industry (Mercer [1991], Intermodal Index [1992, 1993]).

2.3 Relative Importance of Transit Time and Reliability for Different Shippers : Total Logistics Cost Analysis

Several researchers have suggested that the total logistics cost analysis, including both transportation and inventory issues, can be used as an analytical framework to understand shipper's transportation decisions and to examine the shipper/carrier interaction in transportation (Roberts, et al. [1976], Sheffi [1977]). Blumenfeld, et al. [1985] analyzed the trade-offs between transportation, inventory and production costs, and Allen, et al. [1985] analyzed the effect of transit time and reliability on a shipper's logistics cost.

For some commodities, the value of inventory or the needs of the production process lead shippers to demand short transit times with little or no variation in transit times (such as “just-in-time” processes). For other commodities, most notably bulk goods, the value of the commodity may be considerably lower than the equipment in which it moves, so that shippers do not object to holding inventories and safety stocks, and require only that a certain volume is moved within a relatively long window.

The prices customers are willing to pay will also vary. Some customers may be willing to pay a substantial premium to ensure high quality service, while other may not. In both cases, their decisions are based on the logistics costs that they face. If the service provided is matched to the customer’s desires and consistent with expectations, the service may be considered good, even if that service would not be acceptable to a different customer (Martland, et al. [1993]).

In this section, we analyze the relative importance of transit time and reliability to heterogeneous shippers with different characteristics. A detailed analysis of each logistics cost component is performed to examine how logistics cost change under different transit times and reliability of transportation service. Consider a shipper that has a single storage location and continuous review inventory policy and both the demand and the lead time (transit time) are probabilistic. Figure 2.2 shows the inventory level over time, under the conditions of demand and lead time uncertainty, with this inventory policy.

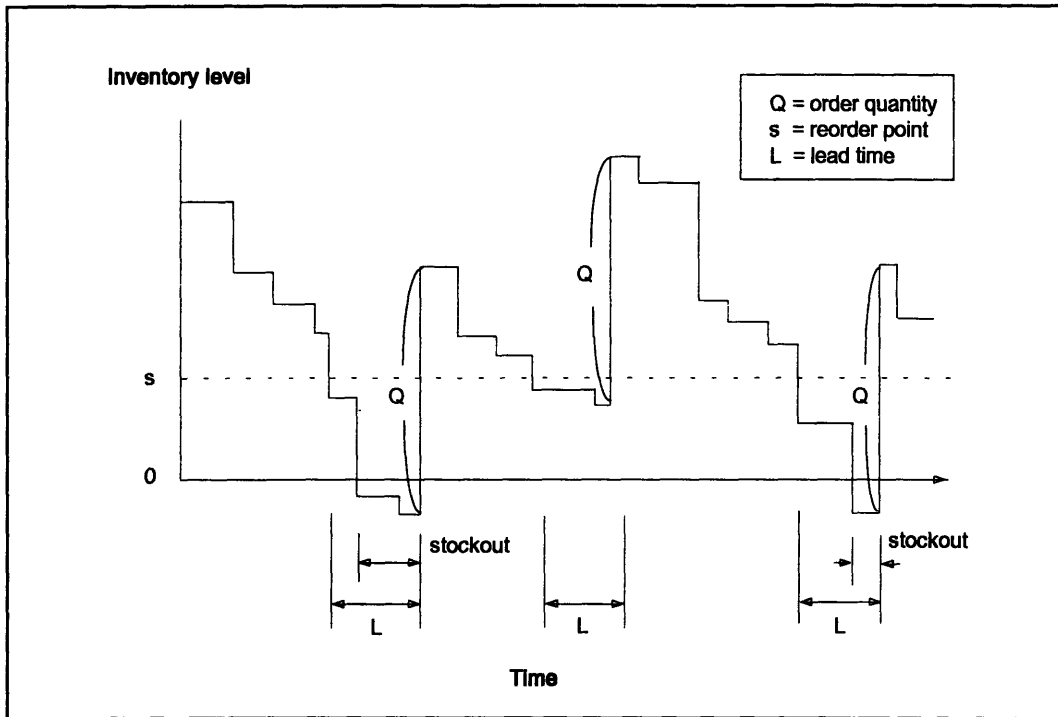


Figure 2.2 : Inventory level over time under conditions of demand and lead time uncertainty - continuous review inventory policy

The total logistics cost function can be represented as the equation (2.1) (see Johnson and Montgomery [1974]). It includes ordering cost, inventory carrying cost, in-transit inventory cost, stock-out cost, and transportation cost.

$$TC = \frac{A\bar{D}}{Q} + VW\left[\frac{Q}{2} + s - \bar{L}d\right] + \frac{VY\bar{L}D}{365} + \frac{K\bar{a}(s)\bar{D}}{Q} + R\bar{D} \quad (2.1)$$

where,

TC : total logistics cost per year (dollars)

Q : reorder quantity (units)

s : reorder point or safety stock (units)

\bar{D} : expected annual demand = $365 \bar{d}$ (units)

\bar{d} : average daily demand (units)

\bar{L} : average lead (transit) time (days)

\bar{Ld} : demand during a lead time (units)

$\bar{a}(s)$: expected shortage per cycle (units)

A : ordering cost (dollars)

V : per unit value of commodity (dollars)

W : annual inventory carrying cost at storage location (percentage)

Y : in-transit inventory cost (percentage)

K : per unit stock-out cost (dollars)

R : per unit transportation rate (dollars)

The stock-out cost per cycle is $K \cdot \bar{a}(s)$ assuming all shortages incur a cost of K per unit.

It means that each shortage incurs a “penalty” of K dollars to a company for not having the product when a customer demanded it. This penalty includes reordering cost, additional transportation and handling cost, cost of emergency supply from an alternative supplier, and costs of present and future lost sales. The level of stock-out cost can be different for different product and firm. A stock-out cost for customer with “just-in-time” processes is much higher than for other customers. In addition, firms that are monopolistic in the market for their products may experience less impact on present and future lost sales than firms that have many competitors that produce substitutable products.

The expected stock-out is a function of the reorder point s . A stock-out can only occur if the demand during the lead time exceeds the reorder point. The amount of shortage when the new order quantity arrives is $a(x, s) = \max[0, x - s]$, which has the expected value $\bar{a}(s)$

$$\bar{a}(s) = \int_s^{\infty} (x - s) f(x) dx \quad (2.2)$$

The derivation of the joint probability distribution of demand during lead time from both the lead time and the daily demand distribution is well known.¹

Given s , $\bar{a}(s)$ can be calculated using a normal approximation

$$\begin{aligned} \bar{a}(s) &= \sigma_x L' \left(\frac{s - \bar{x}}{\sigma_x} \right) \\ &= \sqrt{\bar{L} \text{Var}[d] + \bar{d}^2 \text{Var}[L]} L' \left(\frac{s - \bar{L}\bar{d}}{\sqrt{\bar{L} \text{Var}[d] + \bar{d}^2 \text{Var}[L]}} \right) \end{aligned} \quad (2.3)$$

where,

¹ Suppose the lead time and the daily demand are normally distributed.

$$L \sim N(\bar{L}, \sigma_L^2) = N(\bar{L}, \text{Var}[L])$$

$$d \sim N(\bar{d}, \sigma_d^2) = N(\bar{d}, \text{Var}[d])$$

Then, the demand during lead time is also normally distributed

$$x \sim N(\bar{x}, \sigma_x^2) = N(\bar{x}, \text{Var}[x])$$

where,

$$\bar{x} = \bar{L}\bar{d}$$

$$\text{Var}[x] = \begin{cases} \bar{L} \text{Var}[d] + \bar{d}^2 \text{Var}[L] & \text{if } L \text{ and } d \text{ are independent} \\ \bar{L} \text{Var}[d] + \bar{d}^2 \text{Var}[L] + \text{Cov}[d, L] & \end{cases}$$

Var[d] : variance of daily demand (units)

Var[L] : variance of lead time (days)

$L'(u)$: unit normal linear-loss integral

As a result, stock-out cost can be represented as a function of reorder point s , daily demand distribution and transit time distribution. It will be increased when the average or variance of demand is increased. It will also be increased when transit time or variance of transit time is increased. This thesis focuses considerable attention on studying how railroads provide service to different shippers on these two dimensions.

In all, total logistics cost (2.1) can be viewed as a function of shipper characteristics (annual demand, daily demand distribution, inventory carrying cost, in-transit inventory cost and stock-out cost), product characteristics (value of commodity) and carrier's service level (transportation cost, transit time and reliability measured by variance of transit time). It allows us to evaluate the shipper's logistics cost savings when the service level is improved.

Shippers try to minimize a total logistics cost (2.1). The objective is to find the values of Q^* and s^* that minimize TC . Differentiating with respect to Q and s yields:

$$Q^* = \sqrt{\frac{2\bar{D}[A + K\bar{a}(s^*)]}{VW}} \quad (2.4)$$

$$1 - F(s^*) = \frac{VWQ^*}{K\bar{D}} \quad (2.5)$$

Iterative procedure can be used to solve equations (2.4) and (2.5) simultaneously.

Step 1 : solve Q assuming $\bar{a}(s) = 0$

Step 2 : solve $1-F(s)$ and s given Q

Step 3 : calculate $\bar{a}(s)$ given s using normal approximation

Step 4 : solve Q' given $\bar{a}(s)$ and s' given Q'

Step 5 : convergence test

if $|Q' - Q| < \epsilon$ and $|s' - s| < \epsilon$, go to step 6²

else, go to step 2

Step 6 : obtain optimal solution $(Q^*, s^*) = (Q', s')$

Based on this analytical framework of total logistics cost, the importance of service quality with respect to transit time, reliability and transportation cost to a shipper, and the shipper/carrier interaction can be discussed.

To examine relative importance of transit time and reliability for shippers with different characteristics, three different types of shippers that have different value of commodity and stock-out cost are considered.

Table 2.1 : Characteristics of different shippers examined

	Commodity value (\$/car-load)	Stock-out cost (% of commodity value)
Shipper A	50,000	12
Shipper B	100,000	12
Shipper C	50,000	24

Other characteristics are the same for all three shippers:

- annual demand : 3,650 units (car-loads)
- daily demand is normally distributed

² A constant $\epsilon = 10^{-6}$ is used as a convergence criterion.

$$d \sim N(\bar{d}, \sigma_d) = N(10, 3)$$

- ordering cost : \$500 per order
- annual inventory carrying cost : 10%
- in-transit inventory cost : 10%

To mainly examine the effects of transit time and reliability (measured by variance of transit time) on the shipper's total logistics cost, we exclude transportation cost in the analysis.³

Per unit logistics cost is computed for mean transit time ranging from 1 to 10 and variance of transit time ranging from 0 to 10.

Consider a shipper A. Table 2.2 summarizes the per car logistics cost for different means and variances of transit time. Total logistics cost increased as mean transit time increased and as variance of transit time increased.

Table 2.2 : Per unit logistics cost under different transit time distribution
(value of commodity = \$50,000 per car-load)

E[L]	Var [L]										
	0	1	2	3	4	5	6	7	8	9	10
1	57.4	89.7	105.3	117.3	127.3	136.1	144.1	151.4	158.2	164.6	170.6
2	73.9	104.0	119.4	131.3	141.3	150.1	158.0	165.3	172.1	178.4	184.4
3	89.7	118.2	133.5	145.3	155.3	164.0	172.0	179.2	186.0	192.3	198.3
4	105.2	132.5	147.6	159.3	169.2	178.0	185.9	193.1	199.9	206.2	212.2
5	120.5	146.7	161.7	173.3	183.2	191.9	199.8	207.1	213.8	220.1	226.0
6	135.6	160.9	175.8	187.4	197.2	205.9	213.7	221.0	227.2	234.0	239.9
7	150.6	175.1	189.8	201.4	211.2	219.8	227.7	234.9	241.6	247.8	253.8
8	165.5	189.4	203.9	215.4	225.1	233.8	241.6	248.8	255.5	261.7	267.7
9	180.3	203.6	218.0	229.4	239.1	247.7	255.5	262.7	269.3	275.6	281.5
10	195.1	217.8	232.1	243.4	253.1	261.7	269.4	276.6	283.2	289.5	295.4

³ It can be included to analyze the effect of transportation rate change.

Each component of total logistics cost is analyzed to understand the reasons total logistics cost change under different transit time distributions, especially the reasons for a steeper increase of shipper's logistics cost under higher variance of transit time. As transit time or the variance of transit time increased, a cost-minimizing shipper tends to increase both reorder quantity and safety stock levels to reduce the expected number of shortages. Figure 2.4 shows that an increase of optimal reorder quantity (Q^*) due to unit transit time variance increase is larger than that due to unit transit time increase. On the other hand, an increase of optimal safety stock level (s^*) due to transit time increase is larger than due to transit time variance increase (Figure 2.5).

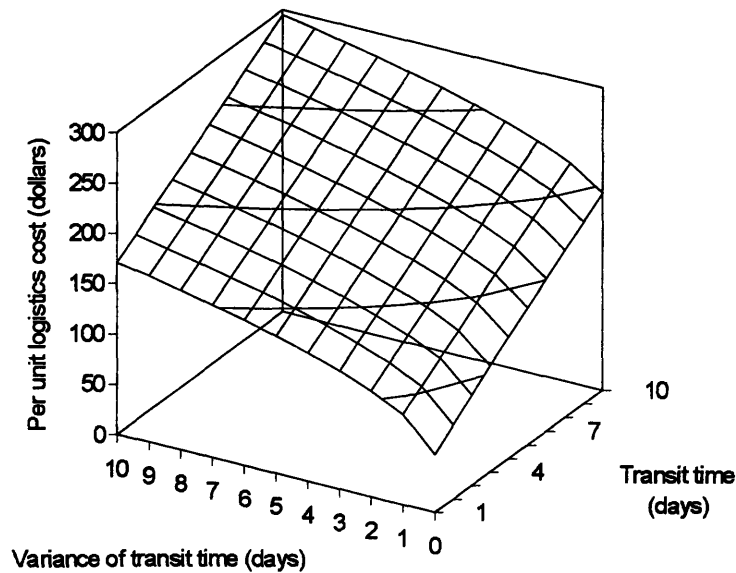


Figure 2.3 : Per unit logistics cost under different transit time distribution (value of commodity = \$50,000 per car-load)

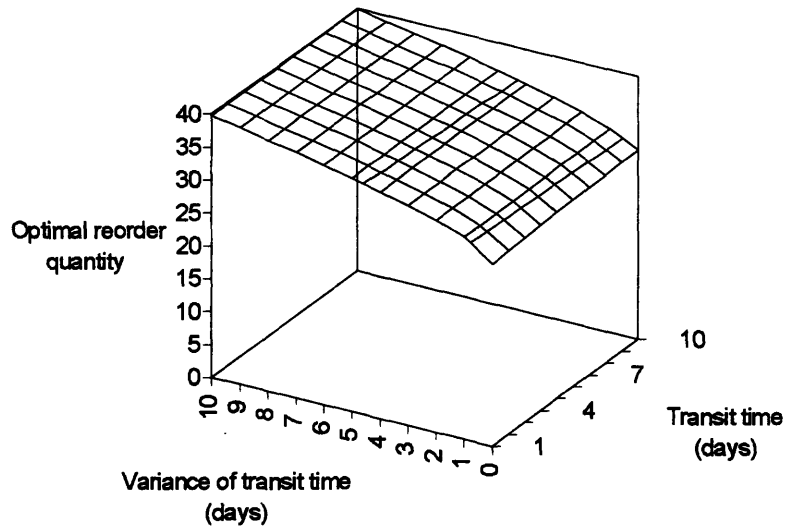


Figure 2.4 : Optimal reorder quantity under different transit time distribution

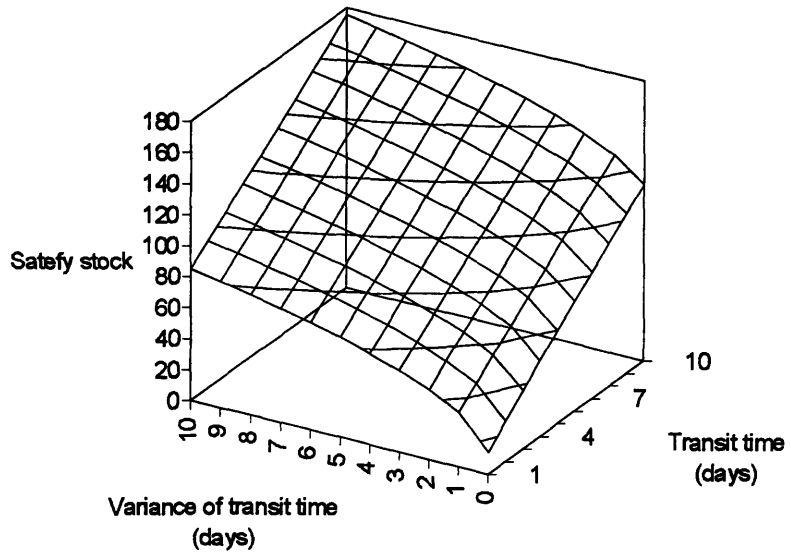


Figure 2.5 : Optimal safety stock level under different transit time distribution

Under such an optimal policy for reorder quantity and safety stock level, we find that the increase of both average inventory level⁴ and expected number of shortage due to unit transit time variance increase are larger than due to unit transit time increase (Figure 2.6 and Figure 2.7).

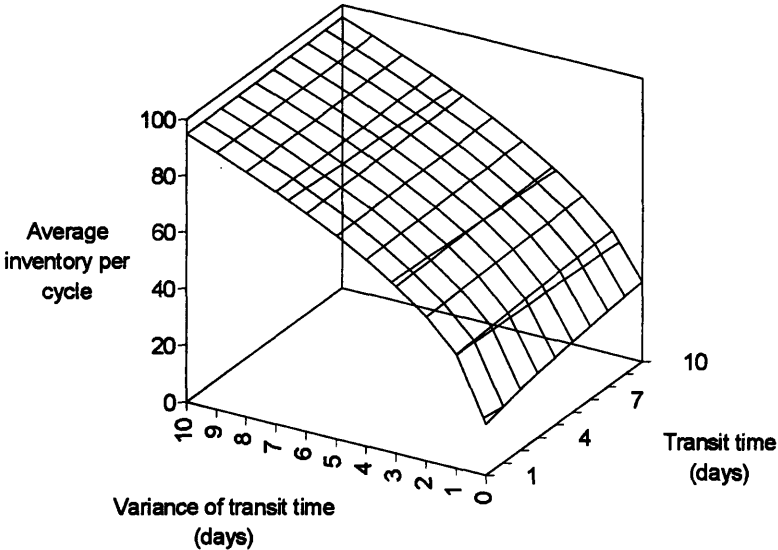


Figure 2.6 : Average inventory per cycle with (Q^*, s^*) under different transit time distribution

⁴ A higher safety stock level (s) under a longer transit time is offset by a higher demand during the lead time (\overline{Ld}) in determining the average inventory per cycle $Q/2 + s - \overline{Ld}$.

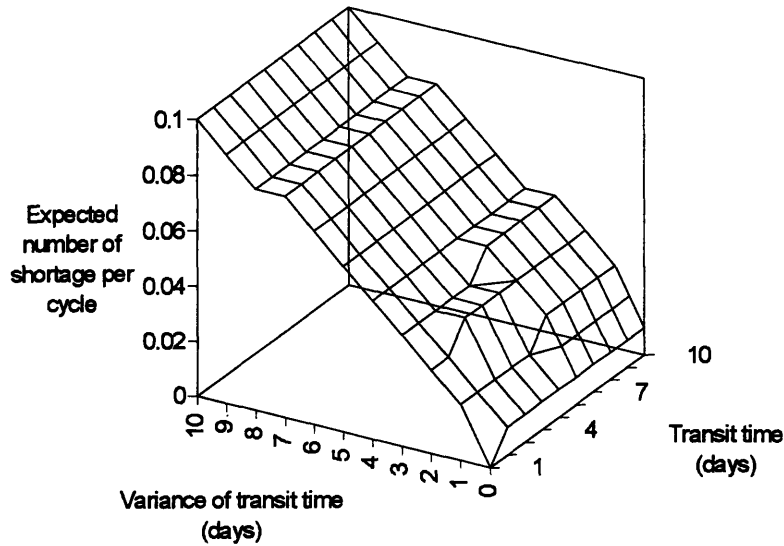


Figure 2.7 : Expected number of shortages per cycle with (Q^*, s^*) under different transit time distribution

Figure 2.8 to Figure 2.11 show the change of each component of total logistics cost under different transit time distributions. An increase of both inventory carrying cost and stock-out cost due to transit time variance increase are significantly larger than due to transit time increase (Figure 2.8 and Figure 2.9). As transit time or variance of transit time is increased, ordering cost is decreased since the optimal reorder quantity is increased (Figure 2.10). In-transit inventory cost is only sensitive to transit time increase (Figure 2.11). As a result, we had per unit logistics cost under different transit time distribution (see Table 2.1 and Figure 2.3).

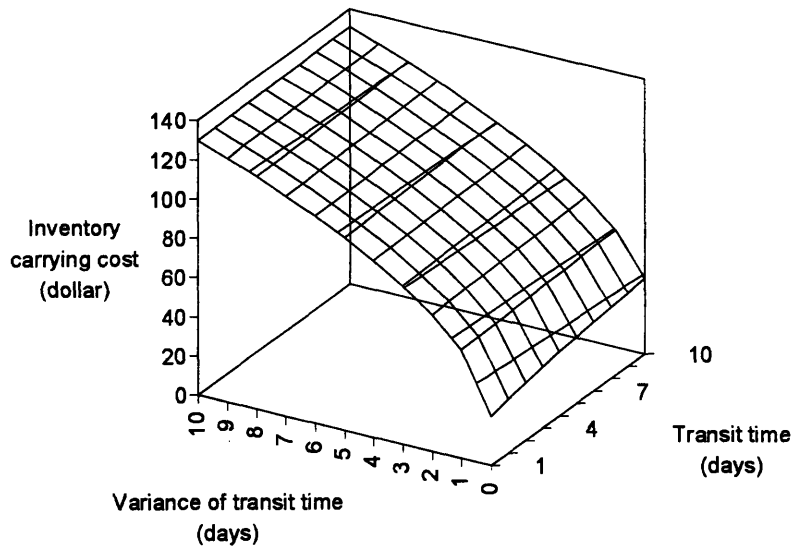


Figure 2.8 : Per unit inventory carrying cost under different transit time distribution

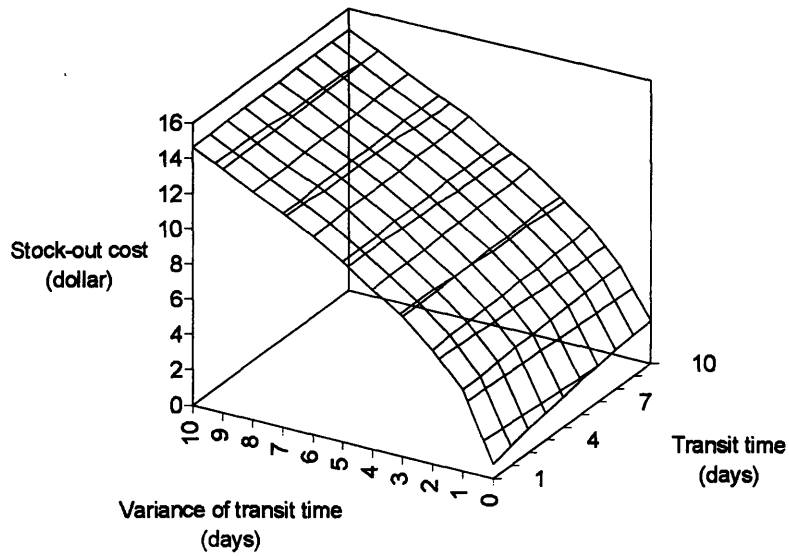


Figure 2.9 : Per unit stock-out cost under different transit time distribution

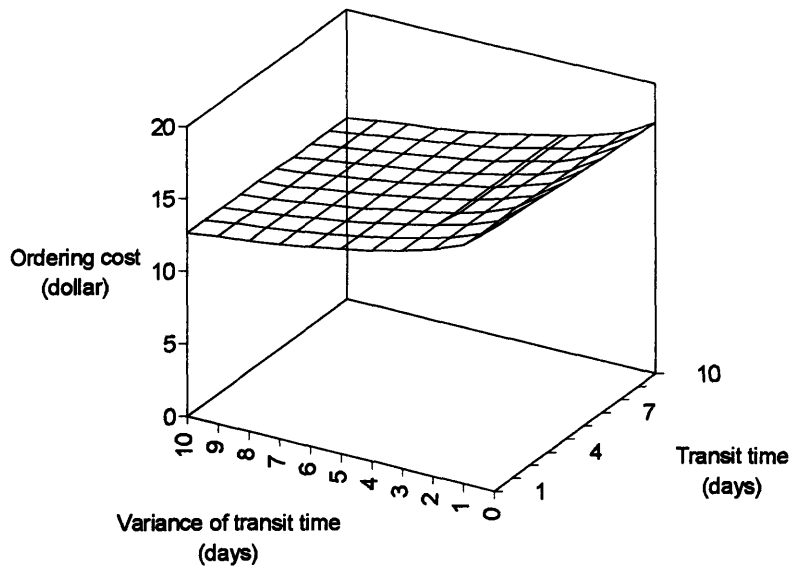


Figure 2.10 : Per unit ordering cost under different transit time distribution

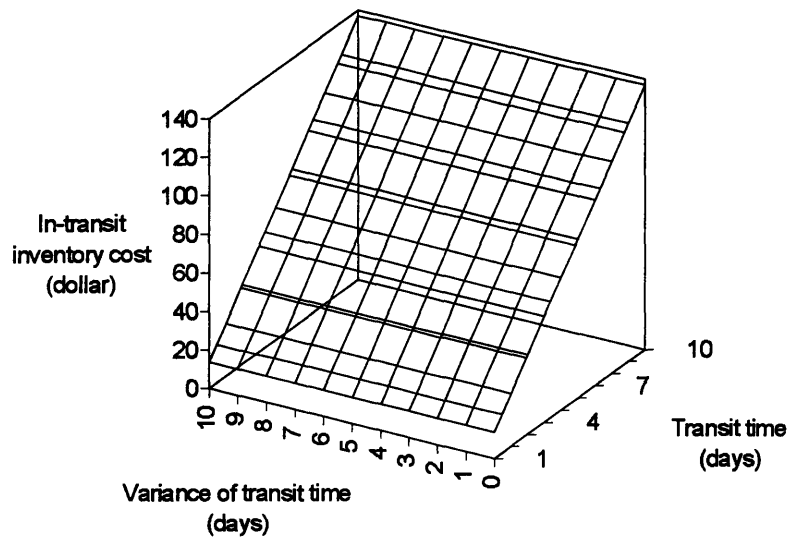


Figure 2.11 : Per unit in-transit inventory cost under different transit time distribution

Using the same analytical framework of total logistics cost analysis, we can also examine the relative importance of transit time and reliability to shippers' different characteristics.

Consider another shipper B that has different commodity values \$100,000 per car-load, assuming other characteristics are the same as before. Table 2.3 summarizes per car logistics cost under different mean and variance of transit time for this shipper. Figure 2.12 and Figure 2.13 show the difference in per unit logistics cost increase among shippers with different commodity values, as transit and variance of transit time increases.

It is clear that the effect of transit time and reliability on the logistics cost of a shipper with higher commodity value is more significant than on a shipper with lower value of commodity. It suggests that different shippers put different values on the improvement of transit time and reliability. This is one of the reasons why railroads need to identify different groups of shippers who have different service requirements and to provide desired levels of service to each group of shippers.

Table 2.3 : Per unit logistics cost under different transit time distribution
(value of commodity = \$100,000 per car-load)

E[L]	Var [L]										
	0	1	2	3	4	5	6	7	8	9	10
1	93.6	160.6	192.9	217.5	238.3	256.5	272.9	287.9	301.8	314.9	327.2
2	126.8	189.2	221.1	245.6	266.3	284.4	300.7	315.7	329.6	342.7	354.9
3	158.6	217.7	249.3	273.7	294.2	312.3	328.6	343.6	357.4	370.4	382.7
4	189.7	246.2	277.5	301.7	322.2	340.2	356.5	371.4	385.2	398.2	410.5
5	220.4	274.7	305.7	329.8	350.2	368.1	384.3	399.2	413.0	426.0	438.2
6	250.7	303.2	333.9	357.8	378.1	396.0	412.2	427.0	440.8	453.8	466.0
7	280.8	331.7	362.1	385.9	406.1	423.9	440.1	454.9	468.6	481.5	493.7
8	310.7	360.2	390.3	413.9	434.1	451.8	467.9	482.7	496.4	509.3	521.5
9	340.5	388.6	418.4	442.0	462.0	479.7	495.8	510.5	524.2	537.1	549.2
10	370.1	417.1	446.6	470.0	490.0	507.6	523.6	538.3	552.0	564.9	577.0

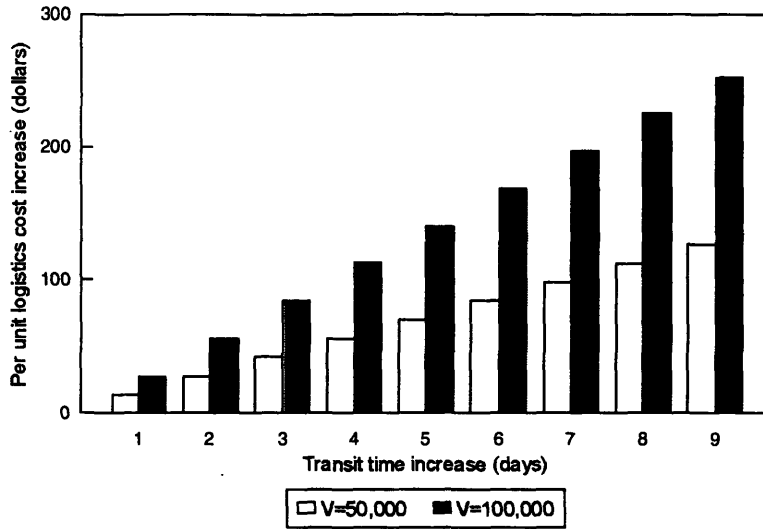


Figure 2.12 : Transit time and logistics cost change with different commodity value (base case : transit time = 1 days, variance of transit time = 3 days)

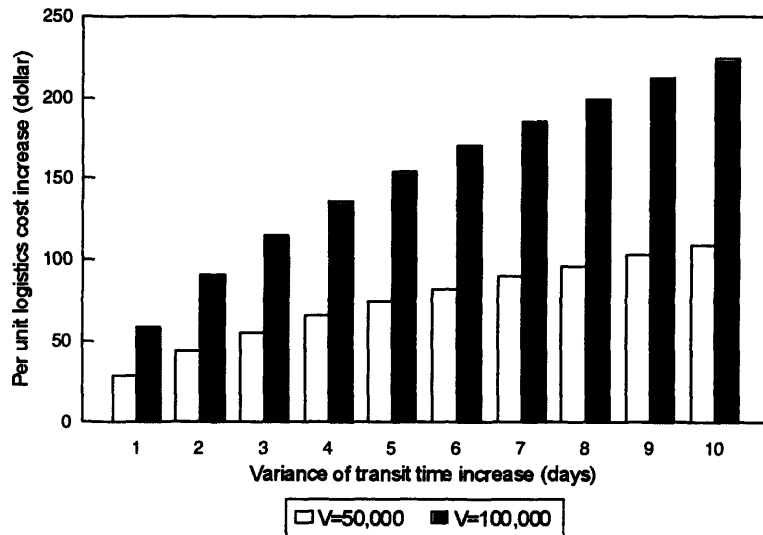


Figure 2.13 : Reliability and logistics cost change with different commodity value (base case : transit time = 3 days, variance of transit time = 0 days)

Consider another shipper C that has commodity value \$50,000 per car-load and stock-out cost 24%, assuming other characteristics are the same as before. Table 2.4 summarizes per car

logistics cost under different mean and variance of transit time for this shipper. Figure 2.14 and Figure 2.15 show the difference in per unit logistics cost increase among shippers with different stock-out cost, as transit and variance of transit time increases. The effect of reliability on the logistics cost of a shipper with higher stock-out cost is more significant than on a shipper with lower stock-out cost.

Table 2.4 : Per unit logistics cost under different transit time distribution
(value of commodity = \$50,000 per car-load, stock-out cost = 24%)

E[L]	<u>Var [L]</u>										
	<u>0</u>	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	<u>6</u>	<u>7</u>	<u>8</u>	<u>9</u>	<u>10</u>
1	57.9	92.8	109.6	122.5	133.4	143.0	151.6	159.5	166.9	173.8	180.3
2	74.6	107.1	123.7	136.6	147.4	156.9	165.5	173.4	180.8	187.7	194.2
3	90.6	121.3	137.8	150.6	161.4	170.9	179.5	187.4	194.7	201.6	208.1
4	106.2	135.6	152.0	164.7	175.4	184.9	193.4	201.3	208.6	215.5	221.9
5	121.6	149.9	166.1	178.7	189.4	198.8	207.4	215.2	222.5	229.4	235.8
6	136.8	164.2	180.2	192.7	203.4	212.8	221.3	229.1	236.4	243.3	249.7
7	151.9	178.4	194.3	206.8	217.4	226.8	235.3	243.1	250.3	257.2	263.6
8	167.0	192.7	208.4	220.8	231.4	240.7	249.2	257.0	264.2	271.1	277.5
9	181.9	207.0	222.5	234.9	245.4	254.7	263.1	270.9	278.2	285.0	291.4
10	196.7	221.2	236.6	248.9	259.4	268.7	277.1	284.8	292.1	298.9	305.3

The analysis shows that the total logistics cost increases as the transit time and/or variance of transit time increases. It also indicates that different shippers put different values on the improvement of transit time and reliability. This brings us back to the issue of market segmentation to identify shippers with different characteristics and the issue of service differentiation to provide more tailored service to different shippers that put different values on various service parameters. When a railroad has a limited physical and operational resources, it needs to consider how to allocate such resources and how to provide different levels of service for heterogeneous shippers that place different values on service improvement.

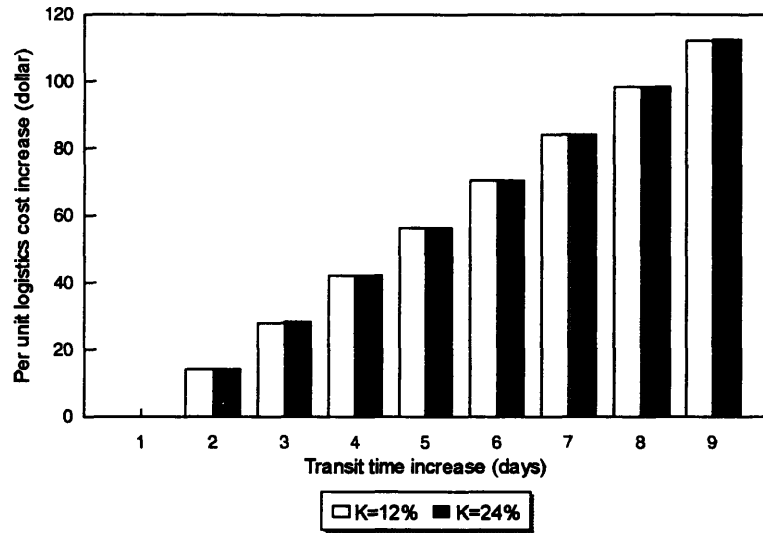


Figure 2.14 : Transit time and logistics cost change with different stock-out cost (base case : transit time = 1 days, variance of transit time = 3 days)

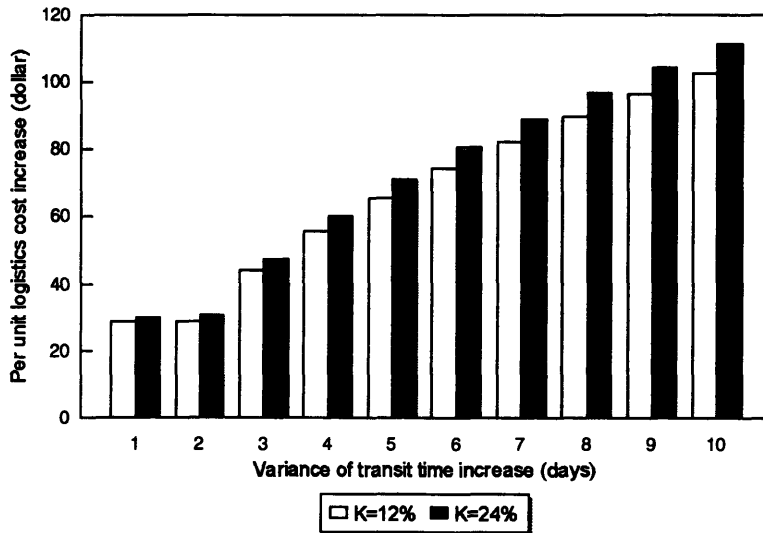


Figure 2.15 : Reliability and logistics cost change with different stock-out cost (base case : transit time = 3 days, variance of transit time = 0 days)

Shippers may use this analytical framework in evaluating alternative transportation modes and carriers. Carriers may use this analysis in evaluating the shipper's potential willingness to pay

for the service improvement (faster or more reliable service). Although the transit time and reliability are mainly focused in this analysis, it should be noted that other elements of service quality (e.g., price, loss and damage, other customer service) may also be important depending upon the characteristics of shippers. In addition, a shipper who is in control of transportation may use this analysis to find the joint optimal transportation and inventory policies (Allen, et al. [1985]).

In the next section, recent studies on freight market segmentation are reviewed and the importance of service differentiation is highlighted.

2.4 Freight Market Segmentation and Service Differentiation

Recent studies have shown that the freight market is divided into a number of market segments that have different elements of service quality and willingness to pay for additional service improvement (McGinnis [1978], Vieira [1992], and Smith and Resor [1991]). Table 2.5 shows the estimated demand elasticities⁵ to service quality (price, transit time and reliability)⁶ from the recent shipper survey by the John Morton Company. It shows that shippers of different commodity groups have different preferences on various service parameters.

⁵ Demand elasticity is a measure of the percentage change in demand brought about by a 1 percent change in some other variable. For example, if reliability elasticity is 6.0, a 1 percent improvement in reliability causes demand to increase by 6 percent. If price elasticity is -1.1, a 1 percent rise in price causes demand to fall by 1.1 percent. Service-price cross elasticity is a measure on how much can price be increased without losing current demand for each 1 percent improvement in service. For example, if reliability-price cross elasticity is 5.5, a 1 percent improvement in reliability allows 5.5 percent increase in price without losing current demand.

⁶ Reliability is measured by the percent of time a loaded car arrives at the customer's dock within the time window desired by the customer. Transit time is measure by the time required for the shipment to move from the shipper's dock to the consignee's dock.

Table 2.5 : Estimation of demand elasticities to service quality for shippers of different commodity groups

	<u>Service elasticity</u>		<u>Price elasticity</u>	<u>Service-price cross elasticity</u>	
	<u>Reliability</u>	<u>Transit time</u>		<u>Reliability</u>	<u>Transit time</u>
Paper	6.0	-1.1	-1.1	5.5	-1.0
Pet Food	6.9	-1.4	-1.5	4.6	-0.9
Aluminum	4.3	-1.3	-1.3	3.3	-1.0
Plastics	4.7	-0.9	-1.6	2.9	-0.6
Tires	6.2	-1.6	-0.9	6.9	-1.8
Average	5.3	-1.2	-1.3	4.1	-0.9

Source : Smith and Resor [1991]

These studies found that shippers' attitudes towards service quality vary significantly. It was also found that within the same group of commodities, it was possible to identify groups of shippers who have service quality needs that are different from other shippers. These studies suggest that there is a heterogeneity in shippers and their traffic. These studies conclude that one service design and marketing strategy will not satisfy all the current and potential shippers. It suggests that there is a potential for carriers to differentiate their service to meet shippers' unsatisfied needs. A carrier must target markets in which to compete and must properly position its services to be competitive in those markets to gain profitable market share. The ability of a carrier to differentiate its services can provide a strategic advantage in the competitive transportation environment.

Market segmentation and service differentiation can be helpful in dealing both with gaining market and improving service quality. Carriers can match their operating capabilities with different needs of shippers in current and potential market segments. In order to match its operating capability with the different needs of shippers in different market segments, a carrier

needs to put its efforts into identifying homogeneous segments of shippers, and improving its ability to design and produce differentiated service products for different market segments.

Service differentiation in general can be defined as a strategy to provide the market-sensitive and tailored service to different shippers who have different service requirements. The underlying concept is that the marketing department can identify different classes of shippers, who have different requirements with respect to cost, transit time and reliability. The operating department can then adjust its operations to provide the desired level of service to each class of shippers, with cost level that are appropriate to those services (i.e., higher cost for higher quality service and lower cost for lower quality service).

Railroad have provided different train services for different classes of freight traffic, e.g., general merchandise train service, bulk unit train service, and intermodal train service. Some railroads have tried to further differentiate services for important shippers. They have done this by giving priorities to certain trains or blocks and by taking special action to make sure that train connections are made and that local deliveries take place on time. For example, CSXT have identified important merchandise O-D flows. To improve service reliability of those traffic, they select 60 merchandise trains as priority “Q-trains” and these trains have priority in getting power, crew and track time. They also take special actions to improve service reliability for these traffic and continuously monitor their service performance (CSXT [1992]).

In this chapter, the importance of incorporating the heterogeneity of shippers and their traffic to the development of a railroad’s service strategy was discussed based on the logistics cost analysis and the findings of recent studies on freight market segmentation, and the needs of market segmentation and service differentiation were discussed. In the next chapter, we examine whether and how railroads are currently differentiating service among different groups of freight traffic.

CHAPTER 3

Empirical Analysis of Service Differentiation in Rail Freight Transportation

3.1 Introduction

Concepts and implications of market segmentation and service differentiation in the context of the rail industry were discussed in Chapter 2. Considering the different service expectations of shippers and their potential willingness to pay for service improvements, it was shown that there is a clear need to differentiate services to different shippers, with cost levels that are appropriate to those services.

In this chapter, we examine how railroads are currently differentiating services among different groups of freight traffic. Railroads already practice some level of service differentiation for broad classes of freight traffic, dividing it into at least three major types: general merchandise train service, unit train service and intermodal train service. An empirical analysis is performed of trip time and reliability of car movements for the three major train services. For each category of train service, a number of different kinds of car equipment can be utilized depending upon the characteristics of the shipments and loading requirements. In this context, car cycle information

for the three car types was collected: box car data for general merchandise train service, covered hopper car data for unit train service, and double-stack car data for intermodal train service. The trip time and various reliability measures are evaluated and compared for different train services.

There have been many empirical studies on the reliability of rail service, but most of these studies analyzed a limited number of O-D pairs (Martland [1972, 1974], Martland, et al. [1981]). To our knowledge, the study described here is the first large-scale systematic assessment of actual O-D trip times and reliability of rail freight service in North American railroads. Some discussion of the causes of unreliability is made with the results found in other studies. Analysis of other elements of service quality (e.g., costs) is not attempted.

3.1.1 Data source

The data was provided through the Association of American Railroads Car Cycle Analysis System (CCAS), which is designed primarily for the analysis of car cycle time. A car cycle begins when a car is placed empty for loading and ends when it is again placed empty for loading. The car cycle time is composed of four basic components: shipper time (i.e., loading time), total loaded time, consignee time (i.e., unloading time) and total empty time. The shipper time begins when a car is placed empty at the shipper's siding and ends when it is released with a load. The loaded time extends from release until placement loaded at the consignee's siding. The consignee time is the time from the placement loaded until the car is unloaded and released to the railroad. The empty time is the time from released empty until it is again placed empty for the next shipper. The empty time can be divided into the empty trip time and the empty terminal time.

The car cycle time is the sum of shipper time, loaded time, consignee time and empty time.

Components of the car cycle are illustrated in Figure 3.1.

Among the components of car cycle time, our primary interest is on the analysis of loaded time and its associated reliability, since the loaded time is directly perceived as service performance by shippers.

In addition, each record in the CCAS includes the Standard Point Location Codes (SPLC) information. For the box car data, each record also includes the Standard Transportation Commodity Codes (STCC) information.

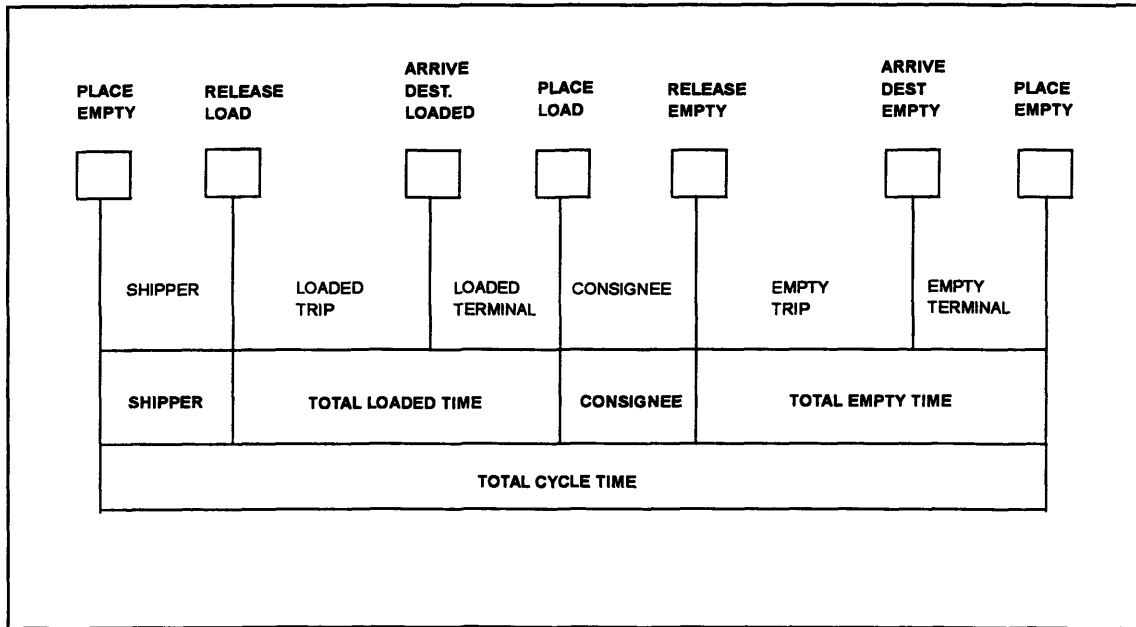


Figure 3.1 : Components of the car cycle

For each type of traffic, car cycle data collection was done in two steps. First, a 10% sample of cars was randomly selected from the Universal Machine Language Equipment Register

(UMLER¹). Second, all of the car cycle records for the selected cars were extracted from the CCAS. The data collection period and the number of car cycle records for each type of traffic are summarized in Table 3.1. Note that only car cycle records with complete information on loaded time were selected.

Table 3.1 : Data source

<u>Car type</u>	<u>Number of records</u>	<u>Period</u>
Box car	252,619	Dec. 1, 1989 - Nov. 30, 1990
Covered hopper car	351,024	Dec. 1, 1990 - Nov. 30, 1991
Double-stack car	23,026	Dec. 1, 1990 - Nov. 30, 1991

Box car data includes both equipped (car type code A) and unequipped (code B) cars. Covered hopper data includes car types C113 (93.2%), C214 and C514.

3.1.2 Selection of O-D pairs

O-D pairs were defined using the 6-digit SPLC, which identifies locations at the station level. The analysis by shipper may be more desirable to take into account of the characteristics of shippers, but shipper information was not available. For the box car data, however, we usually found only one commodity group for each O-D pair, which suggests that most O-D pairs corresponded to movements from one shipper to a single consignee.

¹ A computerized listing of all the cars approved for use in interchange service.

For each car type, trip time and reliability were evaluated for highest volume movements. For the box car and double-stack car, we selected O-D pairs that had more than 30 car moves during the one year period².

The covered hopper data included cars moving in general merchandise trains as well as unit train service. Because we intended to analyze the performance of bulk unit train service using this data, we further processed the data to select cars moving in unit train service. We defined a group of car moves that have the same origin, destination, origin railroad, destination railroad, departure date from origin, and arrival date at destination as a single shipment. We assumed that shipments having at least 4 car movements were unit train moves³, while shipments having less than 4 car movement records likely moved in carload or multi-carload train service. For unit train moves, we identified the total number of shipments over the year for each “service lane”, i.e., for each combination of origin, destination, origin railroad and destination railroad, and considered this to be the number of unit train operations during the year. We found that a large number of service lanes had only a few unit train operations. We selected service lanes that had at least 10 train operations a year to represent regular unit train service. We did not consider the other shipments, which we assumed represented special or non-regular unit train service.

Using these procedures, we selected 477 O-D pairs for box car traffic,⁴ 102 O-D pairs for covered hopper car traffic using unit train service, and 93 O-D pairs for double-stack car traffic.⁵ The selection of O-D pairs and the records sampled for each car type are summarized in Table 3.2.

² Since we used 10 percent random sample, this corresponds to approximately 300 moves a year or nearly 1 move a day.

³ Since the data is a 10 percent random sample of the hopper cars, 4 movement records represented approximately 40 car movements. These shipments are likely to represent unit train moves.

⁴ The O-D selection procedure for the box car was different from the previous analysis (Little, et al., 1992) using the same data set. In the earlier study, the 100 O-D pairs with the highest volume (number of car movements) were analyzed along with a random sample of low volume O-D pairs.

⁵ The 10 corridors (20 O-D pairs) with the highest volume were analyzed (Wang, 1993).

Table 3.2 : Selection of O-D pairs

<u>Car type</u>	<u>Selected O-D</u>	<u>Number of records</u>	<u>Moves/O-D</u>
Box car	477	29,120 (11.53%)	61.0
Covered hopper car	102	11,115 (3.17%)	109.0
Double-stack car	20	10,486 (45.54%)	524.3

() is percent of the sample records for selected O-D pairs

3.1.3 Trip time and reliability measures

We used the following trip time and reliability measures to indicate the performance for the selected O-D pairs.

- Mean trip time
- Standard deviation of trip time
- n-day-percent about mean
- Maximum n-day-percent

The mean and standard deviation can be used to characterize the distribution of trip time. The existence of very long trip times will limit the usefulness of the mean as a measure of central tendency and of the standard deviation as a measure of the compactness of trip time distribution. Therefore, two additional measures of trip time reliability were used. The n-day-percent is a measure calculated from the trip time distribution. The n-day-percent centered about the mean measures the percentage of the cars that arrive within a time window that begins $n/2$ days before

the mean trip time and ends $n/2$ days after the mean trip time. However, since trip time distributions are often skewed to the right, it is often possible to obtain a higher percentage by using a window not centered about the mean. The maximum n -day-percent is the maximum percentage of cars that arrive at the destination within any n -day period. For example, the maximum 2-day-percent measures the largest percent of the cars that can arrive in any 48-hour time window. This measure is independent of predetermined schedules, relatively insensitive to excessive data values or data errors, and not highly related to the mean value.

Consider an example of O-D trip time distribution (Figure 3.2). The mean trip time is 5.0 days and the standard deviation of trip time is 1.7 days. The 3-day-percent about the mean (from day 4 to day 6) is 59.6%. The maximum 3-day-percent (from day 3 to day 5) is 60.6%.

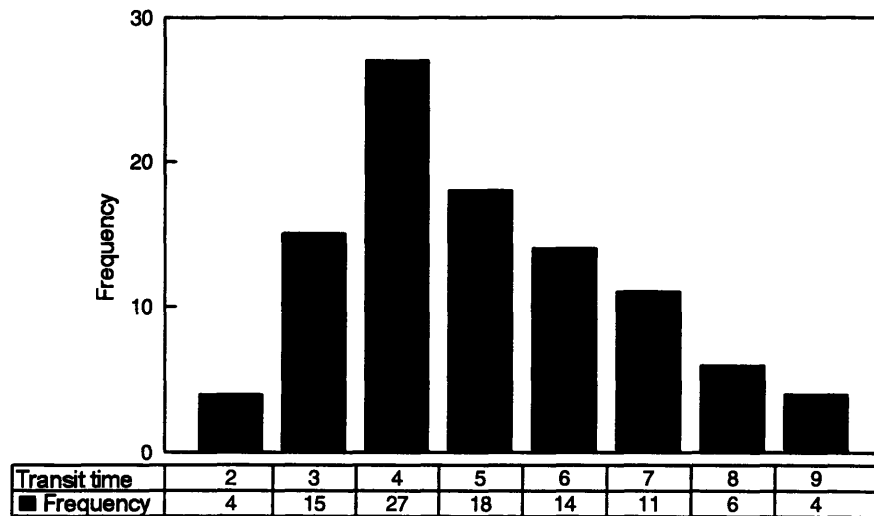


Figure 3.2 : Example of O-D trip time distribution

Shippers are also concerned with performance relative to schedules (or customer commitments). Because car schedule information was not available, we could not analyze

performance relative to schedules. More detailed discussions on the performance measures for trip time in the rail transportation context can be found in Martland [1972].

3.2 Trip Time and Reliability of Box Car Traffic

Car cycle time analysis

Components of the car cycle were analyzed for the entire sample of box cars. We only used “perfect” records that had all the components of the car cycle. The average car cycle time for box cars was 26.9 days. The average loaded time was 8.8 days and the average empty time was 14.5 days. The empty time was much higher than the loaded time largely because there was a surplus of box cars during 1990.

Table 3.3 : Car cycle time : box car service

No. of moves ⁶	48,129
Shipper time	2.15 days
Loaded time	8.77
Consignee time	1.48
Empty time	14.48
Total cycle time	26.88

Car movements can be classified into local and interline moves. Local movements are handled by a single railroad. Interline movements are handled by two or more railroads. The

⁶ It represents the number of "perfect" records (19.05% of 252,619 records).

analysis of the car cycle showed that local movements (23.3 days) had shorter car cycles than the interline movements (28.7 days). Local movements had shorter loaded and empty times than interline movements.

Table 3.4 : Car cycle time of local and interline moves :
box car service

	<u>Local</u>	<u>Interline</u>
No. of moves	16,382	31,747
Shipper time	2.12 days	2.16 days
Loaded time	6.78	9.81
Consignee time	1.49	1.47
Empty time	12.95	15.27
Total cycle time	23.33	28.71

Trip time and reliability analysis

Trip time and reliability were analyzed for the highest volume O-D pairs. Table 3.5 summarizes the aggregate trip time and reliability performance of the selected O-D pairs.⁷ The distance of a typical O-D pair of box car traffic was 788 miles and the average loaded time was 7.2 days. A typical box car spent less than 2 days moving loaded in a train⁸ and the majority of time was therefore spent in other activities, presumably in terminals. The overall reliability level of box car traffic was very low. The average maximum 2-day-percent of box car traffic was 48.6%, which means that only half of the box car traffic arrived within a 2 day window for the typical O-D pair.

⁷ The trip time and reliability measures for each O-D are computed. Then, the aggregate performance is computed as the average of each performance measure for all O-D pairs.

⁸ Assuming a moderate merchandise train speed 20 miles/hour.

Table 3.5 : Aggregate trip time and reliability performance of selected O-D pairs : box car service

Number of O-D	477
Number of moves ⁹	61.0 moves
Number of railroads	2.11 railroads
Distance	788.1 miles
Mean trip time	7.16 days
Std dev of trip time	2.62 days
1-day-% about mean	19.86 %
2-day-% about mean	36.79 %
3-day-% about mean	50.25 %
Maximum 1-day-%	32.42 %
Maximum 2-day-%	48.56 %
Maximum 3-day-%	61.07 %

To examine any meaningful relationship among trip time, reliability performance and other characteristics of O-D car movements (e.g., number of car moves, number of participating railroads and distance), we analyzed the correlation coefficients¹⁰ between variables.

Table 3.6 shows that the number of car moves (i.e., annual shipment volume), number of participating railroads (i.e., number of interchange operations) and distance had significant linear

⁹ This represents approximately 10 percent of the annual shipment size.

¹⁰ The correlation coefficient is a scale-free measure of the linear association between two variables. Let X and Y be any two variables. The correlation coefficient of X and Y is denoted $\rho(X, Y)$ and is given by

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}$$

where, $\text{Cov}(X, Y)$ is the covariance of X and Y, σ_X is the standard deviation of X and σ_Y is the standard deviation of Y.

relationship with trip time. It means that O-D pairs that have longer distance, larger number of interchange operations or smaller volumes tend to have longer trip times.

It also shows that a larger number of interchange operations and longer distances had significant linear relationship with reliability (measured as maximum 2-day-percent).

The correlation between the number of moves and reliability was not significant. This result is consistent with the results of a previous analysis which is summarized in Table 3.7. The comparison of trip time and reliability between large and small O-D pairs showed that large O-D pairs clearly had shorter trip times than small O-D pairs. However, the difference in reliability was not clear (Little, et al. [1992]).

Table 3.6 : Correlation coefficients between variables :
box car service

	<u>Mean</u>	<u>Std dev</u>	<u>2-day-% mean</u>	<u>Max. 2-day-%</u>
No. of moves	-0.22233 (0.0001)	-0.07619 (0.0965)	0.08971 (0.0502)	0.06029 (0.1758)
No. of railroads	0.45653 (0.0001)	0.18144 (0.0001)	-0.24192 (0.0001)	-0.30515 (0.0001)
Distance	0.63421 (0.0001)	-0.08251 (0.1408)	-0.04399 (0.4330)	-0.15649 (0.0050)
Mean trip time	-	0.66655 (0.0001)	-0.55654 (0.0001)	-0.61875 (0.0001)

() is the probability that a null hypothesis $H_0 : \rho = 0$ can be rejected

Table 3.7 : Trip time and reliability between large and small O-D pairs

	<u>Large O-D pairs</u>		<u>Small O-D pairs</u>	
	<u>Local</u>	<u>Interline</u>	<u>Local</u>	<u>Interline</u>
Total number of moves	5,526 moves	6,593	247	398
Distance	n/a	n/a	n/a	n/a
Mean trip time	4.95 days	5.65	5.27	8.18
Std dev of trip time	1.69 days	2.14	1.58	2.81
Maximum 2-day-%	59.5 %	45.4	54.8	42.1

Source : Little, et. al. [1992]

It was also found that the correlation between the reliability and the mean trip time was highly significant (see Table 3.6). Figure 3.3 plotted the mean trip time and maximum 2-day-percent of all O-D pairs. Typically, railroad analysts assert that long trip times are acceptable to shippers if the reliability is good. However, we could not find any distinct cluster of O-D pairs that had both long trip time and good reliability. Figure 3.4 plotted the time spent not moving loaded in a train but in other activities and maximum 2-day-percent of all O-D pairs.¹¹ It clearly shows that reliability deteriorated as the time spent in other activities increased. It suggests that the reliability of car movements can be improved by reducing the time spent in those other activities or by making those other activities more reliable.

This assertion is supported by other previous studies. Previous studies on O-D trip time performance showed that the majority of trip time was spent in terminals (Lang and Martland [1972]). A recent study on the causes of unreliable service, based upon the data of a major railroad, showed that terminal and train delays accounted for more than 40 percent of the delays to shipments (Little and Martland [1993]). This study concluded that unreliable service is more

¹¹ Assuming a moderate merchandise train speed 20 miles/hour.

closely related to the management of resources (terminal management, train management, and power distribution) than to deficiencies in the technology or hardware of railroading.

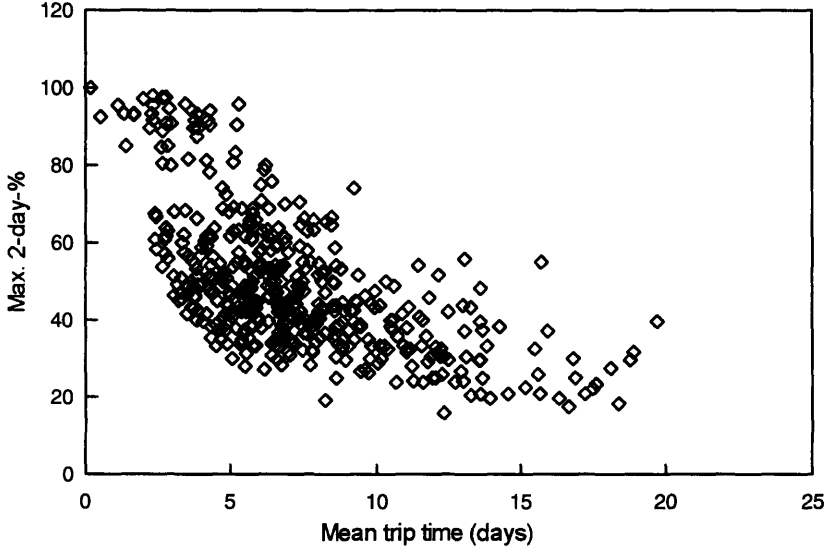


Figure 3.3 : Relation between mean trip time and reliability :
box car service

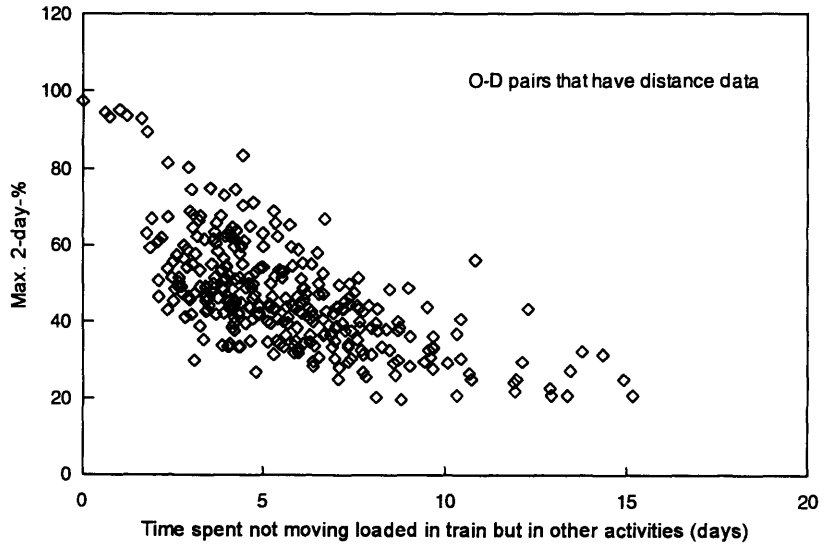


Figure 3.4 : Relation between time spent in other activities and reliability

Figure 3.5 shows the distribution of O-D pairs in terms of the maximum 2-day-percent. It indicates that there is significant variability of performance among different O-D pairs.

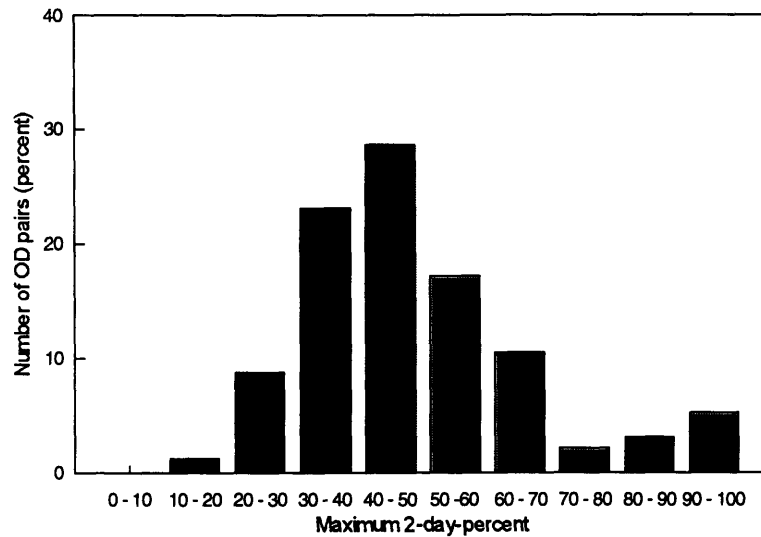


Figure 3.5 : Distribution of O-D pairs among different ranges of reliability performance : box car service

We further identified and analyzed reliability performance for the combination of O-D and commodity¹². Table 3.8 summarizes the distribution of O-D and commodity groups among different ranges of reliability performance. These results show there is a certain degree of service differentiation at the level of commodity groups and individual shippers. It also shows that there is significant variability of performance among O-D pairs even in the same commodity group.

It was not clear, however, if such differentiated service levels are the result of intentional efforts to differentiate service considering service requirements of individual O-D pairs, or if they simply reflect the results of day-to-day operations reacting to the daily traffic variability and the

¹² We analyzed the trip time and reliability performance for 433 combinations of O-D and commodity group that have more than 10 moves. We may presume that each combination of O-D and commodity group corresponds to a single shipper. The commodity group was determined by the 2-digit STCC.

uncertainty in various stages of operations. Due to the aggregate nature of the data, we could not further analyze the reasons or causes of this differentiation in service.

Table 3.8 : Distribution of O-D and commodity groups among different ranges of reliability performance

<u>Commodity</u>	<u>Maximum 2-day-percent</u>				
	<u>0-20</u>	<u>20-40</u>	<u>40-60</u>	<u>60-80</u>	<u>80-100</u>
Farm products	-	6 (54.5)	3 (27.3)	-	2 (18.2)
Food or kindred products	-	9 (60.0)	4 (26.7)	2 (13.3)	
Lumber or wood products	-	3 (23.1)	6 (46.2)	1 (7.7)	3 (23.0)
Pulp and paper	1 (0.8)	45(34.4)	66(50.4)	13(9.9)	6 (4.5)
Chemicals	-	2 (40.0)	1 (20.0)	2 (40.0)	-
Rubber or plastic products	-	1 (12.5)	5 (62.5)	2 (25.0)	-
Clay, concrete, glass, stone	-	2 (20.0)	7 (70.0)	1 (10.0)	-
Primary metal products	-	1 (12.5)	3 (37.5)	4 (50.0)	-
Electrical machinery	1 (7.1)	4 (28.6)	4 (28.6)	4 (28.6)	1 (7.1)
Transportation equipment	-	25(17.4)	85(59.0)	31(21.5)	3 (2.1)
Waste and scrap	-	2 (33.3)	4 (66.7)	-	-
Hazardous materials	-	-	-	3 (16.7)	15(83.3)

() is the percent

3.3 Trip time and Reliability of Covered Hopper Car Traffic

Car cycle time analysis

Components of the car cycle were analyzed for covered hopper cars moving in unit train service.

The average car cycle time was 15.3 days. The average loaded time was 5.3 days and the average

empty time was 6.8 days. Covered hopper cars had a much shorter car cycle than box cars.

Furthermore, all time components of the covered hopper car were shorter than those of the box car.

Table 3.9 : Car cycle time : covered hopper car service

No. of moves ¹³	6,799
Shipper time	1.92 days
Loaded time	5.33
Consignee time	1.27
Empty time	6.76
Total cycle time	15.27

The analysis showed that local movements (14.8 days) had shorter car cycles than interline movements (17.2 days). Local movements were shorter in each component of the car cycle except shipper time.

Table 3.10 : Car cycle time of local and interline moves : covered hopper car service

	<u>Local</u>	<u>Interline</u>
No. of moves	5,397	1,402
Shipper time	2.04 days	1.46 days
Loaded time	5.19	5.85
Consignee time	1.19	1.57
Empty time	6.35	8.34
Total cycle time	14.77	17.23

¹³ The number of "perfect" records (61.17% of 11,115 records).

Trip time and reliability analysis

Trip time and reliability were analyzed for the selected O-D pairs (Table 3.11). The distance of a typical O-D pair of covered hopper traffic was 831 miles and the average trip time was 5.2 days. The reliability of covered hopper car moves was higher than the box car moves. The average maximum 2-day-percent of covered hopper cars was 60.9%. Because cars moved by unit trains do not go through any classification process in intermediate terminals, most of trip time is spent in line activities and in the origin and destination terminals. Trip time and reliability performance is therefore closely related with how a railroad prioritizes the unit train operation in line (e.g., priority in track assignment) and terminal activities (e.g., power and crew assignments). Some railroads also hold groups of 40 or more cars at a terminal for several days until they can be combined with similar groups to form a unit train.

Table 3.11 : Aggregate trip time and reliability performance :
covered hopper car service

Number of O-D	102
Number of moves	108.9 moves
Number of railroads	1.47
Distance	831.0 miles
Mean trip time	5.25 days
Std dev of trip time	2.04 days
1-day-% about mean	23.35 %
2-day-% about mean	46.41 %
3-day-% about mean	62.57 %
Maximum 1-day-%	41.90 %
Maximum 2-day-%	60.95 %
Maximum 3-day-%	73.21 %

We analyzed the correlation coefficients between variables (Table 3.12). The correlation analysis showed that the number of railroads and distance had a significant linear relationship with the mean trip time. The results showed that the reliability (maximum 2-day-percent) deteriorated for O-D pairs with longer distance and mean trip time.

The correlation between the number of car moves or the number of participating railroads and reliability were not significant. It was also found that the correlation between the reliability and the mean trip time was highly significant (Figure 3.6). The covered hopper car service has a stronger linear relationship between the reliability and the mean trip time ($\rho = -0.77$) than the box car service ($\rho = -0.62$).

Table 3.12 : Correlation coefficients between explanatory variables and reliability : covered hopper car service

	<u>Mean</u>	<u>Std dev</u>	<u>2-day-% mean</u>	<u>Max. 2-day-%</u>
No. of moves	-0.01559 (0.8764)	0.03673 (0.7140)	-0.01632 (0.8707)	-0.10494 (0.2939)
No. of railroads	0.21154 (0.0328)	-0.08271 (0.4085)	-0.05291 (0.5974)	-0.09068 (0.3647)
Distance	0.61274 (0.0001)	0.17274 (0.1139)	-0.37354 (0.0004)	-0.35955 (0.0007)
Mean trip time	-	0.73639 (0.0001)	-0.70767 (0.0001)	-0.77366 (0.0001)

() is the probability that a null hypothesis $H_0 : \rho = 0$ can be rejected

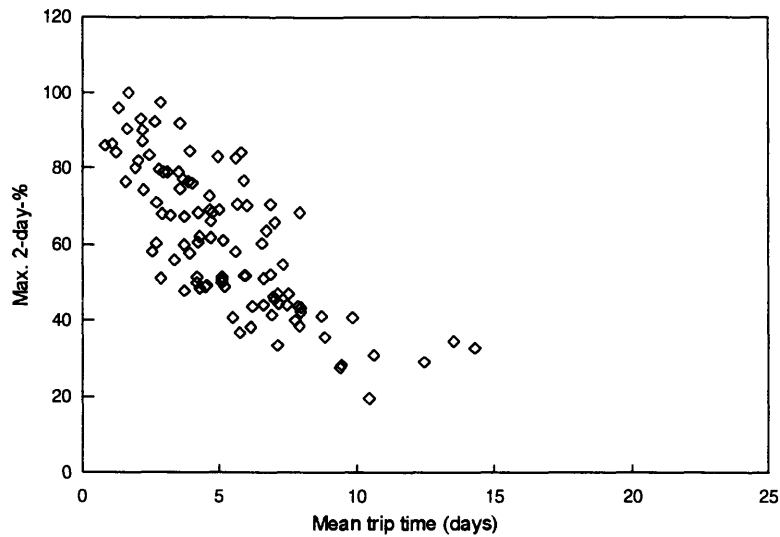


Figure 3.6 : Relation between mean trip time and reliability : covered hopper car service

The analysis showed that there was significant variability of performance among different O-D pairs. Figure 3.7 shows the distribution of O-D pairs of covered hopper car traffic among different ranges of maximum 2-day-percent.

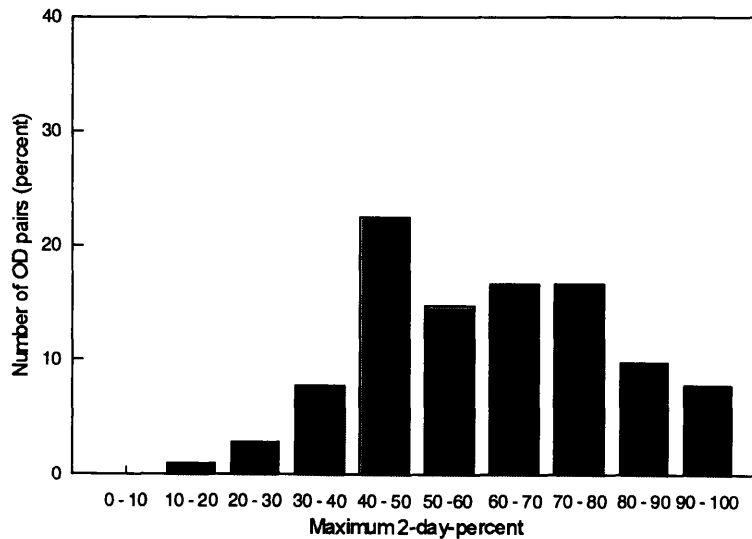


Figure 3.7 : Distribution of O-D pairs among different ranges of reliability performance : covered hopper car service

3.4 Trip Time and Reliability of Double-Stack Car Traffic

Car cycle time analysis

Components of the car cycle were analyzed for the entire sample of double-stack cars. The average car cycle time of double-stack cars was 6.1 days. The average loaded time was 3.2 days and the average empty time was 2.0 days. More than half of double-stack car moves (51.2%) had less than one day of empty time. The double-stack car cycle was less than half of the covered hopper car cycle and only a third of the box car cycle. For double-stack car movement, the empty time was shorter than the loaded time. For this traffic, the empty time is usually incurred with the terminal area, as the double-stack cars are generally reloaded rather than moved empty to another terminal.

Table 3.13 : Car cycle time : double-stack car service

No. of moves ¹⁴	2,573
Shipper time	0.73 days
Loaded time	3.21
Consignee time	0.22
Empty time	1.99
Total cycle time	6.15

Analysis of the car cycle showed that local movements (5.3 days) had shorter car cycles than interline movements (8.0 days).

¹⁴ The number of "perfect" records (11.17% of 23,026 records).

Table 3.14 : Car cycle time of local and interline moves :
double-stack car service

	<u>Local</u>	<u>Interline</u>
No. of moves	1,804	769
Shipper time	0.73 days	0.72 days
Loaded time	2.59	4.67
Consignee time	0.21	0.26
Empty time	1.82	2.38
Total cycle time	5.35	8.04

Trip time and reliability analysis

The trip time and reliability performance of double-stack car movements by unit train service were analyzed for each selected corridor (Table 3.15). The average loaded time was 2.5 days. Double-stack traffic had much faster service than box cars or covered hopper unit trains. The reliability of double-stack car service was also much higher than for the other services. The average maximum 1-day-percent of double-stack car traffic was 89.2%, which means that more than nine of ten double-stack cars consistently arrived within a 1-day window.

Table 3.15 : Aggregate trip time and reliability performance :
double-stack car service

Total number of moves	10,260 moves
Distance	n/a
Mean trip time	60.7 hours
Std dev of trip time	11.9 hours
Maximum 8-hour-%	62.4 %
Maximum 12-hour-%	74.2 %
Maximum 24-hour-%	89.2 %

Source : Wang [1993]

To examine the relationship between the characteristics of intermodal traffic movements and performance, double-stack car movement was classified into eastbound and westbound movements. It was also classified into long and short distance movements, with “long distance” defined as longer than 1,500 miles. The results showed that westbound movements had slightly shorter trip times than eastbound movements. There was no significant difference in reliability between the two directions. Short distance movements had higher reliability than long distance movements, which is consistent with the covered hopper unit train results.

Table 3.16 : Trip time and reliability performance by direction and distance :
double-stack car service

	<u>Eastbound</u>			<u>Westbound</u>		
	<u>Total</u>	<u>Long</u>	<u>Short</u>	<u>Total</u>	<u>Long</u>	<u>Short</u>
Total number of moves	4,387	3,721	666	5,873	4,287	1,586
Distance	n/a	n/a	n/a	n/a	n/a	n/a
Mean trip time	64.4 hr.	70.8	28.6	58.0	67.4	32.8
Std dev of trip time	11.2 hr.	10.9	12.7	12.4	10.7	16.9
Maximum 8-hour-%	61.2 %	60.2	66.7	63.3	59.3	74.1
Maximum 12-hour-%	72.8 %	72.2	76.7	75.1	70.4	88.1
Maximum 24-hour-%	89.4 %	88.7	93.3	89.0	86.3	96.5

Source : ibid.

Table 3.17 shows the trip time and reliability performance of three representative carriers¹⁵. It showed that the performance significantly varied among different carriers. The maximum 1-day-percent ranged from 39% to 99%. Although there was no significant difference in

¹⁵ Carrier K performed best among carriers that operated in long distance corridors. Carrier L performed best among carriers that operated in short distance corridors. Carrier M performed worst among all carriers of selected corridors.

the aggregate, we could observe significant differences in reliability between eastbound and westbound movement in corridor or carrier level. There was, however, no pattern that could be found regarding the difference of reliability between two directions. This also suggests that shippers can reasonably select among carriers on the basis of service level in terms of trip time and reliability.

Table 3.17 : Trip time and reliability of different carriers :
double-stack car service

Carrier Direction ¹⁶	K		L		M	
	E/B	W/B	E/B	W/B	E/B	W/B
Distance	n/a	n/a	n/a	n/a	n/a	n/a
Mean trip time	66.0 hr.	71.4	39.4	38.0	99.4	82.7
Std dev of trip time	4.9 hr.	16.4	7.6	7.7	25.1	22.4
Maximum 8-hour-%	66.7 %	56.3	46.3	63.5	16.6	22.1
Maximum 12-hour-%	83.9 %	65.6	60.6	73.9	23.0	32.2
Maximum 24-hour-%	98.9 %	86.5	86.9	93.2	38.8	58.9

Source : *ibid.*

3.5 Summary and Conclusions

3.5.1 Comparison among different train services

Table 3.18 compares the car cycle components of the three services. The average car cycle was about 1 week for the double-stack car, 2 weeks for the unit train covered hopper car, and 4 weeks

¹⁶ A corridor in each direction included one or more O-D pairs determined by the 6-digit SPLC.

for the box car and the non-unit train covered hopper car. The car cycle time was longer for box car traffic. Moreover, all components of car cycle time were longer for box car traffic than other types of traffic.

Table 3.18 : Car cycle time for different train services

	<u>Box car</u>	<u>Covered hopper</u>	<u>Double-stack</u>
Shipper time	2.15 days	1.92	0.73
Loaded time	8.77	5.33	3.21
Consignee time	1.48	1.27	0.22
Empty time	14.48	6.76	1.99
Total cycle time	26.88	15.27	6.15

The service provided to box car traffic was significantly slower and less reliable than the other types of traffic. The maximum 2-day-percent for a typical box car movement was just under 50%, which means that only half of the cars arrived at the destination within a 2-day-window. It is clear that there was substantial inconsistency in the level of service provided to general merchandise shippers. On the other hand, the service provided to double-stack cars was significantly faster and more reliable than the other two types of traffic. The maximum 1-day-percent for a typical double-stack car movement was 89.2%. Table 3.19 summarizes the trip time and reliability performance of the three train services. The differences in trip time and reliability between box car movements and covered hopper car movements were statistically significant at significance level 0.05.¹⁷

¹⁷ Due to the lack of information, the difference in trip time and reliability between box car service and double-stack car service or between covered hopper car service and double-stack car service were not statistically tested.

Table 3.19 : Trip time and reliability performance of different train services

	<u>Box car</u>	<u>Hopper car</u>	<u>Double-stack</u>
Distance	788.1 miles	831.0	n/a
Mean trip time	7.16 days	5.25	2.53
Std dev of trip time	2.62 days	2.04	0.50
Maximum 1-day-%	32.42 %	41.90	89.2
Maximum 2-day-%	48.56 %	60.95	n/a
Maximum 3-day-%	61.07 %	73.21	n/a

In general, general merchandise shippers were provided significantly less reliable service than shippers using the other two train services. Intermodal shippers were provided much more reliable service than shippers using the other two train services. For a typical intermodal O-D pair, the maximum 1-day-percent was almost 90%; the best intermodal service had a 1-day-percent of nearly 99%.

However, some general merchandise shippers received very reliable service. About 20% of general merchandise shippers were provided more reliable service than a typical bulk unit shipper. About 5% of general merchandise shippers were even provided more reliable service than the average intermodal shipper.

From the analysis of trip time and reliability of different train services in O-D pairs that have similar distance, we could derive the same conclusion on service differentiation among different train services.

3.5.2 Conclusions

A fundamental conclusion from the analysis is that railroads are differentiating services for broad classes of freight traffic. There were clear differences in the trip time and reliability of the three

different train services. It was clear that there was substantial unreliability in the level of service provided to general merchandise shippers. Shippers who use double-stack services, on the other hand, are able to take advantage of much faster and more reliable service.

There was also a certain level of variation in service levels among different O-D pairs for each train service. It was not clear, however, if such differentiated service levels are the result of intentional efforts to differentiate service considering service requirements of individual O-D pairs, or if they simply reflect the results of day-to-day operations reacting to the daily traffic variability and the uncertainty in various stages of operations. Due to the characteristics of the data, no attempt is made herein to determine the causes of unreliability.

To understand the causes of such differentiated service levels, we need additional information on the shipper's service expectation, the carrier's operating policies, the competition among railroads and the competition between rail and truck services. Analysis of other elements of service quality (e.g., costs) is not attempted in this analysis.

It also points out the need to develop more insights on the effects of service differentiation on service levels for different classes of traffic and costs to railroad. In the next chapter, we develop more insights on the effects of service differentiation on the service levels for different classes of traffic and the costs to railroad.

CHAPTER 4

Insights on Effects of Service Differentiation

4.1 Introduction

In the last chapter, an empirical analysis of service levels of different train services was discussed.

In this chapter, the effects of service differentiation on service and cost performance of the rail operation are explored. It is examined how various operating plans considering heterogeneous traffic affect service levels for different classes of traffic and the railroad's operating costs.

Railroads have been able to provide different levels of service to some classes of customers, for example, auto and intermodal. They have done this by making additional blocks, by giving priorities to certain blocks or trains, and by taking special action to make sure that train connections are made and that local deliveries take place on time. It is not clear, however, how much service distinction can be attained, how much it costs overall to provide premium service to some classes of traffic, and how much benefit can be gained in terms of service levels.

Simulation analyses were performed for the two types of rail operations: a direct train service between two terminals and a train service among multiple terminals. For simulation analyses, idealized models, which captured important characteristics of rail operations, were used

to understand fundamental trade-offs between service and cost, when a railroad tries to provide different levels of service to different classes of traffic.

In section 4.2, the results from the analysis of a direct train service is provided between two terminals, is presented. In section 4.3, the results from the analysis of a train service among multiple terminals, is presented. In section 4.4, the conclusions from the two analyses are discussed.

4.2 Case I : Train Service Between Two Terminals

A simple model, where a railroad provides a direct train service between two terminals, was designed and analyzed to obtain some useful insights into the effects of service differentiation strategies in rail freight transportation. It was assumed that shippers specify the priorities of their shipments and are willing to pay for service as a function of service quality. Service quality is defined by average trip time and trip time reliability. The mechanism used to differentiate service by priority class are train make-up rules which are dependent on the shipper supplied shipment priority. These train make-up rules were examined under different resources and operating conditions defined by train capacity, availability of empty cars, and variability of demand.

4.2.1 Case description

Shippers send their products from terminal A to terminal B daily by rail. Shippers specify three levels of priority (i.e., high, medium, low) for their shipments.¹ Daily demand from terminal A to

¹ Each shipment consists of several car loads.

terminal B by priority is probabilistic and is assumed to be normally distributed. A railroad provides service to shipments of different priorities. A train operates each day to terminal B and train transit time is two days. Train transit time is assumed to be deterministic. For each train, train capacity is limited by power availability and is defined in terms of the maximum number of cars it can haul.

A loaded car that arrives at terminal B will return to terminal A after unloading at terminal B. It is assumed that back-haul time from terminal B to terminal A is three days and deterministic. This means a loaded car arriving at terminal B will always be repositioned as an empty car at terminal A three days after its arrival time.

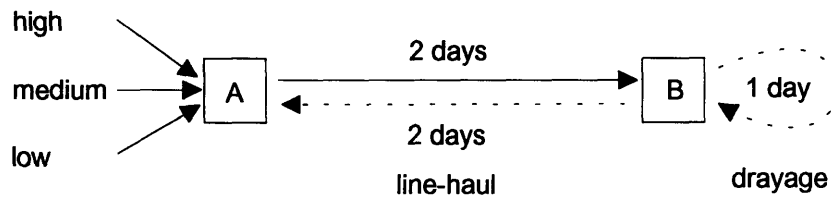


Figure 4.1 : Case network (I)

The railroad has fixed resources in terms of available power and empty cars owned. When the railroad does not have enough empty cars into which to load shipments or enough power to carry the loads from shippers, some loads will be delayed a day or more at terminal A. The railroad tries to provide different levels of service in terms of average trip time and trip time reliability to different priority classes by implementing a train make-up rule as a function of car priority.

It is assumed that the demand pattern is known and fixed during the analysis period.

Average demand from city A to city B is 120 cars per day. Average demand for high, medium and low priorities are 60, 36 and 24 cars per day, respectively. The variability of demand is defined by the coefficient of variation (C.V.) which is equally 0.3 for all priority classes.

Table 4.1 : Demand distribution by priority

<u>Priority</u>	<u>Mean</u>	<u>Std dev</u>	<u>C.V.</u>
High	60	18.0	0.3
Medium	36	10.8	0.3
Low	24	7.2	0.3

4.2.2 Train make-up policies

As examples of decisions for service differentiation, various train make-up policies were studied. Train make-up policy in general determines which blocks of cars are to be connected to a given train. Since traffic conditions vary, the exact number of cars in any block will not be known in advance. Therefore, the total number of cars which need to be connected to a given train may sometimes be more than the train capacity (i.e., maximum number of cars the train can carry). These cars include cars arriving at a given day with different assigned priorities and cars that missed train connections from previous days. Railroads need rational train make-up rule that consider both the heterogeneity of traffic in terms of different assigned priorities and the train capacity.

We compared three heuristic train make-up rules. Train make-up rule 1 makes up trains solely based on priorities assigned by the shipper (high, medium or low). Train make-up rule 2

gives higher preference to the car that missed more train connections on previous days regardless of given priority. Train make-up rule 3 balances the approaches of train make-up rules 1 and 2.

Train make-up rule 1

Train make-up rule 1 makes up trains solely based on priorities assigned by the shipper (i.e., regardless of previous missed connections). For cars that have the same priority, this rule gives higher preference to the car that missed more connections on the previous days.

If there is available capacity in an outbound train

Then, connect high priority cars to train

If there is available capacity

Then, connect medium priority cars

If there is available capacity

Then, connect low priority cars

Train make-up rule 2

Train make-up rule 2 gives higher preference to the car that missed more train connections on previous days regardless of given priority. For cars that have the same number of missed connections, this rule then considers given priority. This rule is similar to FCFS (first-come-first-served) service rule.

If there is available capacity in outbound train

If there are cars that missed previous train connection(s)

Then, connect these cars in the order of more missed connections

If there is available capacity

Then make up train by rule 1

Train make-up rule 3

This rule allows a differentiated number of missed connections to different priority cars. It allows no missed connection for high priority cars. It allows one missed connection for medium priority cars as long as high priority cars can be fully served. It allows up to two missed connections for low priority cars as long as high priority cars can be fully served and there are no medium priority cars that missed more than one connection.

If there is available capacity in outbound train

Then, connect high priority cars

If there is available capacity

If there are medium priority cars that missed more than one connection

Then, connect these cars

If there is available capacity

If there is low priority cars which missed more than two connections

Then, connect these cars

If there is available capacity

Then make up train by rule 1

4.2.3 Performance measures

As an output of the simulation, three categories of performance were evaluated:

- Loaded trip time and trip time reliability, as measures of service
- Rail carrier's operating cost
- Shippers' total logistics costs

In this study, trip time reliability is measured in terms of the variance of trip time.

Rail carrier's operating cost

To see the effects on operating performance, the rail carrier's total and per car operating costs were evaluated. Train cost, car-time cost and car switching cost were considered as important elements of the rail carrier's total operating cost.

Per mile train cost is derived as a function of train-miles, locomotive-miles and locomotive-hours.² Per mile train cost for train capacity 120 cars, 150 cars and 180 cars are \$8.53, \$9.91 and \$11.30 per train-mile respectively. Per mile train cost is increased as train capacity is increased since more power is used.

Per hour car time cost is \$0.75/hour. Car switching cost is derived as a function of switch engine hours required. It is assumed that cars of different priority need different switch engine

² Train transit time from city A to city B is 2 days. Its distance is considered as 1,440 miles assuming average train speed is 30 miles/hour. Each locomotive can carry 30 cars. It means we need 4 locomotive to dispatch train capacity of 120 cars.

hours. Per car switching costs for high, medium and low priority cars are \$25.44, \$19.29 and \$12.86 per car³, respectively. The per car switching cost of a high priority car is higher than that of a low priority car because it is assumed that the handling of high priority car requires more switch engine hours.

Shippers' total logistics costs

To see the potential effects on market performance, the shipper's total logistics cost, resulting from service differentiation strategies, was evaluated. The total logistics cost model was used to predict how cost-minimizing shippers will react as trip time performance of rail service changes. A rail carrier can estimate how much shippers are willing to pay for additional service improvement using this model. Additional information was assumed on the value of commodity by priority class. The value of commodity of high, medium and low priority classes are assigned to be \$200,000, \$100,000 and \$50,000 per car-load, respectively.

The continuous review inventory model with stochastic demand and lead time was used to compute the shipper's total logistics cost (see Chapter 2.4). It was assumed that the ordering cost is \$500, the inventory and in-transit inventory carrying cost are 10%, and the stock-out cost is 12% of the value of commodity. Since each priority has a different demand distribution and value of commodity, we computed total logistics cost for each priority class separately. Then the shippers' total logistics cost was obtained by summing the logistics costs of all priority classes.

³ Switch engine hour for high, medium and low priority car are assumed 16 mins/car, 12 mins/car and 8 mins/car, respectively.

4.2.4 Simulation analysis

Experiment design

To analyze the effects of different train make-up rules, rail operations were simulated for different operating scenarios that are defined by :

- Coefficient of variation of daily demand distribution : 0.3
- Initial empty car inventory at terminal A : 600, 650, 700, 750, 800
- Outbound train capacity : 130, 140, 150, 160
- Three train make-up rules : see section 4.2.2

The simulation result for each scenario is obtained from 100 independent simulation runs.⁴

Each run simulates the 30 days of rail operations.

Simulation results

Service and cost trade-off

The effects of train capacity, initial empty car inventory and train make-up rules on the aggregate trip time performance were examined. Aggregate trip time performance was measured by the average and variance of all traffic (of all priorities).

⁴The sample size can be determined to satisfy requirements imposed on a significance level (α) and a power ($1-\beta$) of the test. For the two-sample *t*-test to test the difference in mean parameters of two experiments, the minimum sample size 10 is required for a significance level $\alpha=0.05$ and power of test $1-\beta=0.95$. For more detailed discussions, see Larsen and Marx, 1986 (p. 395).

As the train capacity increases, average and variance of trip time of all traffic are improved. This is mainly because a car has less chance to miss its appropriate connection as the capacity of an outbound train increases. The system needs more power to dispatch more train capacity. As the system dispatches more train capacity, therefore, the per car operating cost is increased mainly due to the increase of train cost.

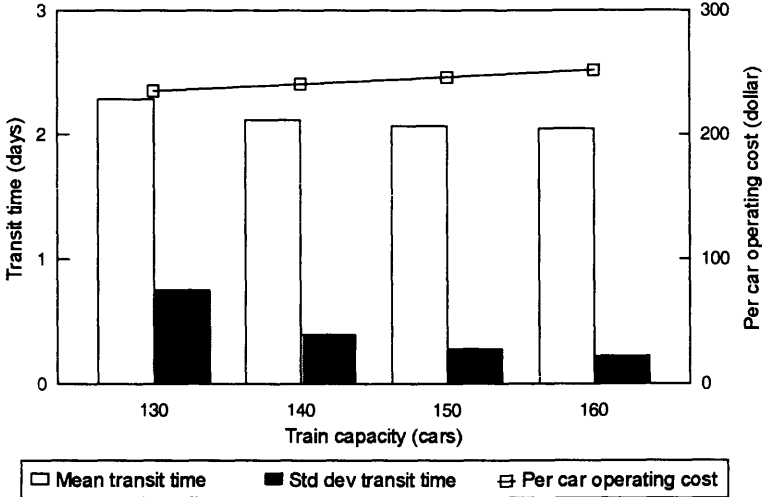


Figure 4.2 : Effects of train capacity on service and cost (initial empty car inventory 700 cars, train make-up rule 1)

The effects of initial empty car inventory were also examined. As the system has more initial empty cars available, average and variance of trip time of all traffic are improved. For a given train capacity, however, average and variance of trip time were not improved beyond a certain level of initial empty car inventory. Although the system has more initial empty cars available, some empty cars are not utilized at all during the 30 day operation period unless the system has more demand and can dispatch more train capacity. For the scenario examined (train capacity 150 cars, train make-up rule 1), average and variance of trip time are not improved after

the system has more than 750 cars as initial empty car inventory. Since we charge a car time cost for the cars whether or not they are utilized, having more than this level of initial empty car inventory simply increases the cost without improving the service.

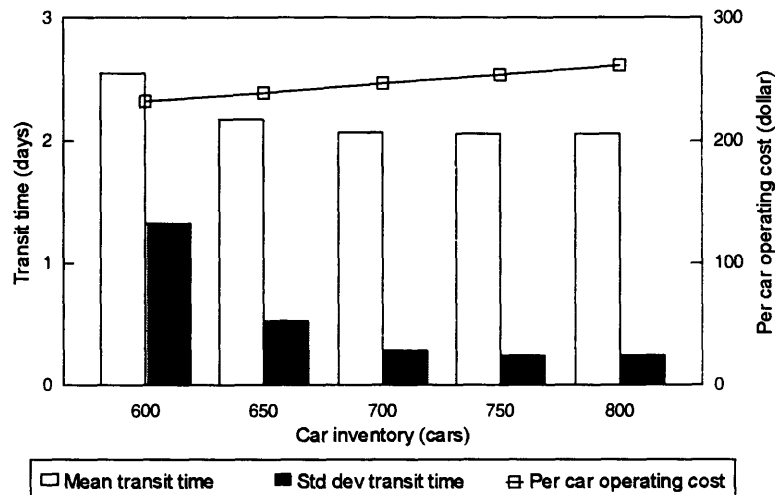


Figure 4.3 : Effects of initial empty car on service and cost (train capacity 150 cars, train make-up rule 1)

Figure 4.2 and 4.3 show that both trip time and reliability for all traffic are improved as the system can dispatch more train capacity or has more empty car inventory. It indicates that there is a clear trade-off between service and cost. This result is consistent with the results of a previous study that showed more direct and frequent train services reduce the mean transit time, but increase costs (see Keaton [1991]).

An important question is whether such a cost increase to improve overall service level can be justified and if so, for which traffic.

Service levels by priority class

The disaggregate performance by priority was examined to develop insights into the effects of service differentiation practices. A system that has 700 cars as initial empty car inventory and 130 cars as train capacity was considered. The effects of the three train make-up rules were examined.

Figure 4 shows trip time distributions for all traffic under different train make-up rules. Different train make-up rules change the variance of trip time although they do not change the average trip time. In the aggregate, for the examined operating condition, implementing train make-up rule 2, which is designed to minimize the total number of missed connections at a terminal, gives the lowest variance of trip time for all traffic (also see Table 4.2). It suggests that a railroad can improve the service reliability (measured by the variance of trip time) for all traffic by implementing an effective operating policy, although it may not improve the average trip time unless it increases the service capacity (e.g., train capacity, the empty car inventory level).

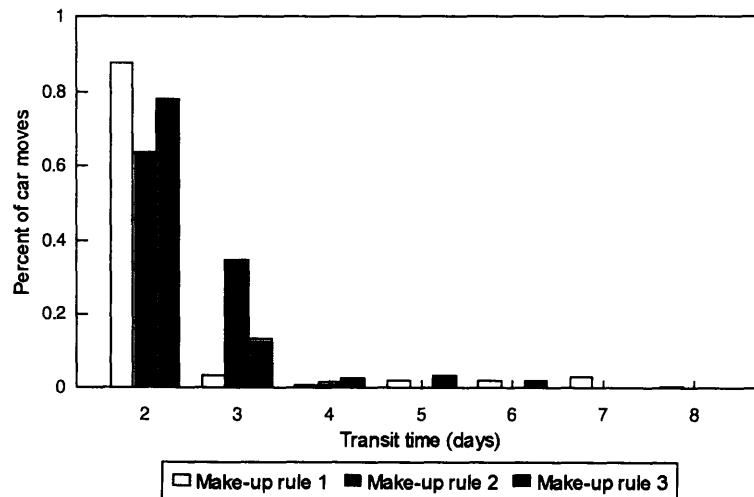


Figure 4.4: Trip time distributions of all traffic under different train make-up rules (initial empty car inventory 700 cars, train capacity 130 cars)

Table 4.2 shows trip time performance by priority class under different train make-up rules. It is found that different train make-up rules cause different average and variance of trip time by priority class. Different levels of service are provided for different priority classes by implementing different operating plans.

Two sample t -tests were done to examine whether the difference in trip time performance among different classes of traffic are statistically significant (Table 4.3).⁵ Under any train make-up rule, the difference in average and variance of trip time among different classes of traffic were statistically significant at significance level 0.05.

Two sample t -tests also were done to examine whether the differences in trip time performance as results of implementing three train make-up rules are statistically significant (Table 4.4). As we discussed, there were no differences in average trip time for all traffic among the three train make-up rules. The differences in variance of trip time for all traffic among the three train make-up rules were statistically significant at significance level 0.05. The differences in average and variance of trip time by priority among the three train make-up rules, in general, were statistically significant at significance level 0.05. There were, however, no difference in both average and variance of transit time of high priority traffic between train make-up rules 1 and 3.

⁵ Two-sample t -test to test the difference in mean parameters of two experiments can be found in any statistics book (for example, see Larsen and Marx, 1986, pp. 362-371). Let $X_1, X_2, \dots, X_n \sim N(\mu_X, \sigma^2)$ and $Y_1, Y_2, \dots, Y_m \sim N(\mu_Y, \sigma^2)$ and let the X 's and Y 's be independent. At the α level of significance, the null hypothesis $H_0 : \mu_X = \mu_Y$ can be rejected if

$$t = \frac{\bar{x} - \bar{y}}{s_p \sqrt{\frac{1}{n} + \frac{1}{m}}} \text{ is either } \begin{cases} \leq -t_{\alpha/2, n+m-2} \\ \geq +t_{\alpha/2, n+m-2} \end{cases}$$

$$\text{where, } s_p^2 = \frac{(n-1)S_X^2 + (m-1)S_Y^2}{n+m-2}$$

For the case we examined, at the significance level 0.05, $t_{\alpha/2, n+m-2} = t_{0.025, 198} = 1.96$.

The comparison of the three train make-up rules indicates that varying the way that different service requirements of traffic are incorporated into an operating scheme results in different levels of service to different classes of traffic. With an operating scheme that properly considers different service requirement of different shippers, a railroad can provide fast and very reliable service to service-sensitive shippers and can provide slow and less reliable service to price-sensitive shippers.⁶ Furthermore, it allows a railroad to utilize existing service capacity more efficiently, avoiding additional investment to increase service capacity to improve the overall service performance for all traffic.

For example, if the railroad want to improve the aggregate trip time performance of all traffic to the service level of high priority traffic with train make-up rule 1 (average trip time 2 days and no variance of trip time), it needs to have an initial empty car inventory of 800 cars and purchase more power to dispatch train capacity 200 cars.⁷

Table 4.2 : Trip time performance under different train make-up rules
(initial car inventory 700 cars, train capacity 130 cars)

<u>Priority</u>		<u>Rule 1</u>	<u>Rule 2</u>	<u>Rule 3</u>
High	Mean	2.00 days	2.10 days	2.00 days
	Std dev	0.00	0.21	0.00
Medium	Mean	2.06	2.34	2.21
	Std dev	0.22	0.43	0.33
Low	Mean	3.32	2.64	3.10
	Std dev	1.16	0.50	0.93
Total	Mean	2.28	2.28	2.28
	Std dev	0.75	0.43	0.63

⁶ Assuming that a shipper sending high priority car is more service-sensitive and a shipper sending low priority car is more price-sensitive

⁷ It was found from the additional simulation runs with different levels of initial empty car inventory and train capacity.

Table 4.3 : Testing differences in trip time performance among traffic classes
(two-sample *t*-test)

		<u>Rule 1</u>	<u>Rule 2</u>	<u>Rule 3</u>
t-value	Mean	-12.00	-8.09	-10.50
(high-medium)	Std dev	-24.44	-10.87	-23.57
t-value	Mean	-13.11	-8.49	-11.91
(medium-low)	Std dev	-18.87	-5.59	-16.32

Table 4.4 : Testing differences in trip time performance among train make-up rules
(two-sample *t*-test)

<u>Priority</u>		<u><i>t</i>-value (rule 1-rule 2)</u>	<u><i>t</i>-value (rule 2-rule 3)</u>	<u><i>t</i>-value (rule 3-rule 1)</u>
High	Mean	-6.25	6.25	0.00
	Std dev	-12.35	12.35	0.00
Medium	Mean	-10.98	4.06	7.28
	Std dev	-14.78	5.60	6.60
Low	Mean	6.85	-6.04	-1.83
	Std dev	13.37	-12.45	-3.86
Total	Mean	0.00	0.00	0.00
	Std dev	7.41	-6.32	-2.32

In the aggregate, implementing train make-up rule 2 gives the lowest variance of trip time for all traffic. This rule, however, improved reliability by significantly improving reliability of low priority traffic while lowering reliability of high priority traffic. As a result, the rail carrier improved the service level for price-sensitive shippers and worsened the service level for service-sensitive shippers. A policy that gives the best reliability in the aggregate is often not be the best policy where there are different classes of traffic. It suggests that simply measuring aggregate reliability performance might mislead the rail carrier in evaluating the current level of service and in improving service. It is important to measure service quality disaggregated by traffic class (or

by market segment). Figures 4.5 to 4.7 graphically show that trip time distribution for all traffic is the result of the summing of trip time distributions for traffic of the three different priority classes.

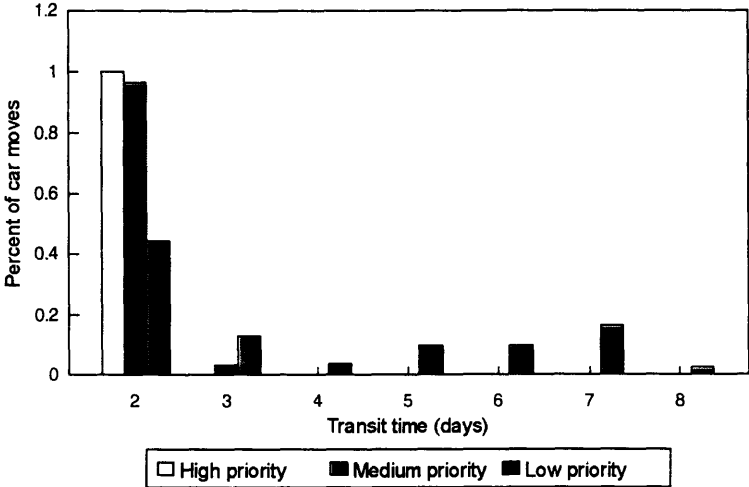


Figure 4.5 : Trip time distribution by priority class under train make-up rule 1

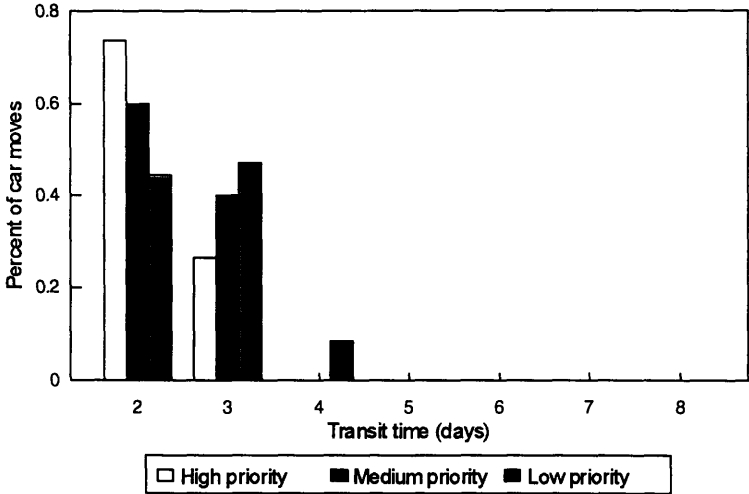


Figure 4.6 : Trip time distribution by priority class under train make-up rule 2

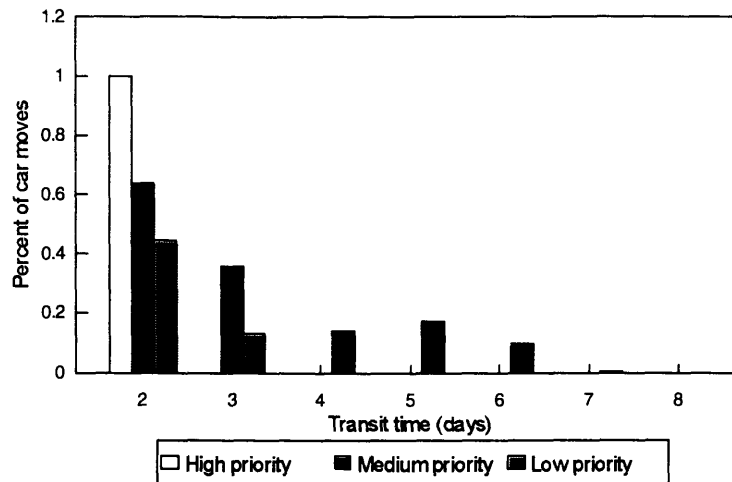


Figure 4.7 : Trip time distribution by priority class under train make-up rule 3

Selecting the best rule

Which of the three train make-up rules should the rail carrier implement? The best rule can be the one that best satisfies shippers' different service expectations without significantly increasing the carrier's operating costs. To examine this issue, the total logistics cost was examined by priority class, based on the trip time performance and assumption on shippers' inventory policies (see "shippers' total logistics costs" in Section 4.2.3).

Table 4.5 shows the per car logistics cost by priority class.⁸ Under any train make-up rule, the differences in logistics cost levels among different classes of traffic were statistically significant at significance level 0.05 (Table 4.6).

Train make-up rule 2 gave the lowest logistics cost only to low priority class. Train make-up rule 1 gave the lowest logistics cost for high and medium priority classes while it gave the

⁸ Assuming the transportation rate is fixed for the duration of each scenario, we computed the total logistics cost, excluding the transportation cost, for 30 days period.

highest logistics cost for low priority class. Overall, train make-up rule 1 gave the lowest total logistics costs for all traffic. Assuming that shippers always try to minimize their total logistics costs, train make-up rule 1 is the best policy, from the shipper's point of view, of the policies examined.

Table 4.5 : Per car logistics cost under different train make-up rules
(initial car inventory 700 cars, train capacity 130 cars)

<u>Priority</u>	<u>Rule 1</u>	<u>Rule 2</u>	<u>Rule 3</u>
High	\$ 210.11	\$ 229.09	\$ 210.11
Medium	122.25	140.62	131.91
Low	116.28	85.32	105.44
Total	165.25	174.12	165.98

Table 4.6 : Testing differences in per car logistics costs among traffic classes
(two-sample *t*-test)

	<u>Rule 1</u>	<u>Rule 2</u>	<u>Rule 3</u>
<i>t</i> -value (high-medium)	184.97	32.53	65.27
<i>t</i> -value (medium-low)	1.97	40.01	10.85

Another important issue is what is the most reasonable investment on service capacity (i.e., train capacity, empty car inventory) to improve the service quality. Already observed is a clear trade-off between the aggregate service level and the total operating cost in Figures 4.2 and 4.3. We raised a question whether such cost increase to improve overall service level could be justified for all traffic classes.

The effects of train capacity increase on service levels are examined by priority class. Figures 4.8 and 4.9 show that additional service capacity increases primarily improve the service

quality of medium and low priority classes for the given system. Note that shippers of different market segments have different expectations for service quality and different willingness to pay for the additional service improvement. Therefore, an additional investment to increase service capacity need to be justified by potential revenue increase with existing market share or by gaining market share.

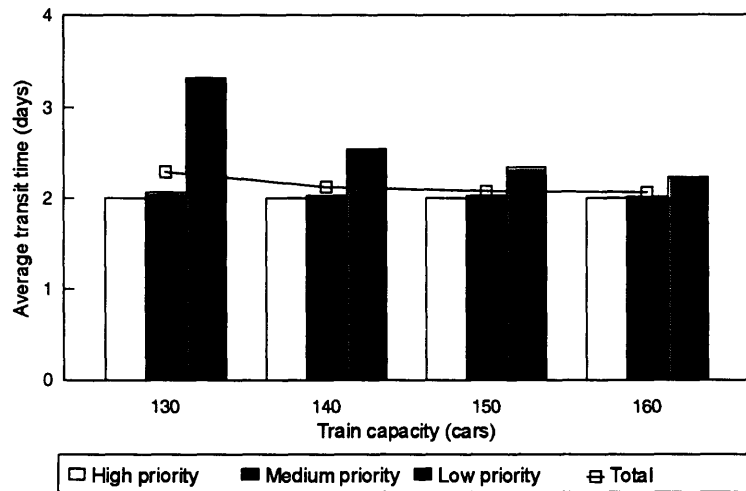


Figure 4.8 : Effect of train capacity on average trip time by priority (initial car inventory 700 cars, train make-up rule 1)

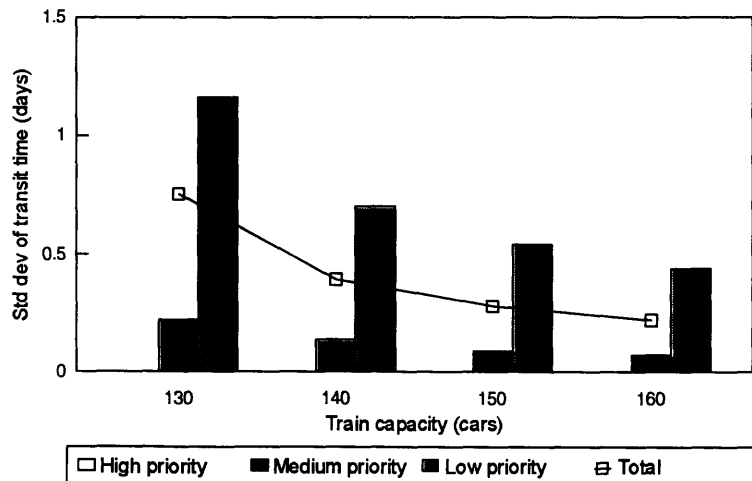


Figure 4.9 : Effect of train capacity on variance of trip time by priority (initial car inventory 700 cars, train make-up rule 1)

Table 4.7 shows the changes in total operating cost and total logistics cost as we add train capacity. Suppose shippers are willing to pay for their logistics cost decrease due to the improvement of transportation service (see section 2.3). An additional investment to dispatch more train capacity needs to be justified by balancing the decrease in total logistics cost with the increase in total operating cost.

As an idealized situation, suppose a shipper operate his own transportation company and is in full control of transportation service. In this case, the transportation cost is simply internal. For such a “closed” system, the best option is to minimize the sum of the total operating cost and the total logistics cost (except transportation cost since it is the internal cost to company).

When the system that has initial empty car inventory of 700 cars and implement train make-up rule 1, the train capacity that minimizes the sum of total operating and logistics costs is 140 cars. Beyond this point, additional investment to dispatch more train capacity exceeds the reduction in total logistics cost.

Table 4.7 : Per car operating and logistics costs change as train capacity increase
(initial car inventory 700 cars, train make-up rule 1)

	<u>130 cars</u>	<u>140 cars</u>	<u>150 cars</u>	<u>160 cars</u>
Per car operating cost	235.35	240.90	246.46	252.01
Per car logistics cost	165.25	158.87	156.93	156.03
Sum of O&L costs	400.60	399.77	403.39	408.04

In this context, the best investment option and train make-up rule may be found for the given system. As the train capacity and/or initial empty car inventory is increased, the total logistics cost is decreased while the total operating cost is increased and vice versa. Figures 4.10

to 4.12 show the sum of total operating cost and total logistics cost by train capacity and initial empty car inventory under different train make-up rules. As a result of examining all simulation scenarios (combinations of train capacities of 130, 140, 150, 160 cars; initial empty car inventory of 600, 650, 700, 750, 800 cars; three train make-up rules), the option that minimizes the sum of total operating cost and total logistics cost is obtained as the combination of train capacity 130 cars, initial empty car inventory 650 cars and train make-up rule 1, given the assumption on the demand variability, carrier’s operating cost structure, and shipper’s characteristics.

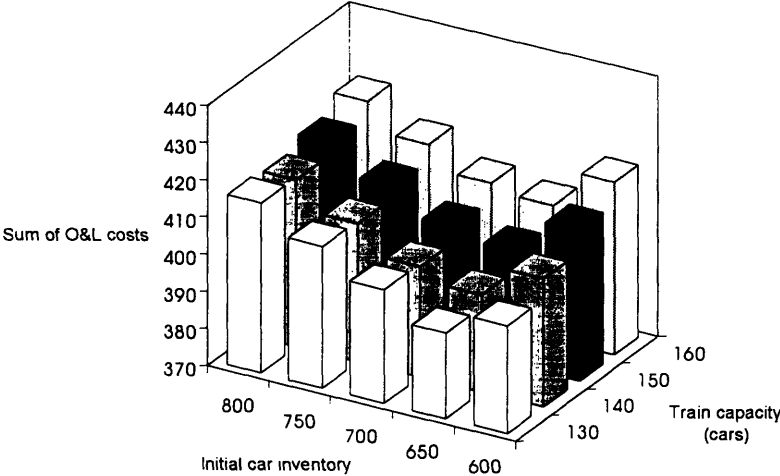


Figure 4.10 : Sum of operating cost and logistics cost by train capacity and initial car inventory : train make-up rule 1

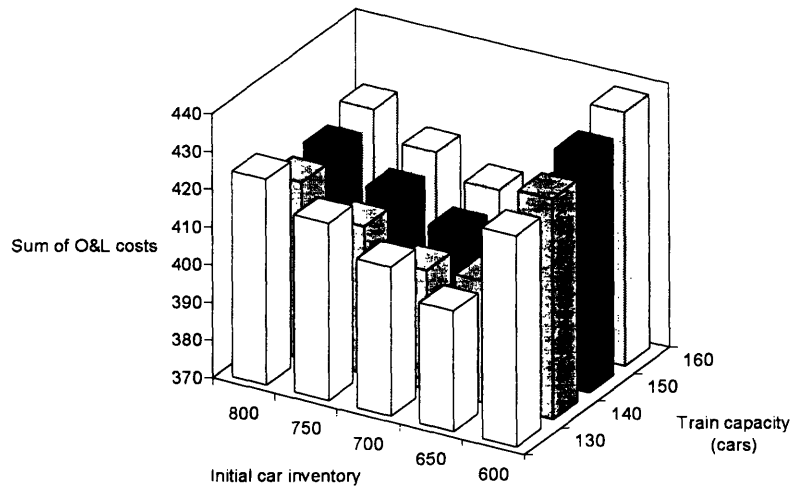


Figure 4.11 : Sum of operating cost and logistics cost by train capacity and initial car inventory : train make-up rule 2

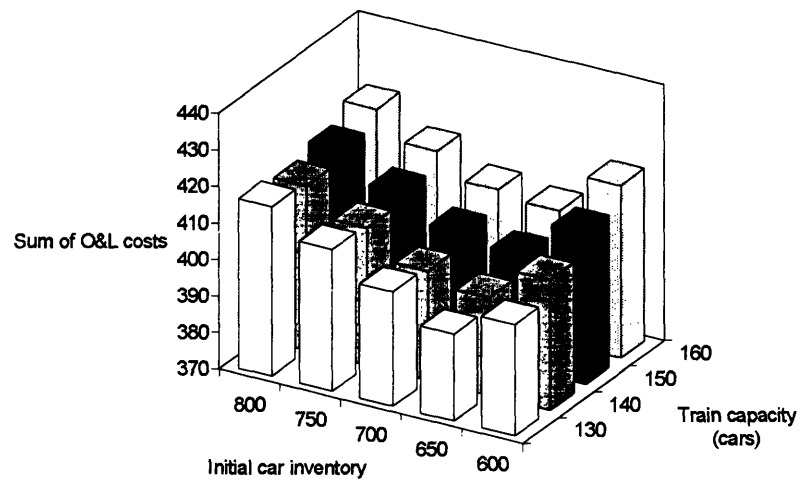


Figure 4.12 : Sum of operating cost and logistics cost by train capacity and initial car inventory : train make-up rule 3

Cost allocation and cost by priority

The operating costs for different operations were examined, with emphasis on the cost per car for each priority class. This is an important measure of effectiveness for service differentiation strategies, especially for pricing decisions and for management control purposes.

Some rational way is needed to allocate common costs to get cost per car by priority class. Cost allocation is not simple, especially when common cost is a large part of the total operating cost. In transportation, capacity related costs (e.g., train cost, train and car delay cost, track cost, etc.) are good examples of common cost. For this study, a heuristic method was used for cost allocation.

Again, a system that has 700 cars as initial empty car inventory and 130 cars as train capacity was considered. The cost allocation procedure will be explained using the result of one simulation; for 30 day operations, train cost was \$388,368, car-time cost was \$378,000, car switching cost was \$75,396 and total operating cost was \$841,764.

Allocation of car-time cost

If the total car-time cost to cars of each priority class is allocated based simply on the amount of total trip time for traffic of each priority, it results in a higher car-time cost for a low priority car and a lower car-time cost for a high priority car.

This does not make much sense because the additional delay for a low priority car actually comes from priority handling of high priority traffic. Under train make-up rule 1, a typical car experiences average trip time 2.38 days. Because the railroad gives different preference to different priority cars in train make-up, a low priority car experiences an average trip time of 3.87

days. On the other hand, high and medium priority cars experience an average trip time of 2 days and 2.04 days (see Table 4.2). It means, on average, high and medium priority cars can save 0.38 days and 0.34 days by delaying low priority cars 1.49 days. Therefore, the additional delay cost of low priority traffic should be redistributed to higher priority traffic on some rational basis.

Under this concept, car-time cost per car was calculated by dividing total car-time cost by total demand, regardless of priorities. The basic idea behind this is that each car, on average, equally contributes to the increase of total car-time. Car-time cost per car is obtained as \$105.9 per car.

Allocation of train cost

A heuristic method was used to allocate train cost based on the following observations:

- Daily train capacity is fixed and therefore daily train cost is fixed
- Low priority cars miss appropriate train connections due to surrendering capacity to high priority cars

On the basis of these observations, the total train cost is allocated to each priority class based on the following steps.

Step 1. Identify number of cars with no missed connection : The number of cars that missed no train connection is identified by priority class.

Step 2. Train cost allocation : Total train cost is allocated to priority classes, based on the number of cars that missed no train connection.

Step 3. Train cost per car : For each priority class, the allocated train cost is divided by the total demand to obtain the train cost per car.

Table 4.8 : Train cost allocation process

	<u>Step 1</u>	<u>Step 2</u>	<u>Demand</u>	<u>Step 3</u>
High	1,797 cars	\$222,260	1,797 cars	\$123.7/car
Medium	1,030	127,395	1,068	119.3
Low	313	38,713	706	54.8
Total	3,140	388,368	3,571	108.8

Table 4.9 summarizes the results of train cost allocation for different train make-up rules. A high priority car always has higher train cost per car than a low priority car. In addition, different train make-up rules result in different train cost per car for each priority. This is mainly because different train make-up rules result in different patterns of car connection by priority class.

Table 4.9 : Train cost per car by priority for different train make-up rules

	<u>Rule 1</u>	<u>Rule 2</u>	<u>Rule 3</u>
High	\$123.7	\$125.6	\$139.1
Medium	119.3	102.3	88.9
Low	54.8	75.7	61.6
Total	108.8	108.8	108.8

Cost by priority

The car-time cost and train cost are heuristically allocated to each priority class. Given the car switching cost information by priority (see Section 4.2.3), total operating cost per car by priority

class is summarized as Table 4.10. We explained the cost allocation procedure using the result of one simulation run.

Table 4.10 : Operating cost per car by priority for different train make-up rules (from one simulation run)

<u>Priority</u>	<u>Rule 1</u>	<u>Rule 2</u>	<u>Rule 3</u>
High	\$255.0	\$256.9	\$270.3
Medium	244.4	227.4	214.1
Low	173.5	194.4	180.4
Total	235.7	235.7	235.7

In each simulation run, the same procedure for cost allocation was done. Table 4.11 summarizes the average cost by priority obtained from all simulation runs. As the result of cost allocation, different levels of cost for different priority services are obtained (high priority service had a higher cost and low priority service had a lower cost). Under any train make-up rule, the differences in operating cost levels among different classes of traffic were statistically significant at significance level 0.05 (Table 4.12).

Table 4.11 : Operating cost per car by priority for different train make-up rules

<u>Priority</u>	<u>Rule 1</u>	<u>Rule 2</u>	<u>Rule 3</u>
High	\$257.1	\$266.4	\$264.6
Medium	243.0	222.8	229.3
Low	168.8	175.5	170.5
Total	235.3	235.3	235.3

Table 4.12 : Testing differences in per car operating costs among traffic classes
(two-sample *t*-test)

	<u>Rule 1</u>	<u>Rule 2</u>	<u>Rule 3</u>
<i>t</i> -value (high-medium)	9.74	17.51	12.53
<i>t</i> -value (medium-low)	25.63	14.02	15.91

It suggests that railroads may differentiate on price for different priority service by providing different levels of service at different costs, and make a profit on all priority classes. For example, a railroad may design differentiated service products (with train make-up rule 1) as shown in Table 4.13 and make a profit on all priority classes.

Table 4.13 : Example of differentiated service products

	Trip time		Cost (\$/car)	Price (\$/car)
	Mean (days)	Std dev (days)		
High	2.00	0.00	257.1	308.52
Medium	2.06	0.22	243.0	291.60
Low	3.32	1.16	168.8	202.06

4.2.5 Findings

The analysis of the example network reveals some important insights on the effects of service differentiation.

- Both trip time and reliability (measured by the variance of trip time) for all traffic were improved as the system can dispatch more train capacity or has more empty car inventory. It indicates that there is a clear trade-off between service and cost. This result is consistent with the results of a previous study (Keaton [1991]) that showed more direct and frequent train services reduce the mean transit time, but increase costs. An important question is whether such a cost increase to improve overall service level is justified for traffic of all priorities.
- Different train make-up policies resulted in different levels of trip time reliability for all traffic. It suggests that a railroad can improve service reliability for all traffic by implementing an effective operating policy, although it may not improve average trip time unless it increases the service capacity (e.g., train capacity, empty car inventory level).
- When the service levels are evaluated by priority class, different train make-up policies resulted in different trip time and reliability for different priority classes. Different levels of service can be provided for different traffic classes by implementing different operating plan. Varying the way that different service requirements of traffic are incorporated into an operating scheme results in different levels of service to different classes of traffic.
- With an operating plan that properly considers different service requirements of different shippers, a railroad can provide fast and very reliable service to service-sensitive shippers and can provide slow and less reliable service to price-sensitive shippers. Furthermore, it allows a railroad to utilize existing service capacity more efficiently, avoiding additional investment to increase service capacity to improve the service performance for all traffic classes that may not require or be willing to pay for higher quality service.
- A policy that gives the best reliability for all traffic might not be the best policy where there are different classes of traffic. A total logistics cost model was used to measure the shippers' potential willingness to pay for the additional service improvement. This measure of

effectiveness allowed us to choose the policy that minimizes the total logistics cost among three train make-up policies.

- The analysis showed that efforts to provide highly reliable service were not justified for all market segments. It also suggests that simply measuring the service performance for all traffic might mislead the rail carrier in evaluating the current level of service and in improving service. The logistics cost analysis also allowed the determination of how much additional investment to increase the service capacity could be justified by potential market reaction in terms of the shippers' potential willingness to pay for the additional service improvement.
- Cost per car for each priority is an important measure of effectiveness of service differentiation strategies, especially for pricing decisions and for management control purposes. A heuristic cost allocation method was used to obtain the operating cost by priority. Using it, high priority service had a higher cost and low priority service had a lower cost. It suggests that railroads may differentiate on price for different priority service by providing different levels of service at different costs, and may make profits on all traffic classes.

4.3 Case II : Train Service among Multiple Terminals

The previous simulation analysis focused on a direct train service between two terminals. Another simulation was designed to further examine the effects of service differentiation for a train service among multiple terminals.

Because of the level of fixed train costs, railroads have an incentive to operate long trains. It is not generally economical to provide direct train services to all pairs of terminals.

To move a larger number of cars on a single train, a railroad consolidates cars for a number of destinations into trains and breakup trains into groups of cars that have common destinations in the terminal. This practice can make rail service more cost effective than competing motor carrier services. On the other hand, rail service is typically slower and less reliable due to the delays involved in consolidation and breakup of trainloads of individual cars. Many cars must change trains in the intermediate terminals, and will usually encounter a delay at the terminal. Rail operations thus involve a trade-off between the economies from shipment consolidation and the resulting delays in trip times (Keaton [1991]).

The mechanism to differentiate service quality by priority class is the train make-up policy as a function of shipment priority, as in the previous model. Two types of train make-up policies were examined under different operating conditions defined by the variability of demand and train transit time. In addition, we sought to determine the amount of traffic that can consistently be provided with highly reliable service by a system that has a limited resources to provide services.

First, we examined whether some of the conclusions we had obtained from the two terminal case are applicable to the multiple terminal case. Specifically, we examined whether train make-up policies, which consider service requirements of different shippers, produce clearly differentiated service levels for different classes of traffic for a train service among multiple terminals where cars move through several intermediate terminal handling. Two train make-up policies were examined. A train make-up policy that has a pre-determined and fixed pull sequencing order in the train make-up process is compared with another policy that dynamically changes pull sequencing order. Second, by examining the two train make-up policies under a system that has different variabilities in demand and train transit time, and different traffic mixes, we examined the “robustness” of these policies in differentiating services.

A rail network simulation model was developed to evaluate the service and cost performance under different simulation scenarios (see Appendix A).

4.3.1 Case description

The example network has 5 terminals. A railroad provides train service for carload traffic among these terminals.



Figure 4.13 : Case network (II)

Each terminal handles traffic that can be segmented into several traffic classes. The network serves 20 origin-destination pairs (every pair of terminals in the network, in both directions). The railroad classifies shipments of each origin-destination pair into two priority classes; high and low priority classes. The demand variability is considered by incorporating a weekly demand pattern and stochastic daily demand pattern. Daily average demand for each origin-destination is 120 cars. Total daily average traffic volume is 2,400 cars. The same weekly traffic pattern is assumed for all origin-destination pairs. It is assumed that the daily demand distribution for each origin-destination is probabilistic and normally distributed.

12 trains (6 trains in each direction) provide train service in the network . Each train stops at all the intermediate terminals (Figure 4.14). The train transit time between any two

terminals is 8 hours. It is assumed that train transit times are probabilistic and normally distributed.

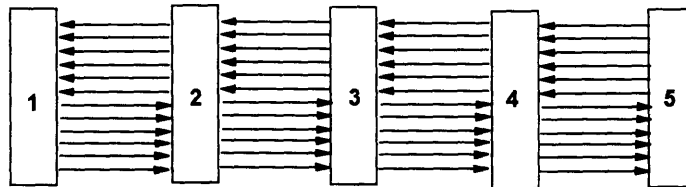


Figure 4.14 : Train service network (II)

For each train, train capacity is limited by power availability and is defined in terms of the maximum number of cars it can haul. Train capacity is designed as 120 cars per train.¹ This designed train capacity can be dispatched only if there are enough locomotives in the terminal. If there are not enough locomotives in the terminal, a shorter train will be dispatched based on the number of locomotives available. When the actual train length is shorter than the designed train capacity, however, the railroad still dispatches the planned locomotives.

It is assumed that each terminal builds blocks only for the adjacent terminals that can be directly reached by trains departing from that terminal. Separate blocks are defined for different traffic classes. The railroad tries to provide different levels of service in terms of average trip time and trip time reliability to different priority classes by implementing a train make-up policy as a function of a car priority.

The two train make-up policies were examined under different operating conditions defined by the coefficient of variation of demand, the coefficient of variation of train transit time, and the shares of different priority classes.

¹ It is assumed that a single unit of locomotive can haul 40 cars. To dispatch 120 cars of train capacity requires 3 locomotive units.

4.3.2 Train make-up policies

A train make-up plan defines which blocks of cars are to be connected to a given train. In current practice, any outbound train has a “take-list” that specifies the list of blocks of cars it may pick up from the classification tracks. Because traffic conditions vary from day to day, the exact number of cars in any one block at any particular moment will not be known in advance. The total number of cars in the blocks designated to be made up into an outbound train may therefore sometimes be more than the defined train capacity.

Static make-up policy

“Static make-up policy” pre-determines the order of preference in pull sequencing for assembling an outbound train. Under this policy, blocks are sequentially pulled based on the order specified in the take list until the dispatchable train capacity is reached.

Using a static make-up policy, low priority blocks may be more likely to miss their appropriate train connections. This policy may result in delays to low priority shipments, while high priority shipments usually arrive at their destination on-time, and often even arrive earlier than scheduled.

Dynamic make-up policy

An alternative policy in train make-up process is designed to reduce the potential excessive delays of low priority shipments while maintaining a satisfactory level of service for high priority

shipments. This policy changes the order of preference of blocks (i.e., changes the pull sequence) in pull sequencing for a train make-up process. The pull sequence order for a train is determined by a “expected penalty” for each car that is computed based on the relative earliness or lateness relative to schedule, and a penalty on early or late arrival at the destination. This policy is referred to as a "dynamic make-up policy."

Dynamic sequencing process

The algorithm used for a dynamic make-up policy is described as follows. First, at terminal n , the expected penalty for each shipment² m is computed as follows.

$$W_m^n = \bar{\phi}_m \max\{0, E[T_m] - \bar{T}_m\} \quad (4.1)$$

$$E[T_m] = \hat{T}_m^n + E[T_m^n] \quad (4.2)$$

where, W_m^n is the expected penalty for shipment m , \bar{T}_m is the scheduled trip time from the origin to the destination for a shipment m , $E[T_m]$ is the expected trip time from the origin to the destination, \hat{T}_m^n is the actual trip time from the origin to the intermediate terminal n , $E[T_m^n]$ is the expected trip time from the intermediate terminal n to the destination, and $\bar{\phi}_m$ is the penalty on early or late arrival at the destination of a shipment m .

In this study, the scheduled trip time from the origin to the destination is obtained by finding the shortest possible schedule from the origin to the destination of a shipment m . The

² A shipment can be defined as a group of cars with the same origin, the same destination, and the same departure time from the origin terminal.

expected trip time from the terminal n to the destination is obtained by finding the expected shortest possible schedule from the terminal n to the destination of a shipment m .

The service level by market (defined by origin, destination and priority) under a dynamic make-up policy depends on how the service standards (i.e., delivery time window and penalty on unreliable service) are established for each market. In this simulation, it is assumed that only late arriving shipments are penalized (see Equation 4.1) and the penalty cost used for the low and high priority shipments are assigned at \$3/car-hour and \$9/car-hour, respectively.

We can find what blocks should be connected to a given train by solving the following knapsack problem.

$$\text{Max} \quad \sum_m W_m^n x_m y_m = \sum_m \bar{\phi}_m \max\{0, \hat{T}_m^n + E[T_m^n] - \bar{T}_m\} x_m y_m \quad (4.3)$$

$$\text{s.t.} \quad \sum_m x_m y_m \leq CAP \quad (4.4)$$

$$y_m = \begin{cases} 1 & \text{if shipment } m \text{ is connected} \\ 0 & \text{otherwise} \end{cases} \quad (4.5)$$

where, x_m is the size (number of cars) of a shipment m , and CAP is the dispatchable train capacity. Thus, at the terminal n , each outbound train is assembled based on the following steps.

Step 1. Compute expected penalty for each car : Compute expected penalty for all cars listed in the take list of a train

Step 2. Determine pull sequence order :

- For cars with positive expected penalty : give higher preference to a car that has a higher expected penalty in pull sequencing for train make-up

- For other cars with no expected penalty : pull cars based on a given take list

Operating difficulties

There exist, however, certain technological difficulties in dynamic train make-up practice. When a car having a higher expected penalty is located in the middle of a classification track, to pull that car requires additional switching. This additional switching requires more switch engine hours and higher car handling costs. The additional switching may also delay the handling of other cars that are to be assembled. For example, car 1 has the highest priority and is located in the middle of a track (between other cars 2 and 3). To pull car 1 first and assemble it to an outbound train requires additional switching (see Figure 4.15).

The simulation model developed in the study simulates the detailed switching operation in the terminal, so that we can examine the additional switching work-load (measured by additional switch engine hours) by implementing a dynamic train make-up policy.

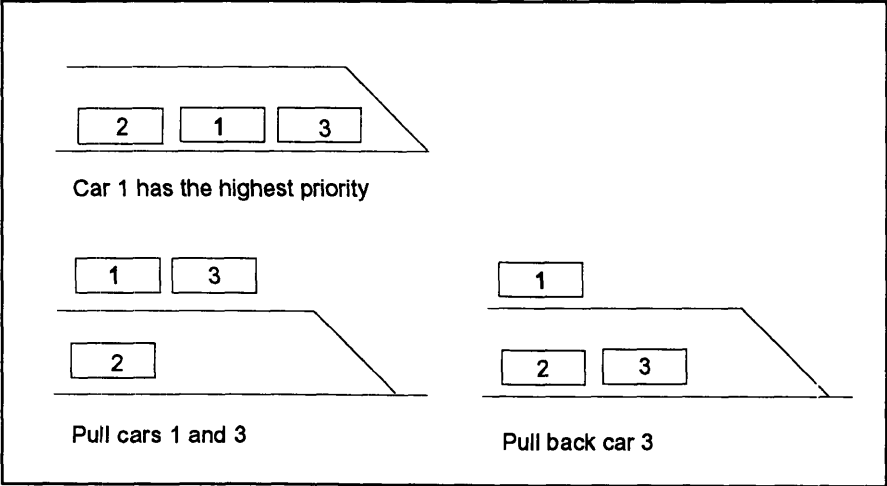


Figure 4.15 : Example of additional switching

4.3.3 Performance measures

As an output of the simulation, three categories of performance were evaluated for all traffic and by priority class.

- Trip time and trip time reliability
- Rail operating costs
- Total logistics costs

Trip time reliability is measured by the variance of trip time and the on-time performance. The rail carrier's total and per car operating costs were evaluated. Train cost, car-time cost and car switching cost were considered as important elements of the rail carrier's total operating cost. The procedure used to compute operating costs and the unit cost information is described in Appendix A. The operating cost by priority class is based on the same heuristic cost allocation method used in Section 4.2.

To see the potential effects on market performance, the shipper's total logistics cost resulting from service differentiation strategies was evaluated. The total logistics cost model was used to predict how cost-minimizing shippers will react as the trip time performance of rail service changes (see Chapter 2.4). A rail carrier can estimate how much shippers will be willing to pay for additional service improvement using the logistics cost model. To compute total logistics costs, additional information was assumed on the value of commodity by priority class. The value of high and low priority classes are assigned to be \$100,000 and \$50,000 per car-load, respectively.

4.3.4 Simulation analysis

Simulation parameters

In addition to demand conditions for the simulation, several parameters for resources and operating conditions were defined.

Resources

The resources for operations were specified as follows.

Locomotives are the essential resource in dispatching planned train capacities. If there are not enough locomotives in the terminal, a shorter train will be dispatched based on the number of locomotives available. The initial inventories of locomotives are: 12 locomotives at terminals 1 and 5; and 24 locomotives at terminals 2, 3 and 4, respectively.

Yard engines are also an important resource for handling cars at the terminal. If there are not enough yard engines in the terminal, car handling times will be increased and some cars will miss their appropriate train connections. The number of yard engines are: 2 engines for terminals 1 and 5; and 4 engines for terminals 2, 3 and 4, respectively.

Input service time

For the simulation of terminal operations, several service time parameters are assumed. It is assumed that both inbound and outbound inspection times of a train are 1 hour regardless of train

length. For the classification and make-up, 30 minutes fixed setup time is assumed. The classification time for each car is assumed to be exponentially distributed and mean service time is 30 seconds per car. The assembly time for each car is also assumed to be exponentially distributed and mean service time is 1.5 minutes per car.³

Experimental design

Several simulation scenarios, having different policies and operating conditions, are examined.

- Coefficient of variation of daily demand distribution : 0.1, 0.2, 0.3, 0.4, 0.5
- Coefficient of variation of train transit time : 0.1, 0.2, 0.3, 0.4, 0.5
- Share of high priority traffic : 0.1 - 0.9
- Two train make-up policies : see section 4.3.2

The simulation result for each scenario is obtained from 20 independent simulation runs.⁴ Each run simulates the 4 weeks of rail operations.

³ These unit service times are realistically determined based on the actual data (September 15 to 22, 1993) of Radnor terminal of CSXT. The average length of inbound and outbound trains were 87 and 66 cars, respectively. The average inbound inspection time was 2 hours; the average classification time was 1 hour and 15 minutes (0.52 minutes/car with 30 minutes fixed set up time); the average assembly time was 2 hours and 20 minutes (1.67 minutes/car with 30 minutes fixed set up time); and the average outbound inspection time was 1 hour and 50 minutes. Since this terminal is one of the busiest terminals of CSXT, we used somewhat shorter unit service times for the simulation.

⁴ See section 4.2.4 for discussions on the minimum sample size.

Simulation results

Effects of demand and train transit time variability

Consider that a railroad implements the static make-up policy to differentiate service among different traffic classes, which control the car-to-train connections by prioritizing cars in pre-determined pull sequencing order for a train make-up at the terminal.

The effects of demand and train transit time variability on the service performance of all traffic were examined. In the aggregate, as the variability in demand or in train transit time increases, the service performance of all traffic deteriorates (Table 4.14 and 4.15). This suggests that reducing the variability in demand and train transit time certainly helps to improve the overall service performance.

Table 4.14 : Effects of traffic variability on service of all traffic :
static make-up policy

Traffic variability (C.V.)	Average trip time (hours)	Std dev of trip time (hours)	Percent on-time (percent)
0.1	28.02	2.53	94.00
0.2	28.38	2.98	93.28
0.3	29.18	3.59	92.13
0.4	30.24	4.39	90.72
0.5	31.39	5.24	89.48

C.V. of train transit time 0.1 and share of high priority traffic 50%

Table 4.15 : Effects of train transit time variability on service of all traffic :
static make-up policy

Transit time variability (C.V.)	Average trip time (hours)	Std dev of trip time (hours)	Percent on-time (percent)
0.1	29.18	3.59	92.13
0.2	29.96	4.61	88.89
0.3	30.64	5.58	83.85
0.4	31.41	6.72	77.84
0.5	32.07	7.79	73.13

C.V. of demand 0.3 and share of high priority traffic 50%

The increase in demand and train transit time variability also increase the per car operating cost (Table 4.16 and 4.17) mainly due to the increase in car time cost. Car time cost per car increases because cars have more chance of missing train connections when either the variability in demand or in train transit time increases. In addition, when the variability in demand increases, it is more likely that a shorter train will be. Since we dispatch the planned locomotives even though the actual train length is shorter than the designed train capacity, the per-car train cost of a shorter train is larger than that of a full train.

Table 4.16 : Effects of traffic variability on per car operating cost :
static make-up policy

Traffic variability (C.V.)	Total cost (\$/car)	Train cost (\$/car)	Car time cost (\$/car)	Car handling cost (\$/car)
0.1	73.79	41.14	28.02	4.61
0.2	74.31	41.30	28.38	4.62
0.3	75.33	41.51	29.18	4.66
0.4	76.74	41.82	30.24	4.47
0.5	78.18	42.06	31.39	4.72

Table 4.17 : Effects of train transit time variability on per car operating cost :
static make-up policy

Transit time variability (C.V.)	Total cost (\$/car)	Train cost (\$/car)	Car time cost (\$/car)	Car handling cost (\$/car)
0.1	75.33	41.51	29.18	4.66
0.2	75.99	41.37	29.96	4.68
0.3	76.18	40.87	30.64	4.68
0.4	76.36	40.27	31.41	4.67
0.5	76.73	39.99	32.07	4.67

The disaggregate performance by priority was examined to develop insights into the effects of service differentiation practices. The analysis of the service performance by priority revealed several interesting results. When either the variability in demand or in train transit time increases, the service performance of all traffic deteriorates. The high priority traffic is, however, “consistently” provided highly reliable service even though the system has more variability in demand and train operations (see Table 4.18 and 4.19).

On the other hand, the service level of low priority class traffic is quite sensitive to the variability of both demand and train transit time. Changes in operating conditions have greatest influence on the service level of low priority traffic. It can be interpreted that, under static make-up policy, the unreliability inherent in a given system is mostly allocated to the low priority traffic. Differentiating service for different classes of traffic can actually be viewed as a process of allocating the unreliability inherent in a given system to different classes of traffic.

Two sample *t*-tests were done to examine whether the difference in trip time performance among different classes of traffic are statistically significant, as in the previous section. Under any level of traffic and train transit time variability, the difference in average and variance of trip time

and on-time performance between high and low priority traffic were statistically significant at significance level 0.05.

Table 4.18 : Effects of traffic variability on service by priority :
static make-up policy

Traffic Variability (C.V.)	Average trip time		Std dev of trip time		Percent on-time	
	High (hr.)	Low (hr.)	High (hr.)	Low (hr.)	High (%)	Low (%)
0.1	26.91	29.12	1.30	3.77	99.99	87.97
0.2	26.91	29.90	1.33	4.63	99.98	86.48
0.3	26.90	31.39	1.34	5.89	99.98	84.14
0.4	26.97	33.66	1.38	7.54	99.94	81.16
0.5	27.02	36.05	1.39	9.32	99.91	78.50

C.V. of train transit time 0.1 and share of high priority traffic 50%

Table 4.19 : Effects of train transit time variability on service by priority :
static make-up policy

Transit time Variability (C.V.)	Average trip time		Std dev of trip time		Percent on-time	
	High (hr.)	Low (hr.)	High (hr.)	Low (hr.)	High (%)	Low (%)
0.1	26.90	31.39	1.34	5.89	99.98	84.14
0.2	27.32	32.66	2.44	6.83	98.55	79.04
0.3	27.62	33.73	3.47	7.73	94.67	72.74
0.4	27.87	35.06	4.57	8.96	89.15	66.19
0.5	28.06	36.24	5.54	10.13	84.27	61.59

C.V. of demand 0.3 and share of high priority traffic 50%

Analysis of cost by priority was done to examine whether cost levels are appropriate to the service provided to each priority class. The heuristics cost allocation method was used to compute

the cost by priority, as in the previous section. The results show that, using this method, high priority service had a higher cost and low priority service had a lower cost (Table 4.20 and 4.21). Under any level of traffic and train transit time variability, the difference in operating cost levels between high and low priority traffic were statistically significant at significance level 0.05.

Table 4.20 : Operating cost by priority under different traffic variability:
static make-up policy

Traffic variability (C.V.)	High priority (\$/car)	Low priority (\$/car)	Total (\$/car)
0.1	76.54	70.98	73.79
0.2	77.05	71.54	74.31
0.3	78.26	72.36	75.33
0.4	80.12	73.30	76.74
0.5	81.93	74.31	78.18

Table 4.21 : Operating cost by priority under different train transit time variability :
static make-up policy

Transit time variability (C.V.)	High priority (\$/car)	Low priority (\$/car)	Total (\$/car)
0.1	78.26	72.36	75.33
0.2	80.80	71.13	75.99
0.3	82.38	69.83	76.18
0.4	83.19	69.37	76.36
0.5	84.02	69.25	76.73

By using an operating plan that properly considers the service requirements of different shippers, a railroad can provide fast and very reliable service to service-sensitive shippers, and

slower and less reliable service to price-sensitive shippers. Furthermore, a railroad can “consistently” provide fast and highly reliable service to service-sensitive shippers (that move 50% of the traffic in this case) regardless of the variability that a system has (e.g., variability on demand, train operation or service time in terminal operations).

A railroad may differentiate on price for different priority service by providing different levels of service at different costs, aiming to make a profit on all priority classes. High priority traffic can be provided fast and highly reliable service at a higher price. Low priority traffic can be provided slower and less reliable service at a lower price.

Effects of traffic mixes

To examine what portion of traffic can consistently be given highly reliable service, a sensitivity analysis on traffic mixes is done by changing the shares of high and low priority traffic. Table 4.22 shows that high priority traffic is provided a highly reliable service until its share reaches 70% of total traffic. When the high priority traffic is beyond this share, however, its service level also begins to deteriorate.

It also shows that the service level of low priority traffic significantly deteriorates as the share of high priority traffic increases. In addition, Figure 4.16 shows that the difference between the service levels of high and low priority traffic becomes larger as the share of high priority traffic increases. For any share of high and low priority traffic, the difference in trip time performance between high and low priority traffic were statistically significant at significance level 0.05.

Table 4.22 : Effects of traffic mix on service by priority :
static make-up policy

Share of high priority	Average trip time		Std dev of trip time		Percent on-time	
	High (hr.)	Low (hr.)	High (hr.)	Low (hr.)	High (%)	Low (%)
0.1	26.99	29.61	1.38	3.81	99.98	87.99
0.2	26.94	29.89	1.37	4.08	99.96	87.34
0.3	26.97	30.18	1.36	4.43	99.96	86.43
0.4	26.96	30.69	1.36	4.99	99.97	85.60
0.5	26.90	31.49	1.34	5.89	99.98	84.14
0.6	26.94	32.37	1.34	7.03	99.91	82.06
0.7	26.99	33.98	1.44	8.86	99.60	77.98
0.8	27.21	36.72	1.73	12.31	98.35	72.61
0.9	27.95	41.09	2.49	17.08	94.47	65.39

C.V. of demand 0.3 and C.V. of train transit time 0.1

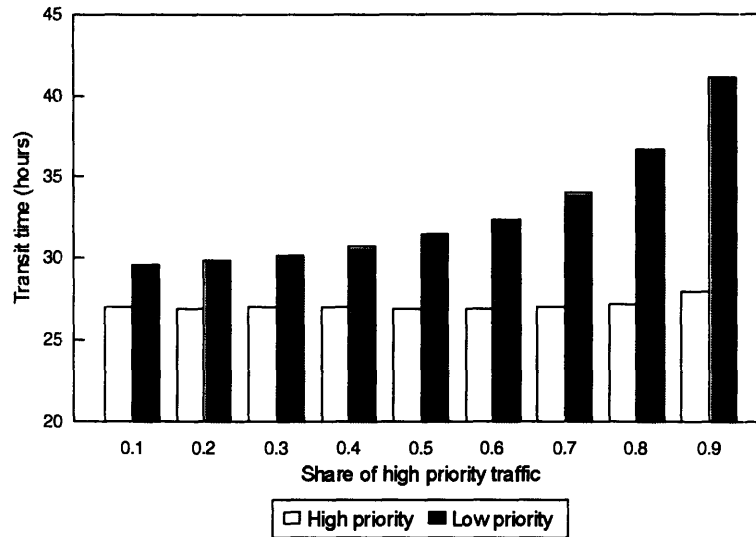


Figure 4.16 : Transit time by priority under different traffic mix

Static vs. dynamic train make-up policies

A dynamic train make-up policy option is designed to reduce excessive delays of low priority traffic while maintaining the desired level of service to high priority traffic.

This policy was examined and we found that it improves the service performance of low priority traffic, at the same time that it worsens the service performance of high priority traffic (see Table 4.22 and 4.23). Under the dynamic train make-up policy, for any share of high and low priority traffic, the difference in trip time performance between high and low priority traffic were statistically significant at significance level 0.05.

Table 4.23 : Effects of traffic mix on service by priority :
dynamic make-up policy

Share of high priority	<u>Average trip time</u>		<u>Std dev of trip time</u>		<u>Percent on-time</u>	
	High (hr.)	Low (hr.)	High (hr.)	Low (hr.)	High (%)	Low (%)
0.1	28.30	30.27	2.38	4.33	94.06	86.99
0.2	28.36	30.64	2.44	4.64	93.92	85.76
0.3	28.57	30.96	2.58	5.04	93.02	84.57
0.4	28.75	31.44	2.73	5.61	92.16	83.28
0.5	29.00	31.87	3.03	6.07	90.93	81.82
0.6	29.51	32.28	3.45	6.61	88.82	79.50
0.7	30.07	32.90	3.86	7.38	86.20	76.76
0.8	30.58	33.54	4.31	7.93	83.28	73.91
0.9	31.68	34.16	5.24	8.03	79.07	70.11

C.V. of demand 0.3 and C.V. of train transit time 0.1

When the share of high priority traffic is low, the low priority traffic has less chance to miss train connections since there is more available train capacity to low priority traffic. In this case, the dynamic make-up policy was not quite effective in reducing the excessive delays of low

priority traffic. As the share of high priority traffic increases, the dynamic make-up policy becomes more effective in reducing excessive delays of low priority traffic. For any traffic mixes, the difference in trip time performance of high priority traffic between static and dynamic make-up policies were statistically significant at significance level 0.05. However, the difference in trip time performance of low priority traffic between two policies became statistically significant after the share of high priority traffic reached about 70% of total traffic.

The comparison of the two train make-up policies suggests that varying the way that the different service requirements of traffic are incorporated into a train make-up policy (i.e., static and dynamic policies) results in different levels of service to different classes of traffic.

An important question is how much distinction should be made between high and low priority shippers. Is there value in reducing excessive delays and improving the service level of low priority traffic? The shipper's total logistics cost is computed as a measure of the potential willingness of shippers to pay for the changes in service performance.

Figure 4.17 shows that the per car logistics cost for all traffic under the dynamic make-up policy is larger than under the static make-up policy. For a specific case we examined (C.V. of demand 0.3 and C.V. of train transit time 0.1), the dynamic make-up policy is not an effective policy option for the purpose of service differentiation. It may suggest that a policy option that requires more frequent car handling at the terminal level may not be effective since it will increase the car handling costs and times at the terminal. As the share of high priority traffic becomes smaller, the difference in per car logistics between the static and dynamic make-up policies becomes smaller.

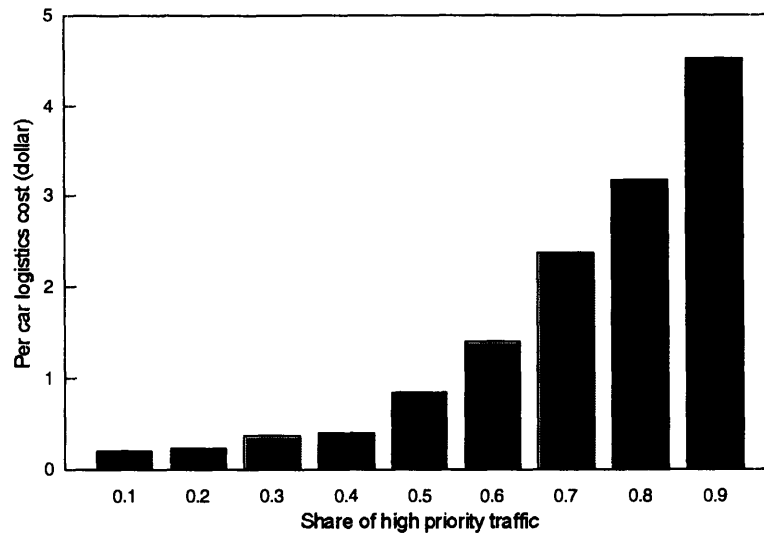


Figure 4.17 : Difference in per car logistics cost between static and dynamic policies: under different traffic mixes

We did another sensitivity analysis on the shipment values of high and low priority traffic. Initially, we assumed the value of the high priority (\$100,000/car-load) to be twice that of the low priority shipment (\$50,000/car-load). As the difference in shipment value between high and low priority shipments becomes smaller, the difference in per car logistics cost between the static and dynamic make-up policies also becomes smaller (Figure 4.18), suggesting that the dynamic make-up policy may become more effective when the share of low priority traffic is large and the difference in shipment value between high and low priority shipments is small (i.e., traffic is relatively homogeneous).

4.3.5 Findings

The analysis of the train service among multiple terminals reveals some important insights on the effects of service differentiation.

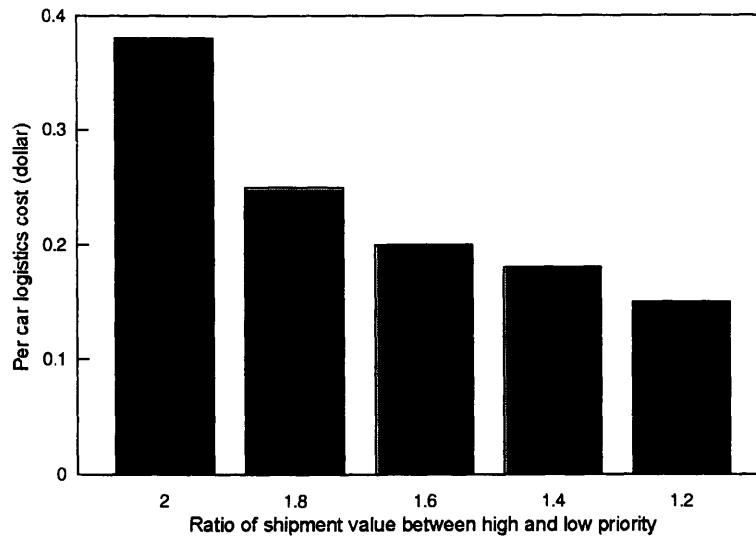


Figure 4.18 : Difference in per car logistics cost between static and dynamic policies : under different ratio of shipment value between high and low priority

- In the aggregate, as the variability in demand or train transit time increases, the service performance of all traffic deteriorates and the per car operating cost is increased.
- Under the static make-up policy, the high priority traffic is “consistently” provided a fast and highly reliable service, even though variability in demand and train operations increase.
- The changes in variability in demand and train operations have the greatest influence on the service level of low priority traffic. It can be interpreted that, under the static make-up policy, the unreliability inherent in a given system is primarily allocated to the low priority traffic. Differentiating services for different classes of traffic can be viewed as a process of allocating the unreliability inherent in a given system to different classes of traffic.
- High priority traffic is provided fast and highly reliable service, although its share becomes relatively high. When the share of high priority traffic goes beyond a certain point, however,

its service level begins to deteriorate. This suggests that a given system will be limited in its capability to provide premium service to shippers.

- Implementing different train make-up policies result in different levels of service to different classes of traffic.
- By using an operating plan that properly considers different service requirements of different shippers, a railroad can provide fast and very reliable service to service-sensitive shippers and slow and less reliable service to price-sensitive shippers. Furthermore, a railroad can “consistently” provide fast and highly reliable service to service-sensitive shippers regardless of a certain variability that a system has (e.g., variability on demand or train operation) with limits, provided that the service-sensitive traffic is not a large fraction of the traffic.
- A railroad may differentiate on price for different priority service by providing different levels of service at different costs, and make profit on all priority classes. High priority traffic can be provided fast and highly reliable service at a higher price. Low priority traffic can be provided slower and less reliable service at a lower price.

4.4 Conclusions

Two probabilistic simulation models were designed and analyzed to develop insights into the effects of service differentiation in the rail freight transportation context.

- Two terminal model where a railroad provides a direct train service
- Five terminal model where a railroad provides car-load train services that require intermediate classification works

As a mechanism to differentiate services, train make-up policies at the terminal, which are dependent on the shipper supplied shipment priority, are examined. These train make-up policies were examined under various resources defined by train capacity and empty car inventory level, and operating conditions defined by the variability of demand and train operations. Various measures for service and cost were evaluated by traffic class.

The results show that varying the way the heterogeneity of traffic is incorporated into an operating scheme results in different levels of service to different classes of traffic. A policy option that improves the service performance of all traffic may not be effective where there are different classes of traffic that have different service requirements. By using an operating plan that properly considers the heterogeneity of traffic, a railroad can provide fast and very reliable service to service-sensitive shippers and slower and less reliable service to price-sensitive shippers. Furthermore, a railroad can “consistently” provide fast and highly reliable service to service-sensitive shippers regardless of a certain variability that a system has (e.g., variability on demand, train operation or service time in terminal operations) with limits, provided that the service-sensitive traffic is not a large fraction of the traffic.

A railroad also can utilize the existing service capacity more efficiently, by focusing resources on the improvement of the service performance for selected market segments that are sensitive to service and willing to pay for these additional service improvements, instead of making significant additional investment to improve the service performance for all traffic. Furthermore, a railroad may improve its profitability by differentiating prices for different market segments based on distinct service levels and the cost of providing such service, and still make a profit on all priority classes.

In conclusion, service differentiation strategies enable a railroad to provide market-sensitive services and to utilize service capacity more efficiently by avoiding additional investment to increase service capacity to improve the service performance for all traffic classes that may not require or be willing to pay for higher quality service.

In this chapter, the train make-up policy was examined as a mechanism for differentiating services for different classes of traffic. There are many other operating plans in which the heterogeneity of traffic can be incorporated (e.g., blocking plan, make-up plan, car scheduling, etc.). In the next two chapters, we examine how to improve the ability of railroads to differentiate service for different classes of traffic by incorporating the heterogeneity of traffic into their operating plans.

CHAPTER 5

Hierarchical Structure of Decisions in Rail Operation and Models for Operating Plans: A Review

5.1 Introduction

In the last chapter, we developed some useful insights on the effects of service differentiation in the rail freight transportation context. It was found that service differentiation can potentially be a very effective strategy for a railroad in gaining market by providing market sensitive services and in enhancing profit. It was also found that a heterogeneity of traffic and varying the way that such heterogeneity is incorporated into operating plans results in different levels of service to different classes of traffic.

With the recognition of heterogeneity of traffic, a railroad needs to properly incorporate such heterogeneity in developing operating plans. In this chapter, a hierarchical structure of decisions in rail operations and previous modeling efforts for each decision area are reviewed and how to incorporate the heterogeneity of traffic into various operating plans is discussed.

5.2 Service Design Process Incorporating Heterogeneity of Traffic

To determine how to design and produce differentiated service products that meet service requirements for different market segments, we need to study how to incorporate the logistics needs of different market segments into the service management process. Martland, et al. [1993] discussed important elements for the improvement of rail freight service management. This section focuses on how to incorporate market segmentation and service differentiation concepts into a service design process.

Figure 5.1 lists the basic elements of service design. The whole process of service design is driven by the objectives pursued by senior rail management. Recent industry-wide efforts to improve service reliability have shown that the objectives of rail management have been more driven by service-focused objectives than before.¹

The first steps in the service design process are to assess the potential demand and to determine the service requirements of shippers. Under market segmentation concepts, a railroad needs to assess the demand and service requirements for specific market segments. A knowledge of the elasticity of demand to elements of service in different market segments will be important in trading off trip times, reliability, price and other elements of service. Based on trading-off different elements of service, a railroad can establish the service standards to be used with different market segments. A railroad then needs to design and produce service products to meet the service standards established for different market segments.

Service differentiation can be defined as a strategy to provide different levels of service to different market segments. The underlying concept is that the marketing department can identify

¹ Current research initiatives by AAR on the service reliability and interline service management partly reflects this trend.

different classes of customers, who have different requirements with respect to cost, transit time and reliability. The operating department can then adjust its operations to provide the desired level of service to each class of customers, with cost levels that are appropriate to those services.

The next steps in the service design process are the development of the capacity plan and operating plan. The capacity plan is the set of investment decision for equipment and infrastructure that can provide adequate resources for implementing the operating plan. Failure to provide adequate resources for implementing the operating plan will lead to significant service problems. Development of a capacity plan is relatively a medium- or long-term decision; and development of operating plan is relatively a short-term decision.

Suppose a railroad develops a plan for equipment and infrastructure and defines a certain service capacity. The operating plan under a service differentiation strategy determines the allocation of this service capacity to heterogeneous traffic from different market segments through various operating procedures. The railroad needs well-defined operating plans that take into account the different service requirements of market segments. The hierarchical structure of decisions in rail operation make this process more difficult as compared with other modes. This issue will be discussed in detail in the next section.

The railroad needs to evaluate the level of service and costs that result of implementing the operating plan. If the predicted level of service does not meet the service requirements of different market segments, the railroad can repeat the service design process entirely or in part.

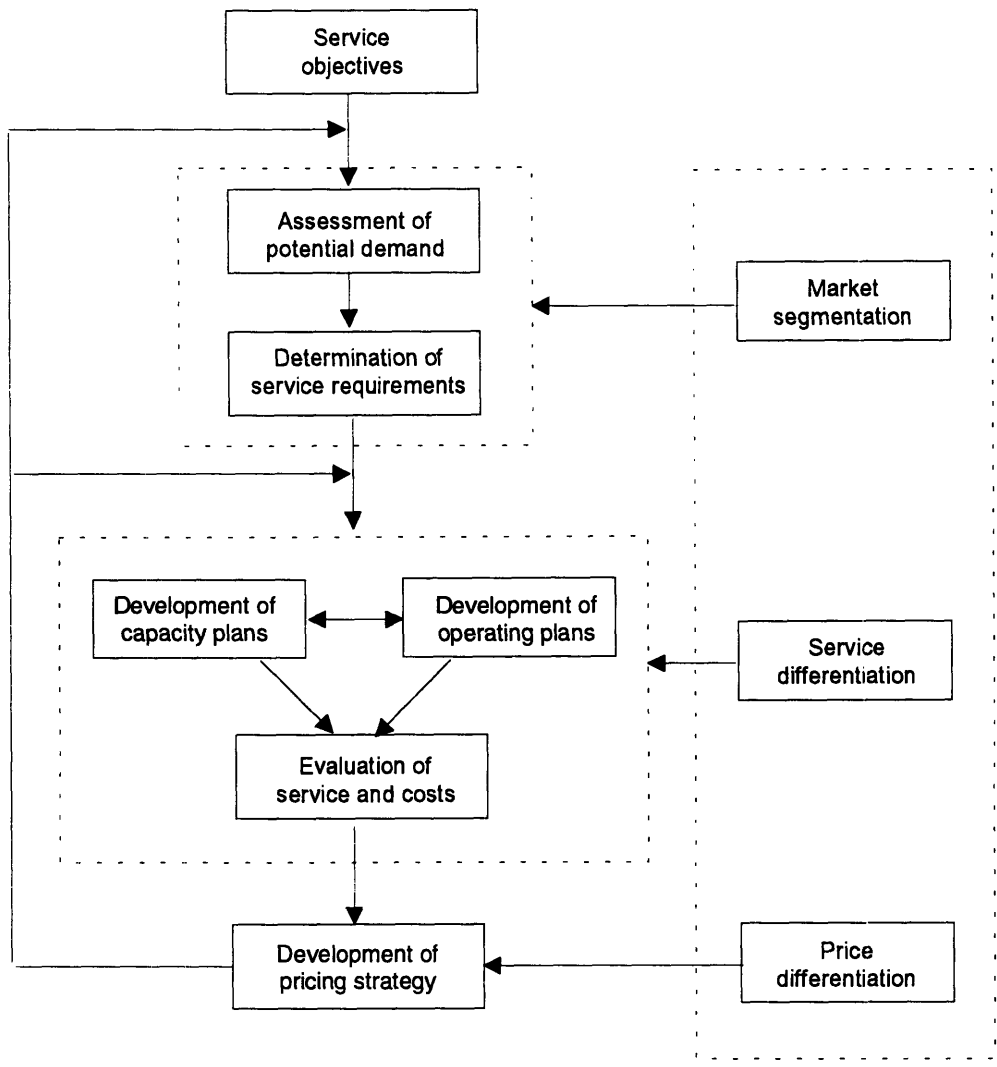


Figure 5.1 : Service design incorporating market segmentation

5.3 Hierarchical Structure of Decisions in Rail Operation

Assad [1980] classified planning decisions in rail operations into strategic, tactical and operational decisions based on the planning horizon, investment requirements, and the level of decision-making. Strategic decisions involve resource acquisition over a long time horizon and typically require major capital investments. Due to the pervasive and long-lasting impact of strategic decisions on the future of the system, top-level management is usually directly involved with their resolution.

Tactical decisions have medium-term planning horizons and focus on effective allocation of existing resources, rather than on major acquisitions. The planning horizon and level of aggregation in model addressing tactical decisions must allow it to take account of broad changes in system parameters and data (such as seasonalities in the traffic volumes and imbalances resulting from lack of uniformity in the geographical pattern of shipments) without having to incorporate day-to-day changes in the data-base.

Operational decisions deal with day-to-day activities in a fairly detailed and dynamic environment. Correspondingly, only lower levels of management (e.g., yard masters) are directly concerned with operational issues. This also includes real-time control decisions needed to manage the uncertainty in real-time rail operations (e.g., train delay due to a train failure). Figure 5.2 shows the elements of hierarchical decisions in rail operations.

Tactical decisions in rail operations include the mid-term decisions, such as train routing, classification, train makeup and traffic routing policies. These decisions are also important inputs for the operational decisions. Train routing decisions deal with the establishment of the service network and the level of service by determining what train services to run, on what routes and with what frequency. Classification policy deals with what blocks to build at each yard and what cars to put into each block (i.e., blocking plan). It also deals with choosing the yards for which blocks

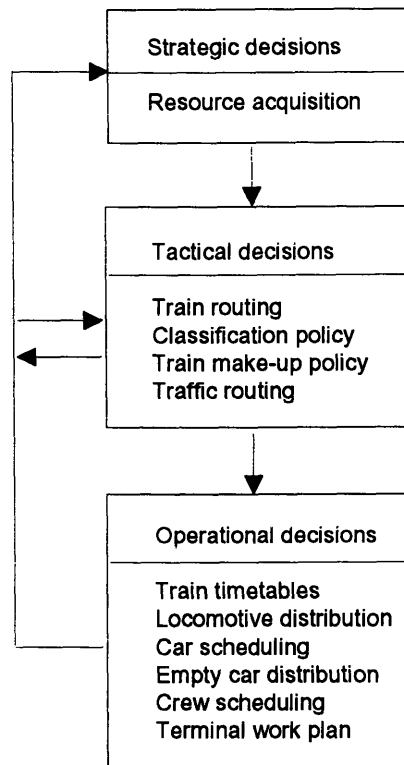


Figure 5.2 : Hierarchical decisions in rail operations

may be formed, and how to distribute the workload among the yards of the system to fully use their capacities and avoid congestion. Train makeup policy deals with what blocks will be carried by each train. Traffic routing determines, for each origin-destination pair, the routing pattern of its traffic (i.e., train sequence used to travel and the yards where the cars are to be classified). Since these decisions are closely related to each other, modeling efforts have been developed to combine a decision model for one or more of these tactical decisions.

Operational decisions in rail operations include short-term decisions such as train timetables, track scheduling and priority policy, locomotive distribution, car scheduling, empty car distribution, terminal work plans, crew scheduling, and maintenance operations. Tactical decisions

will provide important input for operational decisions. After a railroad determines the train route for all train services, it further needs to establish the train schedule for each train. Train timetables can be defined as the set of arrival and departure times at yards, along with train route of each train. Track scheduling and priority policy deals with assigning trains to tracks if track capacity is limited, including the planning for meets and passes according to a priority scheme. Locomotive distribution deals with planning for daily distribution of locomotive over a specified set of train schedules.

The car scheduling system determines how a car is supposed to move from its origin to its destination. The car schedule (or trip plan) information includes the pickup time, the sequence of trains with arrival and departure times, the terminals at which it will be classified or interchanged, and the estimated arrival time at the destination. Car scheduling is becoming a more important part of the operating plan as railroads pursue more scheduled and planned operations, and as shippers demand car schedule information for a seamless procurement, production and distribution plan to improve the logistics process.

Car distribution deals with the problem of distributing or repositioning empty cars over the rail network to meet demand and to rectify imbalances of uneven freight movement. Terminal work plan deals with how to allocate resources to various train and car handling procedures such as receiving, inbound inspection, classification, train make-up, outbound inspection, and maintenance work in the terminal. Crew scheduling deals with how to schedule road and yard crews over the period of time considering both the train and yard work schedule and the collective agreement with the labor union.

A complexity in the development of operating plans for tactical and operational decisions in rail transportation comes from its system characteristics that they need to plan the movements of several distinct entities (i.e., trains, loaded and empty cars, locomotives and crews) on a common

physical network with lines and terminals to produce the service products that meet the service requirements of different shippers.

There are several important issues that need to be considered in the development of the operating plan. The first is the proper objective function to be considered in the development of the operating plan. As we will review, the majority of previous efforts to develop operating plans have focused on cost improvement rather than on service improvement. Some studies that considered service have mainly focused on improvement of average trip time. Recent surveys have revealed that shippers perceived that trip time reliability is much more important than average trip time (Mercer [1991], Intermodal Index [1992, 1993]). This service reliability issue needs to be incorporated into the development of the operating plan.

Other studies have shown that there are a number of market segments in the freight market which have different preferences on factors of service quality. These studies concluded that it is not possible to implement any single marketing and service program that will satisfy all current and potential shippers (McGinnis [1978] and Vieira [1992]). It indicates that, in addition to considering the service reliability issue, railroads need to properly consider different service requirements of different shippers in the development of the operating plan.

In the next section, the previous modeling efforts will be reviewed for each category of hierarchical decision in rail operations. Further research needs will be discussed.

5.4 Review of Modeling Efforts in Developing Operating Plans

There have been continuous efforts to build decision support tools for developing operating plans. Assad [1980, 1981] did a comprehensive review on the early modeling efforts before 1980 for

planning and evaluating yard, line and network operations of rail transportation. In this section, we will review previous modeling efforts in the development of operating plans for tactical and operational decisions in rail transportation.

5.4.1 Models for tactical decisions

There have been two distinct approaches to developing an operating plan. One approach is to redefine the entire operating plan (i.e., “start with a clear sheet of paper”). Most optimization based models belong to this category. Many optimization-based models for tactical decisions have been proposed. Some models address each element of tactical decisions. Other models address combined tactical decisions.

A railroad seldom redefines its entire plan, but frequently modifies the plan to take advantage of marketing opportunities or adjust the plan to current traffic flows. The second approach is to examine and adjust the current plan for incremental improvement. Simulation models, the Service Planning Model (SPM) and the Automatic Blocking Model (ABM) belong to this category. Simulation models and the SPM can also be used as evaluation tools for the plan that was generated from an optimization-based model.

Blocking models

Bodin, et al. [1980] developed a large-scale nonlinear mixed-integer programming model for blocking plan. In the formulation, they considered constraints to 1) limit the minimum and maximum block size, 2) limit the maximum number of blocks that can be made at each yard, 3)

force all cars leaving the same yard with the same destination to use the same block, and 4) limit the total number of cars that can be classified at each yard.

The objective function attempts to take into consideration the cost of handling a car at each yard, the cost of placing a car into a particular block, and the delay costs associated with the classification of a car for a particular block at a particular yard. In a test case of 33 yards based on the Norfolk and Western Railroad, the problem was a mixed integer problem with about 6,500 constraints and 11,000 variables (6,000 0-1 variables). This problem was too complex to be solved by MPSX/370. They relaxed the integrality conditions and solved the problem in iterative manner to obtain the final solution that was within 3% of a tight lower bound.

The Automatic Blocking Model (ABM), developed by ALK Associates, is a heuristic blocking model that focuses more on incremental changes to the plan rather than redefinition of the entire blocking plan (see Van Dyke [1986, 1988]). The ABM uses an iterative heuristic that attempts to improve a blocking plan by using a series of incremental changes starting with a blocking plan supplied by the user, or generated using a set of decision rules. It relies heavily on a shortest path routine to route both blocks and flows. All blocks will be routed on the lowest cost path from their origin to their destination. Each flow then will be routed across the available blocks using the lowest cost combination of blocks subject to block capacity constraints. The advantage of the ABM is that it explicitly considers the priority of cars. It allows a railroad to evaluate various options for service differentiation such as pre-blocking and single priority block. The ABM has been implemented and used by a number of railroads (e.g., CSXT and Norfolk Southern).

Combined tactical planning models

There have been several attempts to model combined tactical decisions. Assad [1980] proposed a model for combined decisions for train routing, traffic routing and train makeup. He formulated it as a nonlinear mixed-integer multicommodity flow problem. The decision variables are frequency for each train service and car flow for each train service. He considered a train capacity constraint to limit the total number of cars on specific train services. In the model, it is assumed that each yard builds blocks only for adjacent yards that can be directly reached by trains departing from the yard. His model thus could not consider the pre-blocking option. The objective function attempts to take into consideration the train cost, the cost of handling cars at each yard, and the delay cost associated with yard congestion.

Crainic, et al. [1984, 1986] developed a comprehensive model for combined decisions of train routing, traffic routing, classification and train makeup. They formulated it as a nonlinear mixed-integer multicommodity flow problem and solved it using a heuristic decomposition technique. The decision variables are train frequency of each train service and car flow on each route. Pre-blocking was considered by including the block swap option in enumerating possible routes. Train capacity constraints to limit the total number of cars on specific trains were also considered. They explicitly considered the heterogeneity of traffic in terms of traffic classes (defined by origin, destination and commodity). Different car time costs are specified for different traffic classes.

The objective function attempts to take into consideration the train cost, the cost of handling cars at each yard, and the car delay cost associated with yard congestion. They used a $M/M/1$ queuing model to consider the relation between yard congestion and delay incurred, and the line delay was considered as a function of train frequency. This work suggests that the solution

can provide the combined information for train routing, traffic routing, classification and train makeup policies. Train routing and traffic routing policies are directly obtained from the solution. Classification and train makeup policies can be established based on the solution. The model and algorithm were tested on a sub-network of the Canadian National (CN).

Keaton [1989, 1992] developed a model for combined decisions of train routing, traffic routing and classification. He formulated it as a linear mixed-integer multicommodity flow problem. In the formulation, he considered constraints that 1) limit the total number of cars on a specific train and 2) limit the maximum number of blocks that can be made at each yard. The objective is to minimize the sum of train cost, car handling cost and car time cost. He assumed a fixed yard process and delay time instead of a function that included congestion effects. By relaxing the train capacity constraints, he obtained a tight lower bound of the optimal solution using a Lagrangian relaxation technique. He tested the model and algorithm on a hypothetical rail network based on a sub-network of the Consolidated Rail Cooperation (ConRail).

Haghani [1989] developed a model for combined decisions of train routing, traffic routing, train makeup and empty car distribution. He attempts to model the tactical train routing and makeup problem in a more dynamic manner by considering traffic variability, and to model the operational car distribution jointly with the tactical decisions. It is the only study that considered the traffic variability among optimization based studies. To capture the effects of temporal traffic pattern, he used the concept of a time-space network. Assuming a temporal but known traffic pattern for the analysis period, he formulated it as nonlinear mixed-integer programming problem. A heuristic decomposition technique was developed to solve the problem. He tested the model and algorithm on a small hypothetical network.

Simulation

Simulation has the capability of representing the complex operating system that can be found in rail operations; this can be a useful technique to evaluate alternative operating plans. General purpose simulation programs were developed by many railroads in the 1970s (e.g., the AAR model [1971]). However, those models were not successfully implemented due to the high computational costs and large data requirements. Large amounts of data is required to calibrate the various model parameters.

Recently, as railroads have focused on the improvement of service reliability, they are seeking more sophisticated evaluation tools that can consider the dynamic and stochastic nature of rail operations. More interest has recently emerged in developing a sophisticated simulation model. A stochastic rail network simulation model can directly incorporate the variability of traffic and train transit time as input distributions. The rail network simulation model discussed in the Chapter 4 and Appendix B can be an example of this approach.

Service Planning Model (SPM)

The Service Planning Model (SPM) was developed to evaluate alternative operating plans with less computational efforts than the simulation model, but still maintaining realistic consideration of rail operations (see McCarren and Martland [1980]). Rather than simulating the movement of cars through the network, the model uses analytical techniques to estimate trip times. The model's major asset is its ability to provide a negotiation framework among various decision making entities (e.g., marketing and operating departments) in developing operating plans.

The model produces detailed service performance, including O-D trip time distribution. The principal input to the model are train schedules, blocking information and probabilistic train connection standards represented as PMAKE function. The detailed description of the PMAKE function can be found in Martland [1982]. The model uses this information, with logic similar to that used in car schedule building, to predict O-D and yard time distribution. Several major railroads have implemented the SPM to evaluate the operating plan and establish service standards under existing or new operating plans.

5.4.2 Models for operational decisions

Operational decisions include short-term decisions such as train scheduling, car scheduling, train control, engine scheduling, empty car allocation, terminal work plans, and crew scheduling.

Tactical decisions will provide important input for operational decisions. Compared with the modeling efforts for tactical decisions, studies on operational planning models explicitly considered the variability of traffic and train operation, and service reliability issues.

Train schedule and timetables

Railroads use train schedules for their train operations. However, it is hard to find literature on how they have developed and implemented train schedules. Assad [1982] studied the stop-schedule problem for single train and two trains on the line of n yards. Keaton [1992] studied the effects of distribution of train arrival and departure times on the average yard time for a single yard.

ALK Associates developed the Train Scheduling System (TSS) for the design, evaluation and maintenance of a set of train schedules (see Van Dyke and Davis [1990]). A train schedule is

defined as having three components: train route, a set of block-to-train assignments, and a set of departure/arrival times. This software is designed to support the modification or redefinition of train routes, train make-ups and train time-tables, given the blocking plan provided by the user.

Train dispatching and train control systems

Train dispatching problems have been extensively studied in both industry and academia. The centralized traffic control (CTC) system has been used for many decades by North American railroads. Train dispatchers on a centralized traffic controlled line control the movement of trains over a line, including the planning of where meets and overtakes are to occur and the alignment of the switches to control each train movement. Recently, the North American railroad industry is beginning to implement new, advanced train control technologies that will significantly change railroad operations. Collectively, these technologies are referred to as the Advanced Train Control System (ATCS). The Burlington Northern Railroad (BN) designed, tested a specific version of ATCS that is called the Advanced Railroad Electronics System (ARES) for computer-assisted train dispatching in real-time operations, but this system was not adopted (Smith and Resor [1991]). The evolution of computer-assisted train dispatching is discussed by Petersen, et al. [1986]. The case study of implementing the computer aided train dispatching system can be found in the paper of Sauder and Westerman [1983].

Most train dispatching studies have focused on finding the optimal train dispatch for single and/or double track lines between two yards. Several mathematical models have been proposed to find the optimal train dispatching plan. Early studies on train dispatching used a simulation model (see Bongaardt, et al. [1980]). Kraft [1987] proposed a train dispatching model that attempted to minimize the total train delays. He used a branch and bound procedure to find the optimal train

dispatching solution. Jovanovic and Harker [1990] developed an optimization-based train dispatching model as part of developing a computer-aided train dispatching system for a major US railroad.

Receiving and dispatching policies in yards

As train dispatching and train control are designed to improve the reliability and safety of train operations in line, yard receiving and dispatching policies are designed to improve the reliability of train and car connections in yard operations. Yard receiving policy determines a priority scheme for processing the queue of trains entering to a classification yard. When a large number of inbound trains are in the receiving yard, a set of rules is needed to determine the sequence of inspection and classification of inbound trains, considering the connection plan between inbound trains and departure trains, as well as certain service priorities. Yagar, et al. [1983] studied several models to determine hump sequence. They tested the model for the Taschereau Yard in Montreal of the Canadian National.

Yard dispatching policy can be defined as controls on train operations that determine when to makeup and dispatch outbound trains based on train schedules, train length limits and other operating factors; it also determines whether or not to cancel outbound when there is not enough traffic. Folk and Sussman [1974] used a simulation model to evaluate the effects of train arrival variability and yard dispatching policy on trip time reliability.

Engine scheduling

Florian, et al. [1976] developed a model to schedule several engine types considering the power requirement and schedule of trains for the Canadian National. They formulated it as a multicommodity flow problem on a time-space network. The objective function was to minimize the total train cost. Bender's decomposition approach was used as a solution algorithm. The developed model, however, was not used to allocate engines to trains on a daily or weekly basis. Rather, it was used to evaluate the classes of trains which are expected to operate in the near or distant future and to estimate their power requirement. Smith and Sheffi [1989] developed a model for the day-to-day power allocation problem considering uncertainty in power requirements. They formulated the problem as a multicommodity flow problem with convex objective function on a time-space network, and solved it using a heuristic algorithm.

Several railroad have developed and implemented the computerized locomotive management system. Chih, et al. [1990] reported on the development and implementation of a real time locomotive distribution system for the Union Pacific (UP). Hornung, et al. [1990] reported on the development of a prototype motive power information and management support system for the Canadian National.

Car distribution

Numerous models of the car distribution process have been developed. Misra [1972] formulated the empty car distribution problem as a transportation problem. Ratcliffe, et al. [1984] developed a simulation model for the prepositioning of empty freight cars. A car allocation algorithm, formulated as a transportation problem, is embedded in a simulation model. These two models

assume that supplies, demands, and travel times are known with certainty and only empty car hours are considered.

Philip [1978] applied inventory control concepts to the problem of sizing empty car inventory from an individual terminal perspective. Based on a simulation model incorporating inventory control concept, he estimated the optimal empty car inventory level as a function of supply and demand variability and the cost of holding empty cars relative to the cost of failing to provide empty cars when required by shippers.

Jordan and Turnquist [1983] developed a dynamic network optimization model that considers the stochastic and dynamic nature of the car distribution problem. In the model, the supply and demand of cars in future periods, as well as their travel times over the network, are represented stochastically to reflect the uncertainty of future conditions. They modeled the distribution of homogeneous cars. The model generates daily flows of empty cars between yards that maximize expected profit.

Chih [1986] and Adamidou, et al. [1993] studied the car distribution problem of multiple railroads with consideration of multiple ownership over multiple time periods. Chih formulated it as a multicommodity flow problem on a time-space network. The objective function attempts to minimize the cost of moving loaded and empty cars for each railroad. They view the car distribution problem from a user optimal point of view rather than a system optimal perspective. The subgradient algorithm was used as a solution algorithm. A case study was done for the car distribution problem during a 15 day period for a selected car type on the network of Southern Pacific (SP), Union Pacific (UP) and Missouri Pacific (MP). Adamidou, et al. view the problem as an N -person noncooperative, nonzero sum game played on a temporal-spatial network. An iterative sequential solution approach was developed and implemented for the case study.

Empty flow management is important for all levels of transportation and logistics systems planning. Major research efforts can also be found in other systems such as truck and container transportation. Dejax and Crainic [1987] did a comprehensive review on the models for empty flows and fleet management in freight transportation.

5.4.3 Conclusions

Table 5.1 and Table 5.2 summarize and compare several important characteristics of practically- and theoretically-developed models for hierarchical decisions for rail operations.

Several common characteristics can be found in previous optimization based tactical planning models. The majority of models considered the rail network operation as a relatively static problem by considering average demand and average train transit time. This is partly because these models consider the nature of tactical plans that support relatively medium-term decisions. Second, the majority of models considered the yard as a “black box” and assumed a certain fixed yard process and delay time. Finally, some of the previous models considered the service reliability issue, but no previous model fully considered the need for establishing different service standards for different shipper groups. The majority of models dealt with homogeneous traffic. This aspect combined with the cost minimization objective generates the design of services that are not sensitive to market service requirements.

On the other hand, the simulation and SPM approaches can explicitly incorporate the stochastic and dynamic nature of rail operations. The simulation approach, in principle, is capable of simulating detailed yard operation. On the other hand, the SPM incorporates the probabilistic yard process and delay time using the estimated PMAKE functions. However, these approaches

can not generate the operating plan by itself. Further research efforts are needed to develop a more robust tactical plan that considers the stochastic nature of rail operations such as variability of demand, transit time, and yard process and delay time.

Compared with the tactical planning models, the majority of operational models explicitly considered the variability of demand and train transit time. Most of the research efforts have focused on train operation, including train scheduling and train control. The empty car distribution problem also has been studied extensively.

Little research can be found in the car scheduling area. The car schedule, in practice, is generated from the train schedule, blocking plan, cutoffs and block-to-train assignments. However, these inputs only define a feasible set of schedules for car moves of each O-D pair. To plan car moves to improve service reliability and to provide market sensitive services, a railroad needs to develop a car scheduling system that also incorporates the service standards of different shipper groups, more realistic train connection performance at yards, and the variability of demand. Yet, no models to support improved car scheduling practice are available.

Another important research area is to develop models for developing terminal work plans. Some of the previous research on yard receiving and dispatching policies addressed the terminal work plan issues. However, they focused more on the train control issue rather than the terminal plan issue. Terminal work plans can be established based on an understanding of the elements and process of detailed terminal operations and controls. Martland, et al. [1993] did conceptual work to develop the terminal control system in conjunction with the advanced line control system.

From the implementation point of view, optimization-based models have not been widely implemented in the rail industry. The majority of railroads are still using heuristic models and simulation models for planning operations. Because those models are commercially available,

railroads can purchase them without investing time and resources in developing more sophisticated decision support tools.

The emergence of new information technology allows new opportunity to implement more sophisticated tools for planning operations. Incorporating advanced decision technology in developing decision support models are also important issue to be studied.

In this chapter, an extensive literature review on previous modeling efforts in developing operating plans for tactical and operational decisions to improve operating and service performance of the rail freight transportation systems was presented. The improvement of freight car routing and scheduling decisions is identified as an important research area.

In the next chapter, a freight car routing and scheduling model for managing heterogeneous traffic on rail freight networks is presented with a case study.

Table 5.1 : Models for tactical decisions

Author	Decisions	Modeling approach	Solution approach	Deterministic vs. Probabilistic	Data needs
Bodin, et al.	Blocking	Non-linear mixed-integer programming model	-	Deterministic	Medium
ABM (ALK) Assad	Blocking Train routing Traffic routing Train make-up	Heuristic Non-linear mixed-integer multicommodity flow problem	- -	Deterministic Deterministic	Medium Medium
Crainic, et al.	Train routing Traffic routing Classification Train make-up	Non-linear mixed-integer multicommodity flow problem	Heuristic decomposition technique	Deterministic	Medium
Keaton	Train routing Traffic routing Classification	Linear mixed-integer multicommodity flow problem	Heuristic decomposition technique	Deterministic	Medium
AAR model SPM	General General	Simulation model Simulation model	- -	Probabilistic Probabilistic	Large Large

Table 5.2 : Models for operational decisions

<u>Author</u>	<u>Decisions</u>	<u>Modeling approach</u>	<u>Solution approach</u>	<u>Deterministic vs.</u>	<u>Data needs</u>
Assad	Train scheduling	Analytical model	-	Probabilistic	-
TSS (ALK)	Train scheduling	Heuristic	-	Deterministic	Medium
Bongaardi, et al.	Train dispatching	Simulation model	-	Deterministic	Large
Kraft	Train dispatching	Integer programming model	Branch-and-bound	Deterministic	Medium
Yagar, et al.	Yard receiving	Analytical model	-	Deterministic	-
Folk and Sussman	Yard dispatching	Simulation model	-	Probabilistic	Medium
Florian, et al.	Engine scheduling	Linear multicommodity flow problem	Bender's decomposition	Deterministic	Medium
Smith and Sheffi	Engine scheduling	Non-linear multicommodity flow problem	Heuristic	Deterministic	Medium
Misra	Car distribution	Linear programming model	-	Deterministic	Medium
Ratcliffe, et al.	Car distribution	Simulation model	-	Probabilistic	Large
Philip	Car distribution	Analytical model	-	Deterministic	Medium
Jordan and Turnquist	Car distribution	Dynamic network optimization model	-	Probabilistic	Medium
Chih	Car distribution	Linear multicommodity flow problem	Subgradient algorithm	Deterministic	Medium
Adamidou, et al.	Car distribution	N-person noncooperative nonzero sum game	Iterative sequential solution approach	Deterministic	Medium

CHAPTER 6

Freight Car Routing and Scheduling for Managing Heterogeneous Traffic on Rail Freight Networks

6.1 Introduction

In the previous chapter, how the heterogeneity of traffic could be incorporated into hierarchical decisions in rail operations was discussed. The theoretically- and practically-developed models for rail operations were reviewed and current freight car scheduling practice was identified as one important area to be improved for the purpose of service differentiation.

The freight car scheduling system determines how a car is supposed to move from its origin to its destination. The car schedule or trip plan lists the pickup time, the sequence of trains on which the car will move on the appropriate day, the intermediate terminals at which it will be classified, and the estimated arrival time at its destination. Car scheduling is taking on a more important role in the operating plan as railroads pursue operations that are better scheduled and planned, and as more shippers demand car schedule information for their procurement, production and distribution plans.

In this chapter, an optimization-based decision support model for routing and scheduling heterogeneous traffic on rail freight networks, considering service requirements of different shippers, variable demand patterns and train capacity limits, is developed. In section 6.2, current freight car scheduling practice and areas for future improvement are discussed. In section 6.3, a dynamic freight car routing and scheduling model is developed. A time-space network representation of freight car moves on a rail freight network, a mathematical formulation, and a solution approach are presented. In section 6.4, the model developed in this study is tested on a hypothetical rail network based on the sub-network of a major US railroad. In section 6.5, conclusions from the case study and areas for further research are discussed.

6.2 Current Freight Car Scheduling Practice and Areas for Future Improvement

In current car scheduling systems, car schedules are usually derived from the operating plan and certain train connection standards at terminals. The local train schedule determines when a car can be picked up and delivered; the blocking plan determines car-to-block assignment; the train makeup plan determines block-to-train assignment; and the through train schedule and train connection standards at terminals determine the corresponding arrival/departure times of a car at intermediate terminals (Figure 6.1). For a typical merchandise car movement, the car schedule can be as shown as Figure 6.2.

Several railroads have developed and implemented their own freight car scheduling systems. Missouri Pacific Railroad (MP) was the first railroad to develop a freight car scheduling system (Missouri Pacific Railroad [1977]). Several railroads have such systems or are developing them, e.g., Union Pacific (UP), CSXT, ATSF, ConRail, Burlington Northern (BN).

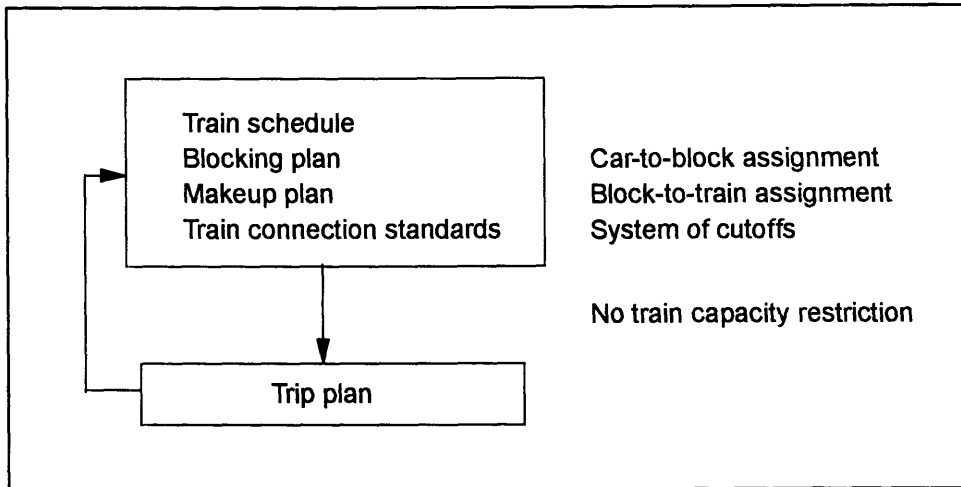


Figure 6.1 : Current car scheduling system

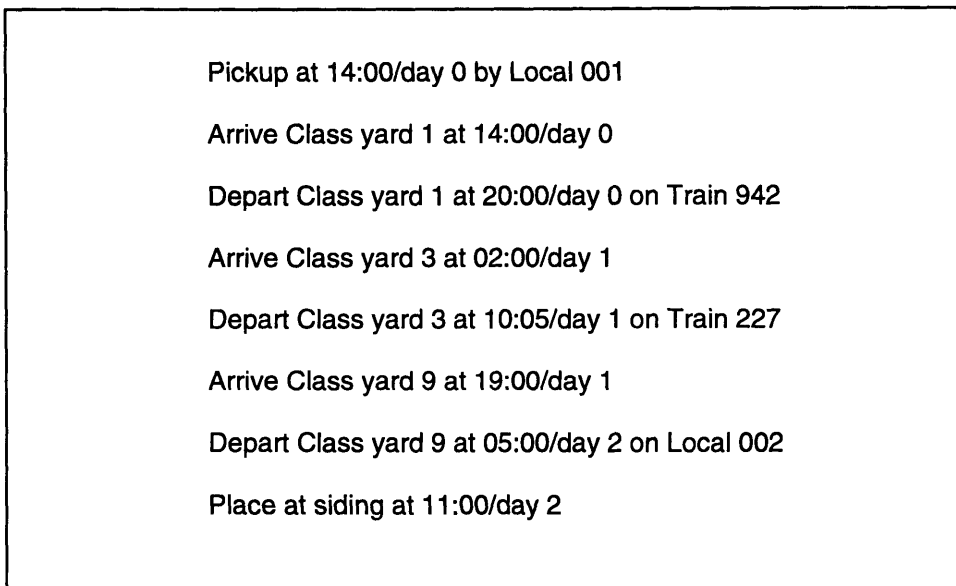


Figure 6.2 : Example of car schedule

There are some system-wide obstacles in current rail operations to building a freight car scheduling system. First, there are still many unscheduled trains (e.g., unit trains, extra trains) in rail operations. These unscheduled trains add a certain complexity to building car schedules for all

traffic. Recently, there has been a trend toward more scheduled operations as part of the railroad's efforts to improve service reliability. Many railroads, however, build car schedules and monitor the schedule performance for only their most important traffic.¹ Second, there is a large portion of traffic that consists of interline moves (using more than one railroad). To build a complete car schedule for a shipment, we need car scheduling capabilities for all participating railroads. Recently, there has been an industry-wide effort to improve interline operations (AAR [1991, 1994]).

Even if a railroad has car scheduling capability, there are still other important issues to be considered in building a car scheduling system that produces car schedules that are more reliable and sensitive to the service requirements of particular shippers.

The first issue is to consider is heterogeneity of traffic. Different shippers have different service expectations, and a railroad needs to develop different service standards for them. Some shippers expect fast and highly reliable service and are willing to pay for it. For example, manufacturers with the JIT (Just-In-Time) production system require a tight delivery time window and a very high standard of on-time performance. Other shippers require less costly service and may accept slower and less reliable moves. Therefore, such heterogeneity of traffic, with respect to trip time and reliability, needs to be considered in building car schedules.

The second important issue is how to consider certain variability in traffic (e.g., weekly and daily traffic variability) in building car schedules. The existing car scheduling systems simply assign an inbound car to the first feasible outbound train, which is defined by blocking and make-up plans, if the available yard time is greater than the cutoff. In certain peak periods, however, the number of cars that can be assigned to an outbound train might exceed the designed (or actual)

¹ C. D. Martland, Rail Session at TRF, Internal Memo, October, 1993

capacity of the train. Some cars can not be connected and will be delayed until the next train departure from the terminal. When train frequency is low, this delay can be very long.

In current car scheduling practice, car schedules are based on blocking and make-up plans that are usually built upon assuming the average traffic pattern, and largely neglects potential train capacity problems due to traffic variability. A scheduling practice without consideration for such traffic variability is likely to produce unreliable car schedules (or, car schedules with unnecessary slack time to avoid possible en route delays), especially when the system is congested.

In this research, we developed a dynamic car scheduling system that considers the heterogeneity of shippers and their traffic, the traffic variability, and the train capacity restrictions so that it produces schedules that are more achievable and sensitive to service requirements of shippers (Figure 6.3).

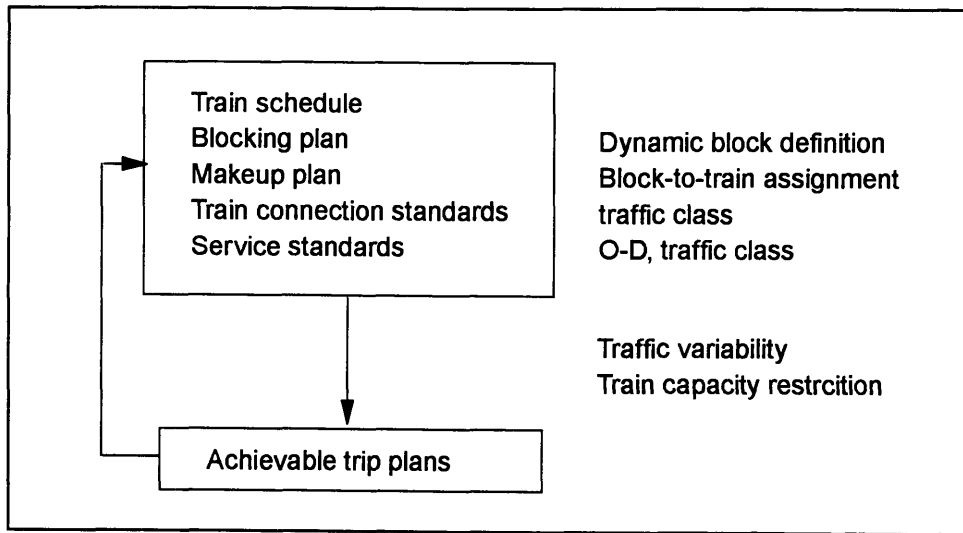


Figure 6.3 : Dynamic car scheduling system

6.3 Dynamic Freight Car Routing and Scheduling Model

We developed a mathematical model for supporting dynamic decisions for routing and scheduling heterogeneous car movements, which have different service requirements, during a certain period of time (e.g., one or two weeks) on a rail freight network that has limited physical and operational resources to provide services.

6.3.1 Incorporating service standards and traffic variability

The heterogeneity of traffic is considered by incorporating different service standards for different shippers. In the model, the service standards for different shippers are represented by delivery time windows. The traffic variability is considered by incorporating a weekly traffic pattern. A dynamic management scheme for car-to-block assignment is suggested to handle the traffic variability problem.

Incorporating service standards

A railroad can establish service standards for different shippers by defining a delivery time window based on the explicit requirements of shippers or negotiations with shippers. The delivery time window is specified by earliest and latest acceptable delivery times. Specifically:

$$t^k \in [t_{MIN}^k, t_{MAX}^k] \quad \forall k \in K \tag{6.1}$$

where, K is the set of shippers, t^k is the delivery time for shipper k , t_{MIN}^k is the earliest acceptable delivery time, and t_{MAX}^k the latest acceptable delivery time.

Different shippers may impose different explicit or implicit costs to a railroad if unreliable service is provided. Explicit costs may include a penalty cost for unreliable service. Implicit costs may include loss of goodwill and long-term market loss to competitive carriers or modes.

Depending upon how they measure and evaluate service, some shippers may impose penalties for both early and late arrival. Other shippers may impose penalties only for late arrival.

A railroad may estimate penalty costs for different shippers based on the explicit penalty cost specified by the shipper or independent study of demand elasticities. Specifically, the penalty imposed by shipper k (φ^k) can be represented as:

$$\varphi^k = \varphi(t^k) = \begin{cases} \phi_1(t^k - t_{MAX}^k), & \text{if } t^k > t_{MAX}^k \\ 0, & \text{if } t^k \in [t_{MIN}^k, t_{MAX}^k] \\ \phi_2(t_{MIN}^k - t^k), & \text{if } t^k < t_{MIN}^k \end{cases} \quad (6.2)$$

where, $\varphi(t^k)$ is the penalty cost when the delivery time is t^k , $\phi_1(\cdot)$ is the penalty cost function when the actual delivery time is later than the latest acceptable delivery time, and $\phi_2(\cdot)$ is the penalty cost function when the actual delivery time is earlier than the earliest acceptable delivery time.

Unreliable service also influences equipment utilization and imposes certain costs on the railroad. Therefore, it is important to establish a car scheduling system that produces car schedules, considering both the service requirements of different shippers and the potential costs of unreliable compliance with schedules.

Incorporating traffic variability

The other important issue to consider is traffic variability and possible train capacity problems. There are two possible types of traffic variability: one variability comes from relatively regular traffic variability, e.g., seasonalities and weekly traffic pattern; other variability has more stochastic nature, e.g., daily traffic variability. In this model, only a weekly demand pattern is incorporated as one aspect of traffic variability.

We may need to adjust schedules for certain peak periods to prevent possible excessive delays at congested terminals or due to limited train capacities in certain train segments. We need to “reschedule” some traffic (e.g., less important traffic) by holding it at intermediate terminals or we may even need to “reroute” them to an alternative route (i.e., block sequence and corresponding train sequence) in a peak period.

Current practice of predetermined and fixed car-to-block assignments² based on the blocking plan cannot deal with this problem. A dynamic management scheme for car-to-block assignment is suggested to incorporate the traffic variability in developing the car schedule for different days of week and different departure times at the origin terminal.

6.3.2 Time-space network representation

To represent temporal and spatial train and car movements on the rail network and to find the optimal pattern of car routes and schedules during a planning horizon, the time-space network representation technique is used. The time-space network representation technique has been used

² It is specified by “tag table” in the block definition that lists final destinations of cars that can be assigned to a specific block.

in many areas such as railroad engine scheduling problem (Florian, et al. [1976]), empty container allocation problem (White [1972], Crainic, et al. [1990]), truck fleet management (Powell, et al. [1988]), and airline crew scheduling (Barnhart, et al. [1993]).

Two time-space networks for train movements and for freight car movements are built to represent the rail operation over a certain period of time. A procedure for building a time-space network for rail operation is explained using a example of train schedule, block definition and make-up plan (Table 6.1 and Table 6.2). Some blocks (blocks number 6 and 8) are designed to bypass classification at intermediate terminals.

The blocking plan determines car-to-block assignments at all terminals. In current use of the blocking plan, a block is defined by several parameters such as origin, next common destination, priority and final destinations. By specifying final destinations for each block, the current blocking plan allow only fixed car-to-block assignment. To allow dynamic car-to-block assignment, the final destinations in the block definition are not specified, i.e., the “tag table” that lists final destination of cars that can be assigned to a specific block is eliminated.

Table 6.1 : Train schedule : example

<u>Train number</u>	<u>Train capacity (cars)</u>	<u>Terminal</u>	<u>Arrival</u>	<u>Departure</u>	<u>Day</u>
908	200	1	-	02:00	1
		3	04:40	05:00	1
		9	08:40	-	1
927	160	1	-	09:30	1
		3	11:55	12:25	1
		9	18:05	-	1
942	80	1	-	20:00	1
		3	24:00	-	1

Table 6.2 : Block definitions and train make-up plan : example

<u>Block number</u>	<u>Origin</u>	<u>Destination</u>	<u>Trains carrying block</u>
2	1	3	942
3	1	3	927
4	1	3	908
6	1	9	908
8	1	9	927
230	3	9	927, 908

Time-space network for train operations

A time-space network to represent scheduled train movement during the planning horizon and can be built from the train schedule.

Each node on the network represents a terminal at a specific point in time. Each link represents a scheduled train segment. The flow on each link denotes cars assigned to a specific train segment. The flow upper bound is the train capacity (in terms of the number of cars) constrained by the number of locomotives.¹ Each train capacity functions as a “bundle constraint” that limits the total number of cars in blocks that are connected to the train. The link cost is the train cost for a corresponding train segment. The train cost is a function of the distance, number of road crews and locomotives, train length, and tonnage. It is assumed that the train cost is fixed during a planning period.

¹ It also constrained by the track geometry and length of sidings.

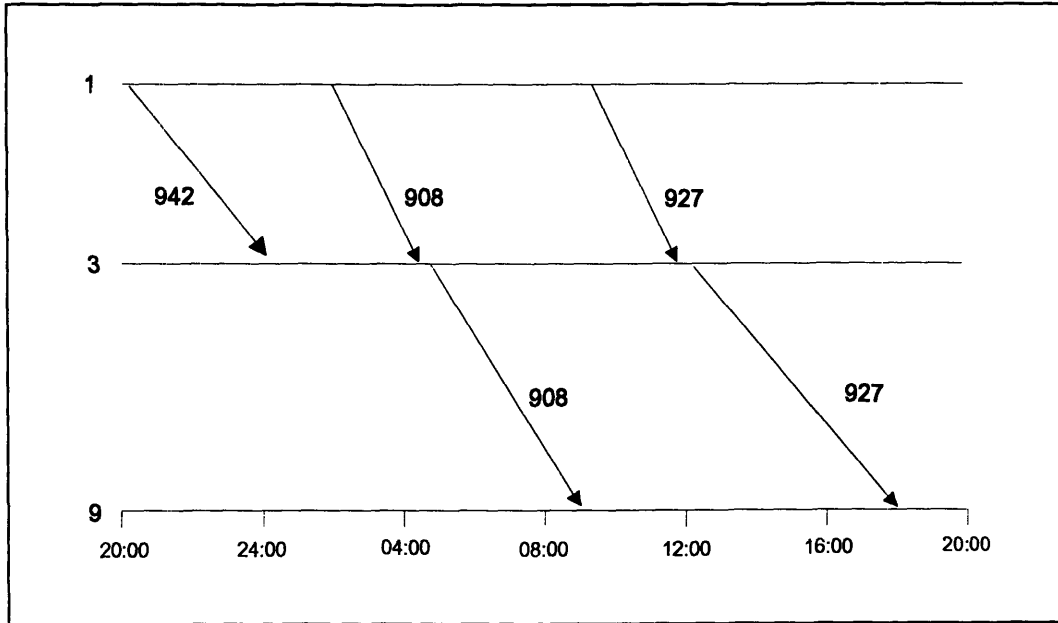


Figure 6.4 : Representation of train schedule

Time-space network to represent freight car movements

A time-space network is built to represent car movements through possible sequences of car-to-block and block-to-train assignments, car processing activity at a terminal, and possible holding at a terminal. Three different types of links are used: movement link, processing link and holding link.

A movement link represents a defined block with a corresponding train schedule. A processing link represents car processing activity at the terminal. A holding link represents car holding until the next available train departure at the terminal.

Links to represent defined blocks

Figure 6.5 shows the representation of all defined blocks in the example on a time-space network. When a specific block can be assigned to several trains at a terminal, we generate separate movement links for each train. For example, block number 230 at terminal 3 can be assigned to two trains (train 927 and 908). In such a case, we generate two movement links (one for block 230 on train 927 and other for block 230 on train 908). Some blocks are defined to use multiple train segments bypassing classification works at intermediate terminals. For example, block number 6 uses two train segments of train 908 (one segment from terminal 1 to 3 and the other from terminal 3 to 9) bypassing the classification at terminal 3. In addition, some blocks are defined to use more than one train. It simply requires a “block swap” operation that transfers the block to an other train without any classification work at a terminal. Each movement link, therefore, corresponds to one or more train segments. The flow on each movement link is constrained by capacity limits of one or more train segments. Each block defined in the blocking plan is represented as a single movement link.

This representation technique allows us to trace both car-to-block and block-to-train assignments and to build car schedules (i.e., sequence of blocks, trains and terminals) from the optimal solution of the model. The flow on the movement link denotes cars that are assigned to a specific block. The flow upper bound can be the maximum block size constrained by corresponding classification track length at terminal. Since the train capacity is the stronger constraint, the block size limit is not considered in the model.

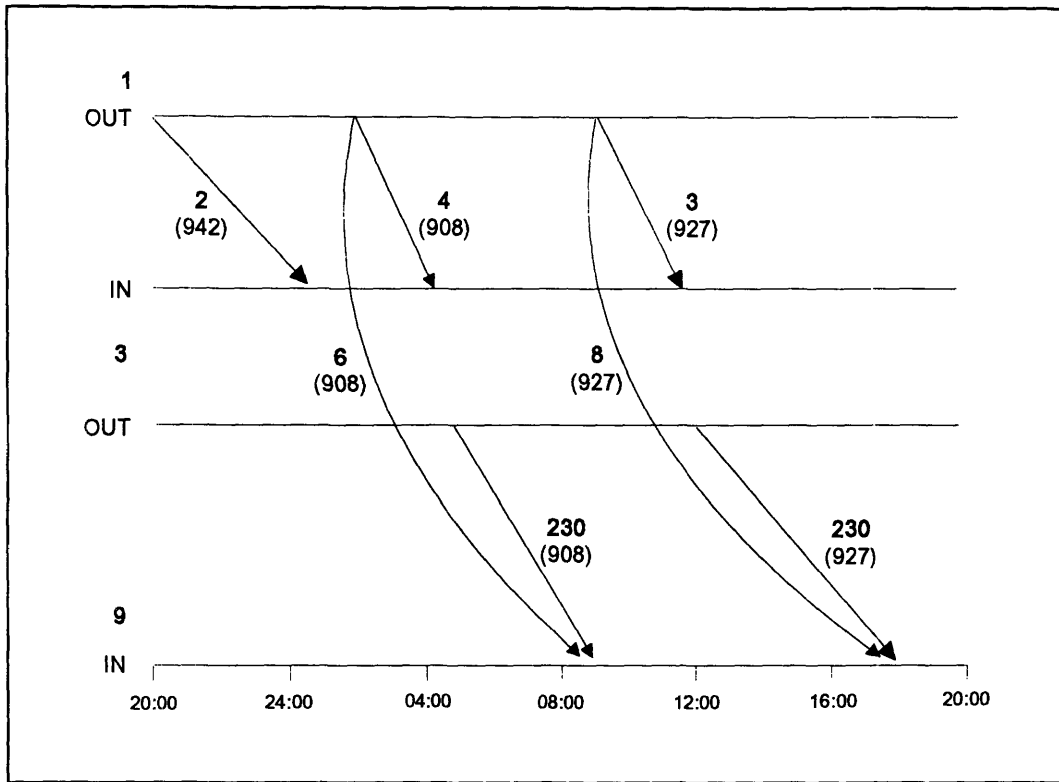


Figure 6.5 : Representation of blocks

Links to represent terminal processing

We represent each terminal as two separate nodes (IN and OUT nodes). An IN node is used to represent train arrival at terminal. An OUT node is used to represent both the train departure from terminal and the time to finish necessary terminal processing time for different traffic classes.

Empirical analyses of terminal performance have shown that different classes of traffic have different train connection performance (i.e., high priority traffic has more reliable train connection performance than low priority traffic) and different terminal processing times (see Kerr, Martland and Sussman [1976]). For example, suppose there are two traffic classes (e.g., high and

low priority traffic) and their terminal processing times are 3 and 6 hours respectively.² Then the estimated time to finish necessary terminal processing for each traffic class is computed and corresponding nodes are generated. Processing links are then generated by connecting nodes for train arrival and nodes for time to finish necessary terminal processing for all traffic classes. In the model, it is assumed that terminal processing times by traffic class are fixed considering the nature of short-run routing and scheduling decisions. The consideration of train capacity constraints in the model helps the daily total traffic volume handled at each terminal and the corresponding terminal processing times to be more stable.

Flow on the processing link denotes traffic volume by traffic class handled through yard operation. Actual flow on this link is always equal to the traffic volume of a corresponding traffic class (e.g., high priority class) that arrived on a specific train. The per unit car handling cost may be different at different terminals. Total car handling cost for the planning period is assumed to be fixed.

Links to represent car holding at terminal

When the number of cars that are expected to be moved on a specific train is more than the train's capacity, we need to hold some cars at the terminal until the next available train. For example, the number of cars in terminal 3 that arrived and finished the necessary terminal processing before departure time of train 908 is more than the capacity of train 908. Therefore, some cars need to be held at the terminal to connect to train 927 or other trains scheduled to depart later.

² Since terminal holding time is explicitly considered in the model, the delay due to missed train connection should not be included in the terminal processing time. It includes inbound and outbound inspection time, classification time, and make-up time.

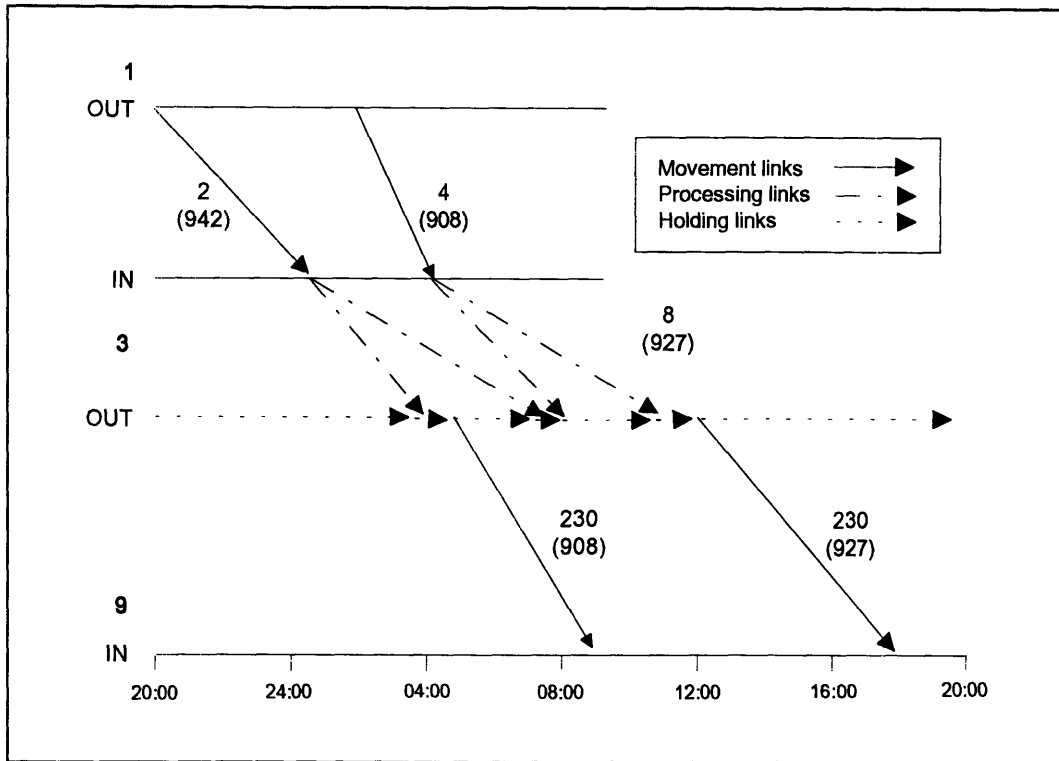


Figure 6.6 : Representation of terminal processing

The flow on the holding link denotes cars held in the terminal either waiting for the first available train connection or for the next available train connection if those cars missed their previous train connections due to a train capacity problem. The flow upper bound is set to infinity, assuming the terminal always has enough track capacity to hold cars. This assumption makes the train capacity problem as only reason that a car miss a connection in the terminal.

For a further model enhancement, we may set the flow upper bound of a holding link as the maximum number of cars that can be held at the terminal to consider the terminal capacity problem as other reason that a car miss a connection in the terminal.

Service standards and delivery time windows

To consider different service standards for different commodities, the delivery time window is defined by both the earliest acceptable arrival time and the latest acceptable arrival at destination for each commodity (see equation (6.1)).

A sink node and number of dummy links are used to represent possible car arrivals at the destination. Suppose a commodity (for example, a group of cars from origin 1 to destination 3 that is scheduled to depart on train 942) has a delivery time window defined by t_{MIN}^1 and t_{MAX}^1 . We generate one sink node for one commodity. Then we generate dummy links that connect all nodes for possible train arrivals at the destination and the sink node. The flow on each dummy link represents cars arriving at different times at the destination terminal for a corresponding commodity. The flow upper bound is set to infinity. The sum of flows on all dummy links to a specific sink node is equal to the volume of the corresponding commodity. The cost of a dummy link is the penalty cost on early or late arrival at the destination.

We make two assumptions about the penalty cost function in the model. First, we assume that only late arrivals are penalized by shippers. Second, we assume that the penalty cost is linearly proportional to the lateness.³ The delivery time window and penalty cost function are represented as follows:

$$t_p^k \in [t_{MAX}^k) \quad \forall p, \forall k \quad (6.3)$$

$$\varphi(t_p^k) = \bar{\Phi}_k \cdot (t_p^k - t_{MAX}^k) \quad \forall p, \forall k \quad (6.4)$$

³ By assuming the linear penalty cost function, we can use standard shortest-path algorithm to find a path that has minimum penalty cost from the origin to the destination of each commodity.

6.3.3 Problem formulation : multicommodity flow problem

We formulated a multicommodity flow problem on a time-space network to determine combined routes and schedules of car movements for a given planning period.

A commodity can be defined as a group of cars that has the same origin, destination, traffic class and departure time.⁴ Since individual commodities share common train capacity, we can formulate our problem as a multicommodity flow problem. We formulate the multicommodity flow problem using path flows.

The following notation is used for a path flow formulation.

c_p^k : per unit cost of commodity k on path p

c_a^k : per unit cost of link a

$\varphi(t_p^k)$: per unit penalty cost of commodity k by using a path p

f_p^k : flow of commodity k on path p

x_a : size of block a

y_t : length of train segment t

d^k : demand of commodity k

u_t : capacity of train segment t

δ_a^p : indicator variable for car-to-block assignment

$$\delta_a^p = \begin{cases} 1 & \text{if itinerary } p \text{ uses block } a \\ 0 & \text{otherwise} \end{cases}$$

⁴ It also can be defined as a shipment.

ξ_t^a : indicator variable for block-to-train assignment

$$\xi_t^a = \begin{cases} 1 & \text{if block } a \text{ is connected to train } t \\ 0 & \text{otherwise} \end{cases}$$

K : set of commodities

P^k : set of paths from the origin to the destination of commodity k

P : set of paths of all commodities (= $\{P^1, P^2, \dots, P^K\}$)

V : set of nodes

A : set of links (including defined blocks)

T : set of train segments

Objective function

The objective is to find the optimal pattern of commodity flows on the specified time-space network and to build corresponding car routes and schedules so as to minimize the total penalty costs. It is assumed that train cost, car handling cost and car time cost are fixed during the planning period. So, the objective function becomes:

$$\text{Min} \sum_{k \in K} \sum_{p \in P^k} c_p^k f_p^k \quad (6.5)$$

where, the path cost c_p^k is the penalty cost by using path p .

$$\begin{aligned} c_p^k &= \sum_{a \in p} c_a^k = \varphi(t_p^k) \\ &= \max\{0, \bar{\varphi} \cdot (t_p^k - t_{MAX}^k)\} \end{aligned} \quad (6.6)$$

Demand constraints

To ensure that all cars reach their destinations, the following demand constraint is required for each commodity.

$$\sum_{p \in P^k} f_p^k = d^k \quad \forall k \in K \quad (6.7)$$

Car-to-block and block-to-train assignments

A car moves from its origin to its destination via a sequence of blocks and trains. The car-to-block and block-to-train assignments are explicitly considered in the model for realistic representation of rail operations. In the formulation, an indicator variable δ_a^p is used to represent car-to-block assignment and another indicator variable ξ_t^a to represent block-to-train assignment. The size of block a (x_a) and the length of train segment t (y_t) can be represented using the path flow variables.

$$x_a = \sum_{k \in K} \sum_{p \in P^k} \delta_a^p f_p^k \quad (6.8)$$

$$y_t = \sum_{a \in A} \xi_t^a x_a = \sum_{a \in A} \xi_t^a \left(\sum_{k \in K} \sum_{p \in P^k} \delta_a^p f_p^k \right) \quad (6.9)$$

For example, consider a sub-network of Figure 6.8 that includes 2 train segments and 3 blocks. From this network configuration, the following variables are defined for the corresponding blocks and train segments.

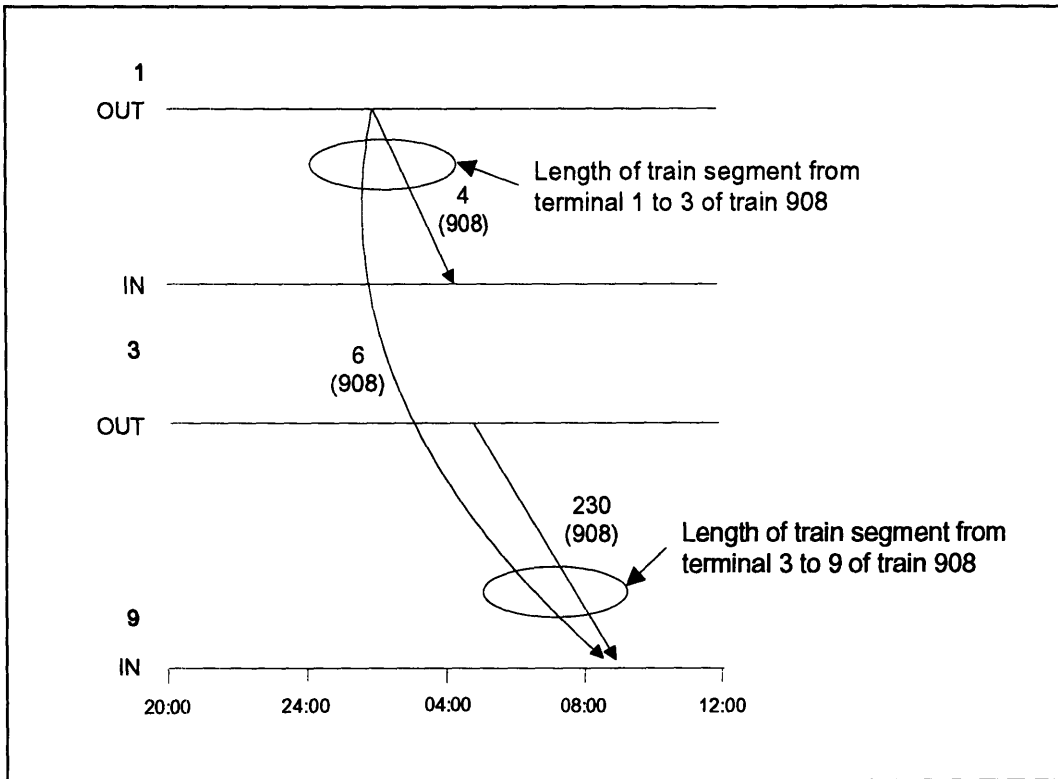


Figure 6.8 : Representation of car-to-block and block-to-train assignments

x_1 : size of block 4 from terminal 1 to 3

x_2 : size of block 230 from terminal 3 to 9

x_3 : size of block 6 from terminal 1 to 9

y_1 : length of train segment from terminal 1 to 3 of train 908

y_2 : length of train segment from terminal 3 to 9 of train 908

There are two possible paths from terminal 1 to 9 and one path from terminal 3 to 9. The following path variables are defined.

f_1^{1-9} : volume of path 1 from terminal 1 to 9

f_2^{1-9} : volume of path 2 from terminal 1 to 9

f_1^{3-9} : volume of path 1 from terminal 3 to 9

The indicator matrices can be defined, for car-to-block and block-to-train assignment based on the blocking and train make-up plans, as follows.

$$\begin{bmatrix} \delta_1^1 & \delta_1^2 & \delta_1^3 \\ \delta_2^1 & \delta_2^2 & \delta_2^3 \\ \delta_3^1 & \delta_3^2 & \delta_3^3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

$$\begin{bmatrix} \xi_1^1 & \xi_1^2 & \xi_1^3 \\ \xi_2^1 & \xi_2^2 & \xi_2^3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

From these indicator matrices, we can derive the block volume and the train length of each train segment using path variables. It means that we can formulate the problem using only path flow variables.

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \delta_1^1 & \delta_1^2 & \delta_1^3 \\ \delta_2^1 & \delta_2^2 & \delta_2^3 \\ \delta_3^1 & \delta_3^2 & \delta_3^3 \end{bmatrix} \begin{bmatrix} f_1^{1-9} \\ f_2^{1-9} \\ f_1^{3-9} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} f_1^{1-9} \\ f_2^{1-9} \\ f_1^{3-9} \end{bmatrix} = \begin{bmatrix} f_1^{1-9} \\ f_1^{1-9} + f_1^{3-9} \\ f_2^{1-9} \end{bmatrix}$$

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \xi_1^1 & \xi_1^2 & \xi_1^3 \\ \xi_2^1 & \xi_2^2 & \xi_2^3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 + x_3 \\ x_2 + x_3 \end{bmatrix} \text{ or,}$$

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \xi_1^1 & \xi_1^2 & \xi_1^3 \\ \xi_2^1 & \xi_2^2 & \xi_2^3 \end{bmatrix} \begin{bmatrix} \delta_1^1 & \delta_1^2 & \delta_1^3 \\ \delta_2^1 & \delta_2^2 & \delta_2^3 \\ \delta_3^1 & \delta_3^2 & \delta_3^3 \end{bmatrix} \begin{bmatrix} f_1^{1-9} \\ f_2^{1-9} \\ f_1^{3-9} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} f_1^{1-9} \\ f_2^{1-9} \\ f_1^{3-9} \end{bmatrix} = \begin{bmatrix} f_1^{1-9} + f_2^{1-9} \\ f_1^{1-9} + f_2^{1-9} + f_1^{3-9} \end{bmatrix}$$

Based on this discussion, the train capacity constraints can be represented using indicator variables for car-to-block and block-to-train assignment, and path flow variables.

$$\begin{aligned} y_t &= \sum_{a \in A} \xi_t^a x_a \\ &= \sum_{a \in A} \xi_t^a \left(\sum_{k \in K} \sum_p \delta_a^p f_p^k \right) \leq u_t, \quad \forall k \in K \end{aligned}$$

Path flow formulation

As a result, the path flow formulation of the multicommodity flow problem is represented as follows.

$$\text{Min} \quad \sum_{k \in K} \sum_{p \in P^k} c_p^k f_p^k \quad (6.10)$$

$$\text{s.t.} \quad \sum_{p \in P^k} f_p^k = d^k \quad \forall k \in K \quad (\sigma^k) \quad (6.11)$$

$$\sum_{a \in A} \xi_t^a \left(\sum_{k \in K} \sum_{p \in P^k} \delta_a^p f_p^k \right) \leq u_t, \quad \forall t \in T \quad (w_t) \quad (6.12)$$

$$f_p^k \geq 0 \quad \forall p \in P^k, \forall k \in K \quad (6.13)$$

In a matrix form, it can be written as

$$\begin{aligned}
 & \text{Min} \quad \underset{(|P| \times |P|)}{c} \underset{(|P| \times 1)}{f} \\
 \text{s. t.} \quad & \underset{(|K| \times |P|)}{D} \underset{(|P| \times 1)}{f} = \underset{(|K| \times 1)}{d} \\
 & \underset{(|T| \times |A|)}{\xi} \underset{(|A| \times |P|)}{\delta} \underset{(|P| \times 1)}{f} \leq \underset{(|T| \times 1)}{u} \\
 & \underset{(|P| \times 1)}{f} \geq \underset{(|P| \times 1)}{0}
 \end{aligned} \tag{6.14}$$

The path flow formulation of this multicommodity flow problem has a very simple constraint structure. The problem has a single constraint for each train segment t which states that the sum of car flows using train segment t is at most the capacity of train u_t . The sum of car flows using train segment t can be obtained from the car-to-block assignment (6.8) and the block-to-train assignment (6.9). Moreover, the problem has a single constraint for each commodity k which states that the total car flow on all the feasible paths connecting the source node s^k and the sink node t^k of commodity k must equal the demand d^k for this commodity.

For a network with $|V|$ nodes, $|K|$ commodities and $|T|$ train segments, the path flow formulation contains $|K| + |T|$ constraints (in addition to the nonnegativity constraints). In constraint, the link formulation contains $|V||K| + |T|$ constraints since it contains flow conservation equations for every node and commodity combination. This savings in the number of constraints does come at a cost since the path flow formulation has a variable for every feasible path connecting a source and sink node for each of the commodities. The number of variables will typically be enormous, growing exponentially with the size of the network. This characteristics of path flow formulation leads us to use the column generation technique as a solution approach.

Optimality conditions

Let's define a dual variable σ^k for each commodity k and another dual variable w_t for each train segment t . With respect to these dual variables, the reduced cost $c_p^{\sigma,w}$ for each path flow variable f_p^k can be defined as

$$\begin{aligned} c_p^{\sigma,w} &= \sum_{a \in p} (c_a^k + \sum_{t \in a} w_t) - \sigma^k \\ &= \sum_{a \in p} c_a^k + \sum_{a \in p} \sum_{t \in a} w_t - \sigma^k \end{aligned} \quad (6.15)$$

That is, the reduced cost of path p is the cost of that path with respect to the modified costs $c_a^k + \sum_{t \in a} w_t$ minus the commodity cost σ^k .

Complementary slackness conditions

The complementary slackness conditions for the path flow formulation (6.10 - 6.13) can be described as follows. The path flows f_p^k are optimal in the path flow formulation (6.10 - 6.13) of the multicommodity flow problem if and only if for some dual variables w_t and σ^k , the reduced costs and link flows satisfy the following complementary slackness conditions.

$$w_t [\sum_{a \in A} \xi_t^a (\sum_{k \in K} \sum_{p \in P^k} \delta_a^p f_p^k) - u_t] = 0 \quad \forall t \in T \quad (6.16)$$

$$c_p^{\sigma,w} \geq 0 \quad \forall p \in P^k, \forall k \in K \quad (6.17)$$

$$c_p^{\sigma,w} f_p^k = 0 \quad \forall p \in P^k, \forall k \in K \quad (6.18)$$

6.3.4 Column generation solution procedure

Dantzig and Wolfe [1960] developed a technique to solve specially structured large linear programs. Their technique solves the LP by alternately solving a coordinating restricted master problem and smaller linear sub-problems. Column generation methods, based on the decomposition principle of Dantzig and Wolfe, recognize that it is not necessary to have the entire constraint matrix available during the time of computation; columns need to be generated only as and when necessary.

Column generation technique has been applied in many areas such as railroad operations (Florian, et al. [1976], Crainic and Rousseau [1986]), airline crew scheduling (Crainic and Rousseau [1987], Barnhart, et al. [1993]), urban transit crew scheduling problem (Desrochers and Soumis [1989]), and vehicle routing problem with time windows (Desrosiers, et al. [1984], Desrochers, et al. [1992]).

The key idea in column generation is never to include explicitly all of the columns of the problem formulation, but rather to generate them only as needed. Each iteration of the column generation algorithm involves the solution of each of the following two parts:

1. The restricted master problem (RMP) that determines the optimal selection of paths. It also determines a set of dual solution w_i and σ^k .
2. The subproblem which generates one or more paths with negative reduced cost, i.e., paths that can potentially reduce the cost of the current solution generated by RMP. New paths can be generated by solving the shortest-path problem.

For the path flow formulation of the multicommodity flow problem with respect to the current basis at any step, the revised simplex method defines the simplex multipliers w_t and σ^k so that the reduced cost of every variable in the basis is zero. If a path p connecting the source s^k and sink t^k of commodity k is one of the basic variables, then $c_p^{\sigma, w} = 0$. Therefore, the revised simplex method determines the simplex multipliers w_t and σ^k so that they satisfy the following equations.

$$\sum_{u \in p} (c_u^k + \sum_{t \in u} w_t) = \sigma^k \text{ for every path } p \text{ in the basis} \quad (6.19)$$

The solution defined by the current basis always satisfies conditions (6.16) and (6.18). Therefore, it is optimal if it satisfies condition (6.17). This condition requires that the reduced cost of every path flow variable is nonnegative. We need to check to see if each commodity k satisfies the following inequality.

$$c_p^{\sigma, w} = \sum_{u \in p} (c_u^k + \sum_{t \in u} w_t) - \sigma^k \geq 0 \quad \forall p \in P^k \quad (6.20)$$

or, equivalently,

$$\min_{p \in P^k} \sum_{u \in p} (c_u^k + \sum_{t \in u} w_t) \geq \sigma^k \quad (6.21)$$

The left-hand side of inequality (6.21) is just the length of the shortest path connecting the source and sink nodes, s^k and t^k , of commodity k with respect to the modified costs $c_u^k + \sum_{t \in u} w_t$.

We solve a shortest path problem from source node s^k to sink node t^k for each commodity k with respect to the modified costs $c_a^k + \sum_{t \in a} w_t$. If for all commodities, the length of the shortest path for that commodity is greater than or equal to σ^k , current path flows f_p^k satisfy the condition (6.17) and are the optimal solution.

Otherwise, if for some commodity k , q denotes the shortest path with respect to the current modified costs $c_a^k + \sum_{t \in a} w_t$ and the reduced cost of path q is negative

$$c_q^{\sigma, w} = \sum_{a \in q} (c_a^k + \sum_{t \in a} w_t) - \sigma^k < 0 \quad (6.22)$$

Then we would perform a basis change introducing the path q into the current basis. It will determine a new set of simplex multipliers w_t and σ^k . We would then, as before, solve a shortest path problem for each commodity k and do the optimality check. We will continue by (1) finding new values for w_t and σ^k , and (2) solving shortest path problems alternately until we find the optimal solution.

The column generation algorithm is implemented as follows:

Step 1. Initialization : Choose an initial set of paths as an initial basic feasible solution.

Step 2. Restricted master problem : Solve RMP to optimality using the revised simplex method.

It will determine a set of dual solution w_t and σ^k .

Step 3. Column generation : Solve a shortest path problem from source node s^k to sink node t^k for each commodity k with respect to the modified costs $c_a^k + \sum_{t \in a} w_t$.

- If for all commodities, the length of the shortest path for that commodity is greater than or equal to σ^k , the current solution is the optimal solution.
- Otherwise, add paths that have negative reduced cost to RMP and go to Step 2.

6.3.5 Inputs and outputs of the model

Table 6.3 summarizes the inputs to the model. The operating plans such as train schedules, blocking plan and train make-up plan are inputs to the model. If a railroad modified one or more of these operating plans, the model produces a different set of car schedules. When a set of car schedules obtained from the model is not satisfactory, we can obtain another set of car schedule that is more satisfactory using an iterative procedure by changing input operating plans.

In addition, terminal processing times, service standards, and predicted demand during the planning horizon are other inputs to the model. The model assumed the predicted but known demand pattern during the planning horizon. The operating plans, terminal processing time and service standards are used to construct the time-space network. The penalty cost information is used to define the cost of dummy links that represent the service standards.

Table 6.4 summarizes the output from the model. The main output from the model is the trip plan (car schedule) information that will be used to make customer commitments. The model generates more achievable trip plans than the current car scheduling system, by considering the expected train capacity problem during the planning horizon. The model generates shipment-based trip plans for shipments from different shippers at different times of the day on different days of the week.

The model computes the train capacity utilization for each train during the planning horizon. It allows more effective planning of train capacity by finding critical capacity bottlenecks

of the system. The model also computes the terminal handling volume for each terminal at different time of the day on different day of the week. With the trip plan information, these information will help terminal managers to build the more effective terminal work plans.

Table 6.3 : Summary of the inputs to the model

Operating plans
• Train schedule (specified by each train segment)
– departure and arrival terminals
– departure and arrival times
– designed train capacity
• Blocking plan
– block definition (without specification of final destinations)
• Train make-up plan
– block-to-train assignment
Terminal processing time
– estimated terminal processing time by traffic class
Service standards for shippers
– delivery time window specified by shipper
– penalty cost on early or late arrival at destination specified by shipper
Demand during planning horizon
– known and predicted demands by shipper

Table 6.4 : List of outputs from the model

Trip plan information
– by shipper
– by day of week
– by departure time at the origin terminal
Train capacity utilization
– by train segment
– by day of the week
Terminal handling volume
– by terminal
– by day of the week
– by time of the day

6.4 Case Study

This section presents the description and results of the case study. The dynamic freight car routing and scheduling model developed in the research has been tested on a hypothetical network based on the sub-network of a major US railroad.¹ The capability of the model to produce clearly differentiated schedules of different traffic classes is discussed. By doing a sensitivity analysis on different traffic mixes (i.e., the shares of different traffic classes), what portion of traffic can be provided a high quality service is examined.

The model is coded using IBM's OSL (Optimization Subroutine Library) and the C programming language. All computational tests were run on an IBM RS 6000/370 workstation.

6.4.1 Case description

The case network has 12 terminals (Figure 6.9). 16 trains are providing car-load (or general merchandise) train service on the network (Figure 6.10).² The train schedule for each train is specified in terms of a set of departure and arrival times along with its train route. For each train, the number of locomotives, which determines its train capacity in terms of the maximum number of cars it can haul, is specified. The blocking plan specifies 59 blocks. Appendix B describes the train schedule, block plan, and train make-up plan.

The network serves 56 markets segmented by origin, destination and traffic class. Traffic is classified into high and low priority traffic. Total traffic volume for a one-week period is 9,712

¹ It is hypothetical because we used hypothetical data of weekly demand, service standards and terminal processing times. The rail network, train schedule, blocking plan and train make-up data are based on the actual data.

² For the case network, there are no trains that bypass one or more terminals.

cars. Daily average traffic volume is 1,387 cars for the entire network. In the network, a corridor between terminal 3 and terminal 9 is somewhat congested and may cause delays due to train capacity problem. Appendix B describes the detailed O-D demand information for the planning period.

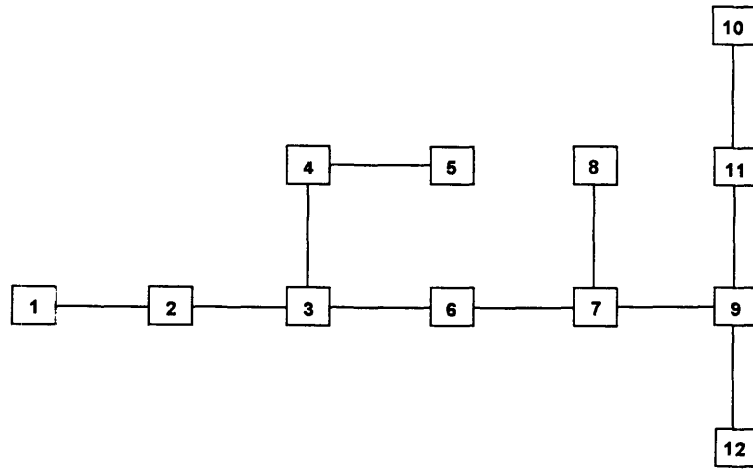


Figure 6.9 : Case network (III)

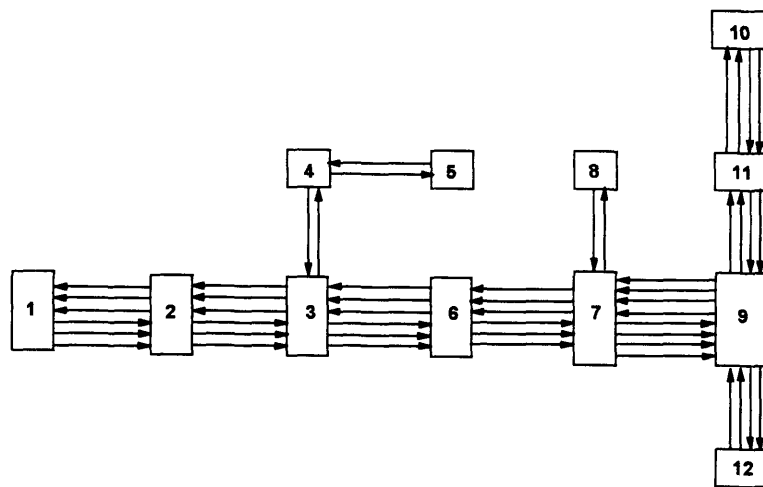


Figure 6.10 : Train service network (III)

For the case study, the delivery time window and penalty cost function are defined as follows. We use the shortest path time as the service standard t_{MAX}^k (i.e., the latest acceptable delivery time) for each commodity k .

We assumed that the per car terminal processing times for high and low priority cars are 8 hours and 16 hours, respectively. Since terminal processing times are different, feasible train connections are different although they arrive at a terminal at the same time. As a result, the shortest path times of different traffic classes for the same origin-destination pair are different. It means that the service standards of different traffic classes for the same origin-destination pair are different. For the same origin-destination pair, the high priority class has a faster service standard than the low priority class.

A different penalty cost $\bar{\Phi}_k$ is assumed for different traffic class k . The penalty cost per hour late³ for the high priority class is \$9. The penalty cost per hour late for the low priority class is \$3. A railroad can use its own set of service standards and a different penalty cost function.

Several test cases that have different traffic mixes (i.e., shares of high and low priority traffic) are defined. With P1 as the base case, they are designed to examine the sensitivity of the model output to the traffic mixes and the operating capability of service differentiation. Table 6.5 describes 8 test cases examined.

6.4.2 Analysis results

Table 6.6 summarizes the sizes of case problems in terms of number of commodities, number of nodes, and number of links; and the computational performance in terms of number of iterations to

³ Hour late from the latest acceptable delivery time.

reach the optimality, total number of columns generated, and computation time in seconds. For all problems, the number of commodities is 1,226 and the number of nodes is 2,946.

Table 6.5 : Test cases with different traffic mixes

<u>Problem</u>	<u>High priority (cars)</u>	<u>Low priority (cars)</u>	<u>High priority share (%)</u>
P1	3,068	6,644	31.6
P2	3,224	6,488	33.2
P3	3,784	5,928	39.0
P4	4,484	5,228	46.2
P5	5,424	4,288	55.8
P6	5,884	3,828	60.6
P7	6,724	2,988	69.2
P8	7,362	2,350	75.8

Table 6.6 : Problem size and computational performance

<u>Problem</u>	<u>No. of links</u>	<u>No. of iterations</u>	<u>No. of columns</u>	<u>Computation time (sec)</u>
P1	27,238	10	4,060	1,563
P2	27,214	9	4,249	1,442
P3	26,461	8	4,012	1,227
P4	26,236	9	4,140	1,392
P5	24,722	9	4,137	1,298
P6	24,510	10	4,069	1,391
P7	24,110	9	4,103	1,249
P8	23,462	9	4,067	1,204

As was discussed in section 6.2, one of the problems of current car scheduling practice is that it does not consider train capacity constraints in generating car schedules. In current practice,

a car scheduling system produces and maintains a fixed set of car schedules, and then assigns traffic on that fixed set of car schedules. This practice at best provides a warning that some trains are over-subscribed.

To examine the range of capacity utilization during a planning period and possible capacity-related problem under different car scheduling practices, we compute the maximum and minimum levels of capacity utilization during a one-week period for each train segment based on the result of assigning traffic on a set of car schedules.

$$\rho_t^{\max} = \max_{n=1 \dots N} \left\{ \frac{y_t^n}{u_t} \right\}$$

$$\rho_t^{\min} = \min_{n=1 \dots N} \left\{ \frac{y_t^n}{u_t} \right\}$$

where, ρ_t^{\max} is the maximum level of capacity utilization for train segment t , ρ_t^{\min} is the minimum level of capacity utilization for train segment t , u_t is the capacity of train segment t , y_t^n is the train length of train segment t at day n , and N is the planning period.

Consider a case problem P1. We compare the projected train capacity utilization by train segment during the planning period when a railroad uses a fixed set of car schedules without considering train capacity constraints and when it uses the optimal set of car schedules generated from the model.

Suppose a railroad uses a fixed set of car schedules that are based on a shortest-time path (i.e., the shortest possible schedule) for each origin-destination pair.⁴ To project train capacity utilization for each train segment during the planning period, we assigned the traffic volume of

⁴ A sequence of car-to-block and block-to-train from the origin to the destination that results in the shortest O-D trip time.

each O-D at different departure day and time to its corresponding shortest-time path without considering train capacity constraints.

Figure 6.11 shows the range of train capacity utilization for each train segment during the planning period when traffic is assigned on a fixed set of car schedules. Several train segments are highly over-subscribed and others are under-subscribed. This is mainly because this car scheduling practice does not consider train capacity constraints.

If there are over-subscribed trains, cars at corresponding terminals are more likely to miss the train connections at terminals, due to train capacity problems, and to have long delays. A railroad needs more controls on car-to-train connections at the terminal-level, which is done usually without considering the impact on system-level service performance, such as origin-destination trip time and on-time performance. For some train segments that are highly over-subscribed, it may not be feasible to operate trains of such long length.

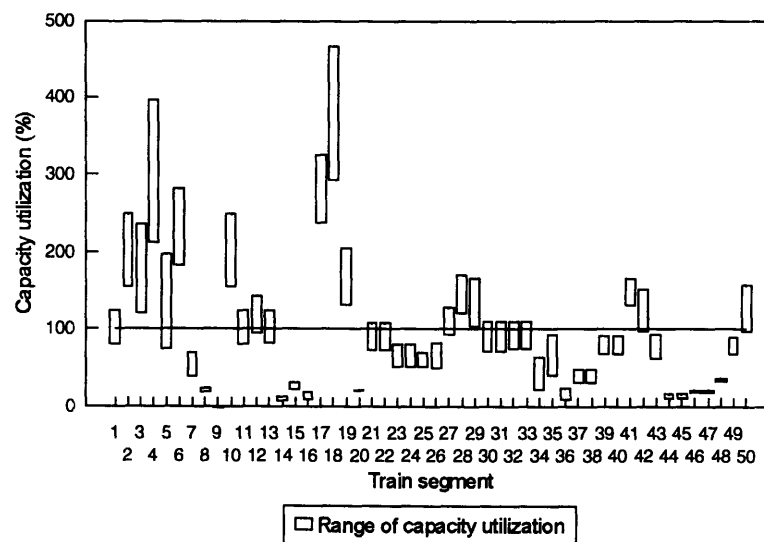


Figure 6.11 : Train capacity utilization with shortest-time path based car schedule

The dynamic car routing and scheduling model developed in this research resolves this problem by explicitly considering train capacity constraints. Figure 6.12 shows the range of train capacity utilization for each train segment when traffic is assigned on the schedules generated from the model. There are no over-subscribed trains during the planning period.

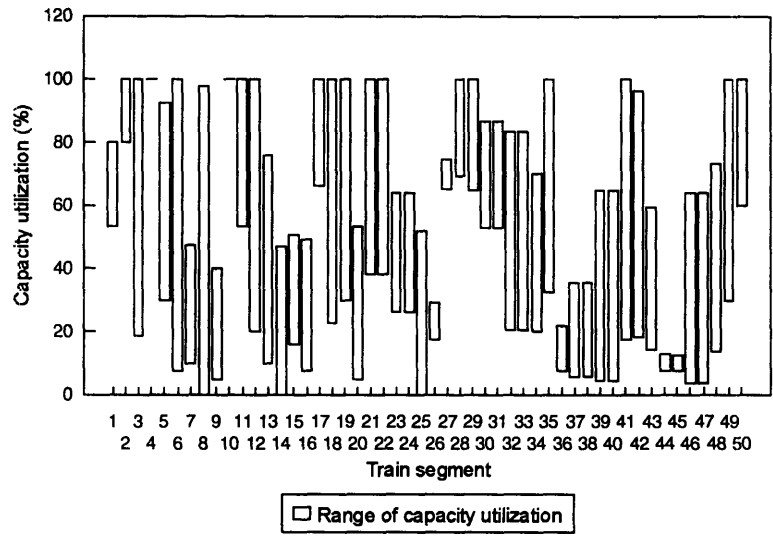


Figure 6.12 : Train capacity utilization with optimal car schedules from the model

The analysis of capacity utilization allows more effective plan on train capacity. From the results of capacity utilization by train segment over the planning horizon, critical train capacity bottlenecks that were fully utilized over a planning period or only during a certain peak days could be found (Figure 6.13). By increasing the capacity of selected bottlenecks, a railroad may use the additional capacity increase to optimal service improvements. Appendix C summarizes the train capacity utilization for all train segments during the planning period.

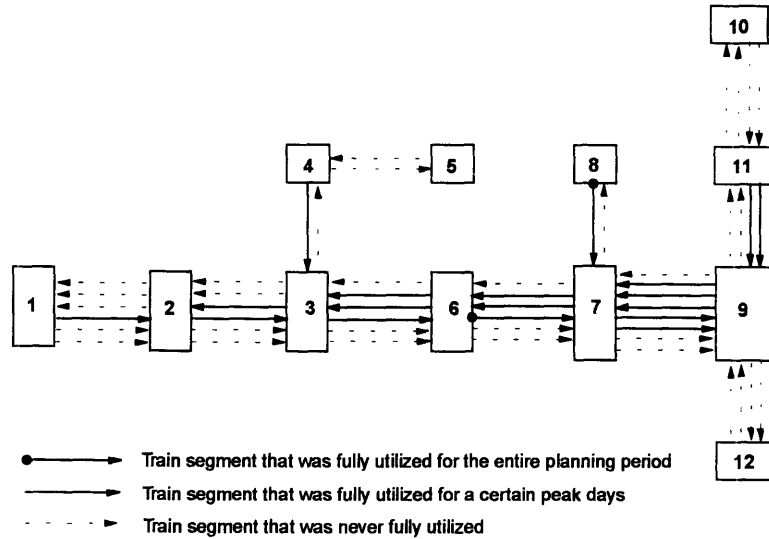


Figure 6.13 : Train capacity bottlenecks

The major output from the model is car schedule information. The model produces distinct car schedules for different departure times at origin, for different days of week for each market, for different traffic classes, and for different O-D pairs.

Again, consider a case problem P1. Table 6.7 and 6.8 show the scheduled O-D trip times of selected O-D pairs during the planning period. High priority traffic (origin 5 to destination 3) has scheduled O-D trip times that are highly consistent for different days of the week (Table 6.7). On the other hand, low priority traffic (origin 5 to destination 7) has scheduled O-D trip times that are varied over different days of the week (Table 6.8). Appendix D shows the detailed car schedule information generated from the model for these specific O-D pairs.

To compare the difference in planned service levels for different classes of traffic, the mean and standard deviation of lateness by traffic class are computed. The lateness is measured by $\max\{0, t_p^k - t_{MAX}^k\}$, i.e., the difference between the scheduled O-D trip time and the latest acceptable delivery time.

Table 6.7 : Scheduled O-D trip times for different days of the week :
origin 5, destination 3, high priority

<u>Day</u>	<u>Traffic volume (cars)</u>	<u>Service standard (hours)</u>	<u>Schedule trip time (hours)</u>
1	16	6.8	6.8
2	16	6.8	6.8
3	24	6.8	6.8
4	28	6.8	6.8
5	24	6.8	6.8
6	16	6.8	6.8
7	16	6.8	6.8

Table 6.8 : Scheduled O-D trip times for different days of the week :
origin 5, destination 7, low priority

<u>Day</u>	<u>Traffic volume (cars)</u>	<u>Service standard (hours)</u>	<u>Scheduled trip time (hours)</u>
1	16	34.2	34.2
2	3	34.2	34.2
	13	34.2	67.2
3	10	34.2	91.2
	14	34.2	178.2
4	28	34.2	154.2
5	24	34.2	82.2
6	16	34.2	58.2
7	16	34.2	58.2

For the case problem P1, mean lateness is 0.1 hours and standard deviation of lateness is 0.9 hours for the high priority traffic. Mean lateness is 10.9 hours and standard deviation of

lateness is 31.4 hours for the low priority traffic. It suggests that a railroad can plan fast and highly reliable service for high priority traffic and slower and less reliable service for low priority traffic, using the model.

We did a sensitivity analysis on traffic mixes by examining all 8 different test cases that have different shares of high and low priority traffic. Table 6.9 summarizes the optimal solutions (minimum penalty cost, mean and standard deviation of lateness for different traffic classes) for case problems. High priority traffic is provided a highly reliable service. As a railroad tries to provide a highly reliable service to more traffic, however, service level of high priority traffic also begin to deteriorates. Service level of low priority traffic, on the other hand, significantly deteriorates as the share of high priority traffic increases. In addition, the difference between service levels of high and low priority traffic become larger as the share of high priority traffic increases (see Figure 6.14 and 6.15). This result is consistent with the previous conclusion we obtained from the sensitivity analysis on traffic mixes for a train service among multiple terminals in the section 4.3.4 (see Table 4.22 and Figure 4.16).

Table 6.9 : Optimal solutions for case problems

<u>Problem</u>	<u>Min. penalty cost (dollar)</u>	<u>High priority</u>		<u>Low priority</u>	
		<u>Mean lateness (hours)</u>	<u>Std dev lateness (hours)</u>	<u>Mean lateness (hours)</u>	<u>Std dev lateness (hours)</u>
P1	220,495.2	0.1	0.9	10.9	31.4
P2	222,674.7	0.1	0.9	11.3	30.8
P3	222,674.7	0.1	0.9	12.3	32.7
P4	224,588.7	0.1	0.8	14.1	34.6
P5	238,542.0	0.5	2.2	16.7	37.6
P6	242,761.5	0.6	2.7	18.4	37.7
P7	255,185.5	0.8	3.4	23.0	43.1
P8	262,529.3	0.9	3.9	28.5	45.7

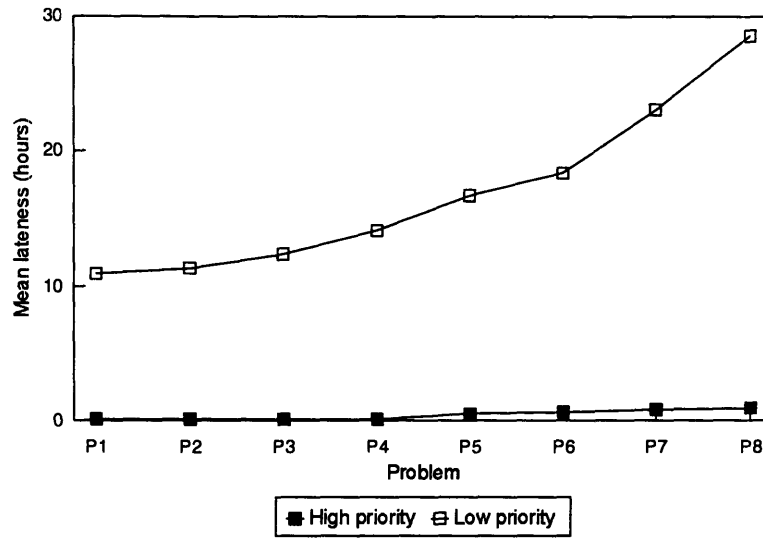


Figure 6.14 : Mean lateness to service standard for different traffic classes

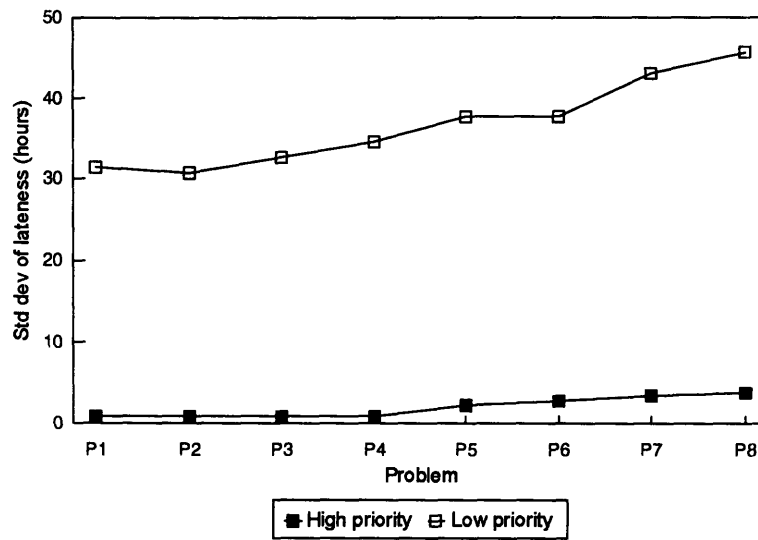


Figure 6.15 : Std dev of lateness to service standard for different traffic classes

6.5 Conclusions

From the case study, it is found that the model produces distinct car schedules for different O-D pairs, different traffic classes, different days of the week, and different departure times at origin terminal.

It maximizes the achievement of established service standards by minimizing total penalty costs. It produces clearly differentiated schedules for different classes of traffic by incorporating different service standards and different penalty costs. It also allows an effective plan on train capacity. From the results of the model, critical train capacity bottlenecks that were fully utilized over the entire planning period or a certain peak period could be found. By increasing the capacity of selected bottlenecks, a railroad may use the additional capacity to optimal service improvements.

In conclusion, the dynamic car routing and scheduling model can effectively support routing and scheduling heterogeneous traffic on rail freight networks and can improve the ability of railroads to differentiate services. By explicitly considering the expected train capacity problem, it produces more achievable car schedules than does the current car scheduling practice. Current car scheduling practice requires more controls for car-to-train connections at terminals, since it results frequently in over-subscribed trains. Considering car schedule information from the model, a railroad can commit more achievable trip plans to shippers and can improve service reliability. As a result of sensitivity analysis on traffic mixes, it is found that a railroad with limited service capacity has a limited capability to provide the premium service to shippers. This analysis allows a railroad to determine the maximum portion of traffic to which they can provide premium service.

The model is flexible for examining other issues. By changing the train capacity, we can examine the relation between capacity increase and service improvement. When a railroad changes the operating plans such as train schedules (e.g., train timetable change, extra trains, annulment),

blocking plan (e.g., new blocks) and make-up plan, the model can produce a new set of optimal car routes and schedules. This can be done by updating the time space network (see section 6.3.2) and re-optimizing the problem. Using an iterative procedure by changing input operating plans, a railroad may obtain a set of car schedule that best achieve the established service standards for current shippers.

To successfully implement the dynamic car scheduling system, a well-established information system plays an important role. The model needs regularly updated information on demand and terminal processing time. In this context, a railroad needs the capability of short-term demand forecasting and of continuous monitoring for terminal processing time. The model produces the information for dynamic car-to-block assignment. A railroad needs to develop a more sophisticated terminal work plan so that it achieves dynamic car-to-block assignment and improves schedule adherence for individual car moves through the terminal.

The dynamic car scheduling model developed in the study needs to be improved in several directions. The model assumed that only late arrivals are penalized and the penalty cost is linearly proportional to lateness. These assumptions need to be relaxed to examine more service design issues (i.e., establishing more tailored service standards for different shippers). It will be a nonlinear problem if they are relaxed. In the model, the block size limits and the integrality constraints are not explicitly considered. Including these constraints are needed for the model enhancement. More realistic representation of terminal operations may be needed for further enhancement of the model. Although the model considers a weekly traffic variation as one aspect of the variability in demand, a more sophisticated modeling approach is needed to fully consider more stochastic aspects in demand and in operations. The computational feasibility for larger networks is another very important issue.

Re-examining and re-structuring operating plans are very important issues for the purpose of service differentiation. This chapter shows that the dynamic car routing and scheduling model can effectively support routing and scheduling heterogeneous traffic on rail freight networks and can improve the ability of railroads to differentiate services.

The next chapter summarizes the important findings of the research and draws several conclusions and makes suggestions for additional research.

CHAPTER 7

Conclusions and Future Research

7.1 Summary and Conclusions

Recent studies have shown that the freight market can be divided into a number of market segments that have different preferences on elements of service quality and different willingness to pay for additional service improvement. These studies suggest that a railroad must target markets in which to compete and must properly position its services to be competitive and profitable in those markets; the ability of a railroad to differentiate its services can provide strategic intermodal and intramodal advantages in the competitive transportation environment.

The research in this thesis aimed to develop empirical and theoretical insights on how service differentiation strategy helps a railroad to gain market share and improve service quality.

Three areas were identified and researched to accomplish these objectives. First, the research examined how railroads are currently differentiating services among different groups of traffic. Second, the research developed insights into the effects of service differentiation on the service provided to segmented markets and the costs to railroad of providing differentiated services. Third, the research examined how to further improve the ability of a railroad to differentiate

services by designing an operating plan that fully considers the service requirements of different market segments. The primary findings and conclusions from the research are summarized as follows.

Empirical analysis of service differentiation

An empirical analysis of transit time and reliability of car movements was performed for three major train services: general merchandise train service, unit train service and intermodal train service. Car cycle information for three car types was collected for this purpose: box car data for general merchandise train service, covered hopper car data for unit train service, and double-stack car data for intermodal train service. Transit time and various reliability measures were evaluated and compared for different train services. Several important findings from the analysis are summarized as follows.

- There were clear differences in the trip time and reliability of the three different train services (Table 7.1). The service provided to box car traffic was significantly slower and less reliable than the other types of traffic. The service provided to double-stack cars was significantly faster and more reliable than the other types of traffic.
- There was also a certain level of variation in service levels among different O-D pairs for the same train service. It was not clear, however, if such differentiated service levels are the result of intentional efforts to differentiate service by considering the service requirements of individual O-D pairs, or if they simply reflect the result of day-to-day operations reacting to daily traffic variability and the uncertainty in various stages of operations.

Table 7.1 : Trip time and reliability performance of different train services

	<u>Box car</u>	<u>Hopper car</u>	<u>Double-stack</u>
Distance	788.1 miles	831.0	n/a
Mean trip time	7.16 days	5.25	2.53
Std dev of trip time	2.62 days	2.04	0.50
Maximum 1-day-%	32.42 %	41.90	89.2
Maximum 2-day-%	48.56 %	60.95	n/a
Maximum 3-day-%	61.07 %	73.21	n/a

We can conclude that railroads are currently differentiating services for major classes of freight traffic, with clear differences in the trip time and reliability of the three different train services. There was substantial unreliability in the level of service provided to general merchandise shippers. Shippers who use double-stack services, on the other hand, are able to take advantage of much faster and more reliable service. There was also a certain level of variation in service levels among different O-D pairs for the same train service.

To understand the reasons for these differentiated service levels between different train services and among different O-D pairs for the same train service, we would need to gather additional information on shipper's service expectation, carrier's operating policies for service differentiation, the competition among railroads, and the competition between rail and truck services. The lack of clarity about why these differentiated service levels currently exist also points out the need to develop better understanding of the effects of service differentiation.

Effects of service differentiation

Two probabilistic simulation models were designed and analyzed to develop insights into the effects of service differentiation in the rail freight transportation context.

- Two terminal model where a railroad provides a direct train service
- Five terminal model where a railroad provides car-load train services that require intermediate classification

These simple models that are designed for developing insights into the effects of service differentiation also can be helpful in looking at some real world situations, for example, unit train operations and general merchandise train operations.

It was assumed that shippers specify the priorities of their shipments and are willing to pay for service as a function of service quality. Service quality are defined by trip times and trip time reliability. As a mechanism to differentiate services, train make-up policies at the terminal, which are dependent on the shipper supplied shipment priority, are examined. These train make-up policies were examined under various resources defined by train capacity and empty car inventory level, and operating conditions defined by the variability of demand and train operations. Service levels of different classes of traffic and the costs to railroad of providing such differentiated services were evaluated. Several important findings from the analyses are summarized as follows.

- Both trip time and reliability for all traffic were improved as the system can dispatch more train capacity or has more empty car inventory. It indicates that there is a clear trade-off between service and cost. This result is consistent with the results of a previous study (Keaton [1991]) that showed more direct and frequent train services reduce the mean transit time, but increase costs. An important question is whether such a cost increase to improve overall service level is justified for traffic of all priorities.

- Different train make-up policies resulted in different trip time and reliability for different priority classes. Different levels of service can be provided for different traffic classes by implementing different operating plans (Table 7.2). Varying the way that different service requirements of traffic are incorporated into an operating plan results in different levels of service to different classes of traffic.

Table 7.2 : Trip time performance under different train make-up rules
(initial car inventory 700 cars, train capacity 130 cars)

<u>Priority</u>		<u>Train make-up rules</u>		
		<u>Rule 1</u>	<u>Rule 2</u>	<u>Rule 3</u>
High	Mean	2.00 days	2.10 days	2.00 days
	Std dev	0.00	0.21	0.00
Medium	Mean	2.06	2.34	2.21
	Std dev	0.22	0.43	0.33
Low	Mean	3.32	2.64	3.10
	Std dev	1.16	0.50	0.93
Total	Mean	2.28	2.28	2.28
	Std dev	0.75	0.43	0.63

- With an operating plan that properly considers different service requirements of different shippers, a railroad can provide fast and very reliable service to service-sensitive shippers and can provide slow and less reliable service to price-sensitive shippers.
- The analysis showed that efforts to provide highly reliable service were not justified for all market segments. A total logistics cost analysis was performed to measure the shippers' potential willingness to pay for the additional service improvement. It also allowed the determination of how much additional investment to increase the service capacity could be justified by potential market reaction in terms of the shippers' potential willingness to pay for the additional service improvement.

- Service differentiation strategy allows a railroad to utilize service capacity more efficiently, by focusing resources on the improvement of the service performance for selected market segments that are sensitive to service and willing to pay for these additional service improvements, instead of significant additional investment to improve the service performance for all traffic classes that may not require or be willing to pay for higher quality service.
- Cost per car for each priority is an important measure of effectiveness of service differentiation strategies, especially for pricing decisions and for management control purposes. A heuristic cost allocation method was used to obtain the operating cost by priority. Using it, high priority service had a higher cost and low priority service had a lower cost (Table 7.3). It suggests that railroads may differentiate on price for different priority service by providing different levels of service at different costs, and may make profits on all traffic classes.

Table 7.3 : Cost per car by priority for different train make-up rules

Priority	Train make-up rules		
	Rule 1	Rule 2	Rule 3
High	\$257.1	\$266.4	\$264.6
Medium	243.0	222.8	229.3
Low	168.8	175.5	170.5
Total	235.3	235.3	235.3

Several important insights were derived from the simulation analyses. Varying the way that heterogeneity of traffic is incorporated into an operating plan results in different levels of service to different classes of traffic. An operating strategy that properly considers heterogeneity of traffic enables a railroad to provide fast and highly reliable service to service-sensitive shippers and slow and less reliable (but less costly) service to cost-sensitive shippers. This strategy also

allows a railroad to utilize service capacity more efficiently, by focusing resources on the improvement of the service performance for selected market segments that are sensitive to service and willing to pay for these additional service improvements, instead of significant additional investment to improve the service performance for all traffic classes that may not require or be willing to pay for higher quality service. Furthermore, a railroad may improve profitability by differentiating prices based on distinct service level and the cost of providing such services for different market segments.

We can conclude that service differentiation strategies enable a railroad to provide market-sensitive services, to utilize service capacity more efficiently, and to potentially enhance profit.

A review of modeling efforts for service differentiation

One of the findings from the previous analysis indicates that varying the way that heterogeneity of traffic is incorporated into an operating plan results in different services and cost levels to different classes of traffic.

In this context, ways of incorporating market segmentation and service differentiation strategies into the service design process in general were discussed. The research also reviewed current hierarchical decisions (tactical and operational decisions) and practically- and theoretically-developed models for rail operations to examine whether these models properly considers the heterogeneity of traffic. Several important findings from the review are summarized as follows.

- In rail operations, the following is a useful classification of decisions. Tactical decisions have medium-term planning horizon and focus on effective allocation of existing resources. They include train routing, classification policy, train make-up policy and traffic routing.

Operational decisions deal with day-to-day activities in a fairly detailed and dynamic environment. They include train scheduling, car scheduling, engine scheduling, empty car allocation, crew scheduling and terminal work plan.

- The majority of models for tactical decisions considered the rail network operation as a relatively static problem by considering average demand and average train transit time. Second, the majority of tactical models considered the yard as a “black box” and assumed a certain fixed yard process and delay time. Finally, some of the previous tactical models considered the service reliability issue, but no previous models fully considered the need for establishing different service standards for different shipper groups. The majority of tactical models dealt with homogeneous traffic. This aspect combined with the cost minimization objective generates the design of services that are not sensitive to market service requirements.
- Compared with the models for tactical decisions, the majority of models for operational decisions explicitly considered the variability of demand and train transit time.
- Simulation and Service Planning Model (McCarren and Martland [1980]) approaches can explicitly incorporate the stochastic and dynamic nature of rail operations. The simulation approach, in principle, is capable of simulating detailed yard operation. On the other hand, the SPM incorporates the probabilistic yard process and delay time using the estimated PMAKE functions. However, these approaches can not be used to directly generate the operating plan.
- Little research can be found in the car scheduling area. The car schedule, in principle, can be generated from the train schedule, blocking plan and train make-up plan. However, these inputs only define a feasible set of schedules for car moves of each O-D pair. To plan car moves to improve service reliability and to provide market sensitive services, a railroad needs to develop a car scheduling system that also incorporates the service standards of different

shipper groups, more realistic train connection performance at yards, and the variability of demand.

The review leads us to conclude that a model for rail operating plan that fully considers the heterogeneity of traffic has not yet been established for the purpose of service differentiation.

Furthermore, we note that car scheduling is becoming a more important aspect of the operating plan, as railroads pursue more carefully scheduled and planned operations, and as more shippers demand car schedule information for planning their procurement, production and distribution processes. Current freight car routing and scheduling practice was identified by the review of models as a critical area to be improved for the purpose of service differentiation. Yet, no models to support improved car scheduling practice are available. This research aimed to fill this gap with a model that is the topic of the next section.

A dynamic freight car routing and scheduling model

Car scheduling can be a mechanism for effective service differentiation. However, current car scheduling practice ignores some important issues: the heterogeneity of traffic in terms of different service requirements, train capacity constraints and the characteristics of traffic variability in rail operation.

In response to this gap, we developed a dynamic car scheduling model that considers the heterogeneity of shippers and their traffic, traffic variability, and train capacity constraints so that it produces more market-sensitive and achievable car schedules.

A linear multicommodity flow problem was formulated to optimally route and schedule car movements for a rail network. A time-space network was used to represent rail operation and temporal car movements. The column-generation technique was used as a solution approach.

A case study was done for a network based on the data of a major US railroad. The network has 12 terminals and 16 trains are providing car-load train services each day. Daily average traffic volume is about 1,400 cars for the entire network. Computational time to obtain the optimal solution is about 1,300 seconds on IBM RS 6000/370 workstation. Several important findings from the case study are summarized as follows.

- The analysis showed that a car scheduling system that does not consider the train capacity constraints routes too much traffic to some trains. If there are over-subscribed trains, cars at corresponding terminals are more likely to miss the train connections due to train capacity problems and have long delays. A railroad needs to have more control of car-to-train connections at the terminal level, which have historically been done without considering their impact on system-level service performance, such as origin-destination trip time and on-time performance. For some train segments that are over-subscribed, such long train length may not be feasible because of power constraints.
- The dynamic car scheduling model developed in the research resolves this problem by explicitly recognizing train capacity constraints. By explicitly considering the expected train capacity problem, it produces more achievable car schedules than do the current car scheduling practices.
- The model generates differentiated car schedules for different classes of traffic for different days of the week. An example of output that summarizes the service level by priority is shown

in Table 7.4. This result supports that the model does allow a railroad to plan clearly differentiated services for different classes of traffic.

- As a result of sensitivity analysis on traffic mixes (i.e., the share of high and low priority traffic), it was found that a railroad can provide clearly differentiated services, provided that high priority traffic is not a large fraction of the traffic.

Table 7.4 : Service level by priority : sensitivity on traffic mixes

<u>Problem</u>	<u>High priority share (%)</u>	<u>High priority</u>		<u>Low priority</u>	
		<u>Mean lateness (hours)</u>	<u>Std dev lateness (hours)</u>	<u>Mean lateness (hours)</u>	<u>Std dev lateness (hours)</u>
P1	31.6	0.1	0.9	10.9	31.4
P2	33.2	0.1	0.9	11.3	30.8
P3	39.0	0.1	0.9	12.3	32.7
P4	46.2	0.1	0.8	14.1	34.6
P5	55.8	0.5	2.2	16.7	37.6
P6	60.6	0.6	2.7	18.4	37.7
P7	69.2	0.8	3.4	23.0	43.1
P8	75.8	0.9	3.9	28.5	45.7

- Although more computational experiments are needed to prove whether the model can be implemented for larger networks, the basic feasibility of the model to support improved car scheduling practice was demonstrated through the case study.

In conclusion, the dynamic car scheduling model developed in the research effectively supports the routing and scheduling of heterogeneous traffic on rail networks. Using the model, a railroad can improve its ability to differentiate services for different classes of traffic. To achieve this, we also need to enhance following elements that support the car scheduling activity.

- A well-established information system is needed to support the dynamic car scheduling activity. The implementation of the model requires the capability of regularly updated short-term demand forecasting and terminal processing times.
- In addition, the dynamic car scheduling system requires a dynamic block definition for dynamic car-to-block assignment. A railroad needs to develop a more sophisticated terminal work plan so that it can achieve the schedule adherence necessary for individual car movements through the terminal.

7.2 Contributions

An improved perspective on service management of rail freight transportation is developed in this thesis. Service differentiation, which incorporates the service requirements of heterogeneous shippers in designing and delivering services, can potentially be a very effective strategy for a railroad to utilize in its efforts to design and delivering market-sensitive services. In this thesis, three specific contributions have been achieved.

- First, to our knowledge, the first large-scale systematic assessment of actual O-D trip times and reliability of rail freight service in North American railroads was made.
- Second, the study developed insights into the effects of service differentiation in the rail freight transportation context and showed that service differentiation can potentially be a very effective strategy for a railroad. It was demonstrated that, with operating plans that are properly considers different service requirements of shippers and their traffic, a railroad can

differentiate services for different classes of traffic at cost levels that are appropriate to those service levels.

- Finally, a practical decision support model for routing and scheduling heterogeneous traffic on rail freight networks was developed. Using the model, a railroad can improve its ability to differentiate services for different classes of traffic.

7.3 Suggestions for Future research

In this section, areas for future research are suggested and discussed.

The empirical analysis of service differentiation in Chapter 3 focused only on trip time and reliability of different groups of traffic. Due to the lack of information, the causes of differentiated trip times and reliability were not fully explained. Collection of more data on possible explanatory variables (e.g., number of intermediate terminal handling) is necessary for the causality analysis to be conducted. The analysis of other elements of service quality (e.g., price, loss and damage, other customer services) is also an important area of future research.

The effects of service differentiation were evaluated using simulation in Chapter 4. As a mechanism to differentiate services, train make-up policies were examined. Other operating plans, such as a blocking plan, train schedules and train dispatching policy at terminals can be examined using a rail network simulation model developed in the study (Appendix A). More simulation experiments which consider other operating plan options on larger and more realistic networks will provide better understanding of the effects of service differentiation on service and cost in rail operations.

A dynamic freight car routing and scheduling model was developed in Chapter 6. Computational experiments for larger networks are needed since the computational feasibility for larger networks is a very important issue from the implementation point of view. The block size and integrality constraints need to be included for the model enhancement. Although the model considers a weekly traffic variation as one aspect of the variability in demand, a more sophisticated modeling approach is needed to fully consider more stochastic aspects in demand and in operations. Incorporating heterogeneity of traffic into decisions in rail operations can also be extended to other operating plans such as a blocking plan and train scheduling.

Finally, a comprehensive case study for a specific railroad, including all the research areas we examined in this study, may allow more concrete and applicable conclusions.

7.4 A Final Comment

The passage of the Staggers Act in 1980, which removed much federal economic regulations from the rail industry, provided the managers of U.S. railroads more freedom to earn a fair return on investment. Railroads have undertaken a number of initiatives to rationalize their rate structure, input utilization, and scale of operations to increase returns to competitive levels. In addition, there have been continuous efforts to improve the overall service reliability in the rail industry to be more service competitive in the market.

The freight market can be divided into a number of market segments that have different preferences on elements of service quality and willingness to pay for additional service improvement. It is not possible to implement any single marketing and service management program that will satisfy all current and potential shippers. A railroad must target markets in

which to compete and must properly position its services to be competitive and profitable in those markets. The ability of a railroad to differentiate its services can provide a strategic advantage in the competitive transportation environment.

This thesis demonstrates that an operating strategy that properly considers distinct service requirements of different market segments allow a railroad to provide market-sensitive services. It also allows a railroad to utilize existing resources to provide services more efficiently, by focusing resources on the improvement of services for selected market segments that are sensitive to service and willing to pay for these additional service improvements, instead of significant additional investment to improve the service performance for all traffic.

Service differentiation has the potential to be a very effective strategy for a railroad as a next step to utilize in its efforts to gain market share and enhance profit. It is hoped that this thesis can be useful to the rail industry in applying service differentiation strategies to its operations.

REFERENCES

- Adamidou, E. L., A. L. Kornhauser, and Y. A. Koskosidis, A Game Theoretic/Network Equilibrium Solution Approach for the Railroad Freight Car Management Problem, *Transportation Research-B*, Vol. 27B, No. 3, pp. 237-252, 1993
- Ahuja, R. K., T. L. Magnanti, and J. B. Orlin, *Network Flows: Theory, Algorithms, and Applications*, Prentice Hall, 1993
- ALK Associates, The Automated Blocking Model, User's Manual, 1986
- Allen, W. B., M. Mahmoud, and D. McNeil, The Importance of Time in Transit and Reliability of Transit Time for Shippers, Receivers, and Carriers, *Transportation Research-B*, Vol. 19B, No. 5, pp. 447-456, 1985
- Armstrong, J. H., *The Railroad What It Is, What It Does: The Introduction to Railroading*, Simmons-Boardman Books, Inc., 1990
- Assad, A. A., Models for Rail Transportation, *Transportation Research-A*, Vol. 14A, pp. 205-220, 1980
- Assad, A. A., Modeling of Rail Networks: Toward a Routing/Makeup Models. *Transportation Research-B*, Vol. 14B, pp. 101-114, 1980
- Assad, A. A., Analytical Models in Rail Transportation: An Annotated Bibliography, *INFOR*, Vol. 19, No. 1, pp. 59-80, 1981
- Assad, A. A., A Class of Train-Scheduling Problem, *Transportation Science*, Vol. 16, No. 3, pp. 281-310, 1982
- Association of American Railroads, AAR Network Simulation Model: User's Manual, 1971
- Association of American Railroads, Ad Hoc Committee to Develop ETA and Trip Plan Capabilities Among Railroads, A Proposal for Systems to Support Interline Service Management, AAR Report R-776, 1991
- Association of American Railroads, ISM Task Force Team, Interline Service Management: Functional Requirements, AAR Report R-836, 1994
- Barnhart, C., E. Johnson, E. L. Anbil, and L. Hatay, A Column Generation Technique for the Long Haul Crew Assignment Problem, *Optimization in Industry, Volume II*, John Wiley and Sons, 1993

Blumenfeld, D. E., L. D. Burns, and J. D. Diltz, Analyzing Trade-Offs between Transportation, Inventory and Production Costs on Freight Networks, *Transportation Research-B*, Vol. 19B, No. 5, pp. 361-380, 1985

Bodin, L. D., B. L. Golden, and A. D. Schuster, A Model for the Blocking of Trains, *Transportation Research-B*, Vol. 14B, pp. 115-120, 1980

Bongaardt, H. L., E. L. Clausing, and G. F. List, Railroad Network Modeling: Recent Practical Applications, *Proceedings of 21st Annual Meeting*, Transportation Research Forum, Philadelphia, Pennsylvania, October, 1980

Chiang, Y., P. O. Robert and M. Ben-Akiva, *Development of a Policy Sensitive Model for Forecasting Freight Demand*, US Department of Transportation, 1981

Chih, K. C., A Real Time Dynamic Optimal Freight Car Management Simulation Model of the Multiple Railroads, Multicommodity Temporal Spatial Network Flow Problem, Doctoral Dissertation, Department of Civil Engineering, Princeton University, 1986

Chih, K. C., M. S. Rothenberg, and A. L. Kornhauser, Implementing of a Real Time Locomotive Distribution system, *Proceedings of the Second International Conference on Computer Aided Design, Manufacture and Operation in the Railway and other Advanced Mass Transit Systems*, Rome, Italy, March, 1990

Crainic, T., J. Ferland and J. Rousseau, A Tactical Planning Model for Rail Freight Transportation, *Transportation Science*, Vol. 18, No. 2, pp. 165-184, 1984

Crainic, T., J. Rousseau, The Column Generation Principle and the Airline Crew Scheduling Problem, *INFOR*, Vol. 25, No. 2, pp. 136-151, 1987

Crainic, T. and J. Rousseau, Multicommodity, Multimode Freight Transportation: A General Modeling and Algorithmic Framework for the Service Network Design Problem, *Transportation Research-B*, Vol. 20B, No. 3, pp. 225-242, 1986

Crainic, T. G., M. Gendreau, and P. Dejax, A Dynamic Stochastic Model for the Allocation of Empty Containers, Publication No. 713, Centre de recherche sur les transports, University of Montreal, 1990

CSX Transportation, Service Design, Overview of Yield Management: Focusing Service Capacity Utilization, January 1992

Dantzig, G. B. and P. Wolfe, Decomposition Principles for linear programs, *Operations Research*, Vol. 8, pp. 101-111, 1960

Dejax, P. J. and T. D. Crainic, A Review of Empty Flows and Fleet Management Models in Freight Transportation, *Transportation Science*, Vol. 21, No. 4, pp. 227-247, 1987

Desrochers, M. and F. Soumis, A Column Generation Approach to the Urban Transit Crew Scheduling Problem, *Transportation Science*, Vol. 23, No. 1, pp. 1-13, 1989

- Desrochers, M., J. Desrosiers, and M. Solomon, A New Optimization Algorithm for the Vehicle Routing Problem with Time Windows, *Operations Research*, Vol. 40, No. 2, pp. 342-354, 1992
- Desrosiers, J., F. Soumis, and M. Desrochers, Routing with Time Window by Column Generation, *Networks*, Vol. 14, pp. 545-565, 1984
- Duffy, M. A., Statistical Process Control Applied to Rail Freight Terminal Performance: A Case Study of CSX's Radnor Yard, MS Thesis, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, 1994
- Florian, M., G. Bushell, and J. Ferland, The Engine Scheduling Problem in a Rail Network, *INFOR*, Vol. 14, No. 2, pp. 121-138, 1976
- Folk, J. F. and J. M. Sussman, Unreliability in Railroad Network Operations, *Systems Planning and Design: Case Studies in Modeling, Optimization, and Evaluation*, Prentice-Hall, 1974
- Friedlaender, A. F., E. R. Berndt, and G. McCullough, Governance Structure, Managerial Characteristics, and Firm Performance in the Deregulated Rail Industry, *Brookings Papers: Microeconomics 1992*, pp. 95-186
- Friedlaender, A. F., Fair Rates and Rates of Return in a Deregulated Rail Industry, MIT Department of Economics Working Paper 576, Massachusetts Institute of Technology, 1991
- Gopal, C. and H. Cypress, *Integrated Distribution Management: Competing on Customer Service, Time, and Cost*, Business One Irwin, 1993
- Haghani, A. E., Formulation and Solution of a Combined Train Routing and Makeup, and Empty Car Distribution Model, *Transportation Research-B*, Vol. 23B, No. 6, pp. 433-452, 1989
- Hornung, M., H. Rosen, and J. Szymkowiak, Design and Prototyping of a Motive Power Information and Management Support System, *Proceedings of the Second International Conference on Computer Aided Design, Manufacture and Operation in the Railway and other Advanced Mass Transit Systems*, Rome, Italy, March, 1990
- IBM, Optimization Subroutine Library: Guide and Reference, Release 2, 1991
- Johnson, L. A. and D. C. Montgomery, *Operations Research in Production, Scheduling, and Inventory Control*, Wiley, New York, 1974
- Jordan, W. C. and M. A. Turnquist, A Stochastic, Dynamic Network Model for Railroad Car Distribution, *Transportation Science*, Vol. 17, No. 2, pp. 123-145, 1983
- Jovanovic, D. and P. T. Harker, A Decision Support system for Train Dispatching: An Optimization-Based Methodology, *Journal of Transportation Research Forum*, Vol. 31, No. 1, pp. 25-37, 1990
- Jovanovic, D. and P. T. Harker, Tactical Scheduling of Rail Operations: The SCAN I System, *Transportation Science*, Vol. 25, No. 1, pp. 46-64, 1991

Keaton, M. H., Designing Optimal Railroad Operating Plans: Lagrangian Relaxation and Heuristic Approaches, *Transportation Research-B*, Vol. 23B, No. 6, pp. 415-431, 1989

Keaton, M. H., Service-Cost Tradeoffs for Carload Freight Traffic in the U.S. Rail Industry, *Transportation Research-A*, Vol. 25A, No. 6, pp. 363-374, 1991

Keaton, M. H., Designing Railroad Operating Plans: A Dual Adjustment Method for Implementing Lagrangian Relaxation, *Transportation Science*, Vol. 26, No. 4, pp. 263-279, 1992

Keaton, M. H., The Impact of Train Timetables on Average Car Time in Rail Classification Yards, *Journal of Transportation Research Forum*, Vol. 32, No. 2, pp. 345-354, 1992

Kernighan, B. W. and D. M. Ritchie, *The C Programming Language*, Prentice Hall, 1988

Kerr, P. A., C. D. Martland, J. M. Sussman, and C. E. Philip, Models for Investigating Train Connection Reliability at Rail Classification Yards, *Studies in Railroad Operations and Economics*, Vol. 14, Massachusetts Institute of Technology, 1976

Kraft, E. R., A Branch and Bound Procedure for Optimal Train Dispatching, *Journal of Transportation Research Forum*, Vol. 28, No. 1, pp. 263-276, 1987

Kraft, E. R., Demand Variability, Service Reliability and Railway Freight Research Needs, Working Paper 94-04-03, Department of Operations and Information Management, The Wharton School, University of Pennsylvania, 1994

Lang, A. S. and C. D. Martland, Reliability in Railroad Operations, *Studies in Railroad Operations and Economics*, Vol. 8, Massachusetts Institute of Technology, 1972

Larsen, R. J. and M. L. Marx, *An Introduction to Mathematical Statistics and Its Applications*, Prentice-Hall, 1986

Law, A. M. and W. D. Kelton, *Simulation Modeling and Analysis*, McGraw-Hill, Inc., 1991

Levin, R. C., Railroad Rates, Profitability, and Welfare under Deregulation, *The Bell Journal of Economics*, Vol. 12, Spring, pp. 1-26, 1981

Little, P., O. K. Kwon, and C. D. Martland, An Assessment of Trip time and Reliability of Box Car Traffic, *Proceedings of 34th Annual Forum*, Transportation Research Forum, St. Louis, Missouri, 1992

Little, P. and C.D. Martland, Causes of Unreliable Services in North American Railroads, *Proceedings of 35th Annual Forum*, Transportation Research Forum, New York, New York, October, 1993

Martland, C. D., Rail Trip Time Reliability: Evaluation of Performance Measures and Analysis of Trip Time Data, *Studies in Railroad Operations and Economics*, Vol. 2, Massachusetts Institute of Technology, 1972

- Martland, C. D., Improving Railroad Reliability: A Case Study of the Southern Railway, *Studies in Railroad Operations and Economics*, Vol. 10, Massachusetts Institute of Technology, 1974
- Martland, C. D., P. A. Clappison, C. D. Van Dyke, and R. J. Tykulsker, Rail Service Planning: A Case Study of the Santa Fe, *Studies in Railroad Operations and Economics*, Vol. 35, Massachusetts Institute of Technology, 1981
- Martland, C. D., PMAKE Analysis: Predicting Rail Yard Time Distribution using Probabilistic Train Connection Standards, *Transportation Science*, Vol. 16, No. 4, 1982
- Martland, C. D., P. Little, M. Duffy, and Y. Dong, Development of Concepts for Linking Railroad Terminal Control Systems to Advanced Line Control Systems, Working Paper, Center for Transportation Studies, Massachusetts Institute of Technology, 1993
- Martland, C. D., P. Little, O. K. Kwon, and R. Dontula, Background on Railroad Reliability, Report No. R-803, Association of American Railroads, 1992
- Martland, C. D., P. Little, and J. M. Sussman, Service Management in the Rail Industry, *AAR/FRA/TRB Conference on Railroad Freight Transportation Research Needs*, Bethesda, Maryland, July, 1993
- McCarren, J. R. and C. D. Martland, The MIT Service Planning Model, *Studies in Railroad Operations and Economics*, Vol. 31, Massachusetts Institute of Technology, 1980
- McGinnis, M. A., Segmenting Freight Markets, *Transportation Journal*, Fall, 1978, pp. 58-68
- McGinnis, M. A., A Comparative Evaluation of Freight Transportation Choice Models, *Transportation Journal*, Winter, 1989, pp. 36-46
- Mercer Management Consulting, The Service Quality Challenge for the 1990s, *The 5th American Railroad Conference*, November, 1991
- Mercer Management Consulting, *Intermodal Index*, sponsored by Intermodal Association of North America and The National Industrial Transportation League, 1992, 1993 issues
- Misra, S. C., Linear Programming of Empty Wagon Disposition, *Rail International*, March, 1972, pp. 151-158
- Missouri Pacific Railroad, Missouri Pacific's Computerized Freight Car Scheduling System: System Functional Requirements, Report No. FRA-OPPD-77-10, Federal Railroad Administration, U.S. Department of Transportation, 1977
- Morgenbesser, M. J. and C. D. Martland, The Economics of Running More Trains: A Case Study of the Louisville and Nashville Railroad, Multisystems, 1979
- Petersen, E. R., A. J. Taylor, and C. D. Martland, An Introduction to Computer-Assisted Train Dispatch, *Journal of Advanced Transportation*, Vol. 20, No. 1, pp. 63-72, 1986

- Philip, C. E., The Application of an Inventory Model to the Railroad Empty Car Distribution Process, *Studies in Railroad Operations and Economics*, Vol. 25, Massachusetts Institute of Technology, 1978
- Powell, W. B., Y. Sheffi, and K. S. Nickerson, Maximizing Profits for North American Van Lines' Truckload Division: A New Framework for Pricing and Operations, *INTERFACES*, Vol. 18, No. 1, pp. 21-41, 1988
- Roberts, P. O., M. N. Terziev, J. T. Kneafsey, and L. B. Wilson, Analysis of the Incremental Cost and Trade-offs between Energy Efficiency and Physical Distribution Effectiveness in Intercity Freight Markets, CTS Report Number 76-14, Center for Transportation Studies, Massachusetts Institute of Technology, 1976
- SAS Institute Inc., SAS Language and Procedures, Version 6, 1989
- SAS Institute Inc., SAS Companion for the CMS Environment, Version 6, 1990
- Sauder, R. L. and W. M. Westerman, Computer Aided Train Dispatching: Decision Support through Optimization, *INTERFACES*, Vol. 13, No. 6, pp. 24-37, 1983
- Shapiro, J. F., *Mathematical Programming: Structure and Algorithms*, John Wiley and Sons, 1979
- Sheffi, Y., Carrier/Shipper Interactions in the Transportation: An Analytical Framework, *Journal of Business Logistics*, Vol. 7, No. 1, 1977, pp. 1-27
- Shen, P., Intercity Freight Transportation Demand Forecast and Model Systems, *Proceedings of 34th Annual Forum*, Transportation Research Forum, St. Louis, Missouri, October, 1992
- Smith, M. E., Keeping Trains On Schedule: On-Line Planning Systems for the Advanced Railroad Electronics System (ARES), *Journal of Transportation Research Forum*, Vol. 31, No. 1, pp. 17-24, 1990
- Smith, M. E. and R. R. Resor, Value of High-Quality Service: How Should the ARES-Equipped Railroad Operate?, *Transportation Research Record 1314*, 1991
- Smith, S. and Y. Sheffi, Locomotive Scheduling under Uncertain Demand, *Transportation Research Record 1252*, pp. 45-53, 1989
- Turnquist, M. A. and M. S. Daskin, Queuing Models of Classification and Connection Delay in Railyards, *Transportation Science*, Vol. 16, pp. 207-230, 1982
- Van Dyke, C. D., M. M. Perkins, and C. D. Martland, User's Guide to the MIT Service Planning Model, Center for Transportation Studies, Massachusetts Institute of Technology, 1981
- Van Dyke, C. D., The Automatic Blocking Model: A Practical Approach to Freight Railroad Blocking Plan Development, *Proceedings of 27th Annual Meeting*, Transportation Research Forum, Seattle, Washington, September, 1986

Van Dyke, C. D., Dynamic Management of Railroad Blocking Plans, *Journal of Transportation Research Forum*, Vol. 29, No. 1, pp. 149-152, 1988

Van Dyke, C. D. and L. Davis, Software Tools for Railway Operations/Service Planning: the Service Planning Model Family of Software, *Proceedings of the Second International Conference on Computer Aided Design, Manufacture and Operation in the Railway and other Advanced Mass Transit Systems*, Rome, Italy, March, 1990

Vieira, L. F. M., The Value of Service in Freight Transportation, Doctoral dissertation, Massachusetts Institute of Technology, 1992

Yagar, S., F. F. Saccomando and Q. Shi, An Efficient Sequencing Model for Humping in a Rail Yard, *Transportation Research-A*, Vol. 17A, No. 4, pp. 251-262, 1983

Yan, S., Intermodal Pricing and Operations Using Network Flow Techniques, Doctoral Dissertation, Department of Civil Engineering, Massachusetts Institute of Technology, 1991

Wang, S., Service Reliability of Double-Stack Container Trains in the United States, MS Thesis, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, 1993

White, W. W., Dynamic Transshipments Networks: An Algorithm and its Application to the Distribution of Empty Containers, *Networks*, Vol. 2, pp. 211-236, 1972

Appendix A

Rail Network Simulation Model

A.1 Structure of simulation model

An event-based stochastic rail network simulation model is developed to evaluate the effects of various operating practices on the service and cost performance of rail freight transportation operations. The model simulates detailed movements of individual cars on a rail network given hierarchical operating plans, resources, and stochastic demand and operating conditions. It evaluates O-D trip time and reliability, facility and equipment utilization and operating costs for a given simulation scenario. The simulation model is coded using the C programming language.

The details of simulating car moves on a rail network is discussed in the section A.2. The input and output of simulation model is described in the section A.3.

A.2 Simulating car movements on the rail network

The simulation model developed in the research focuses more on the terminal activities than the line activities. A railroad can influence car movements in different locations of a rail network.

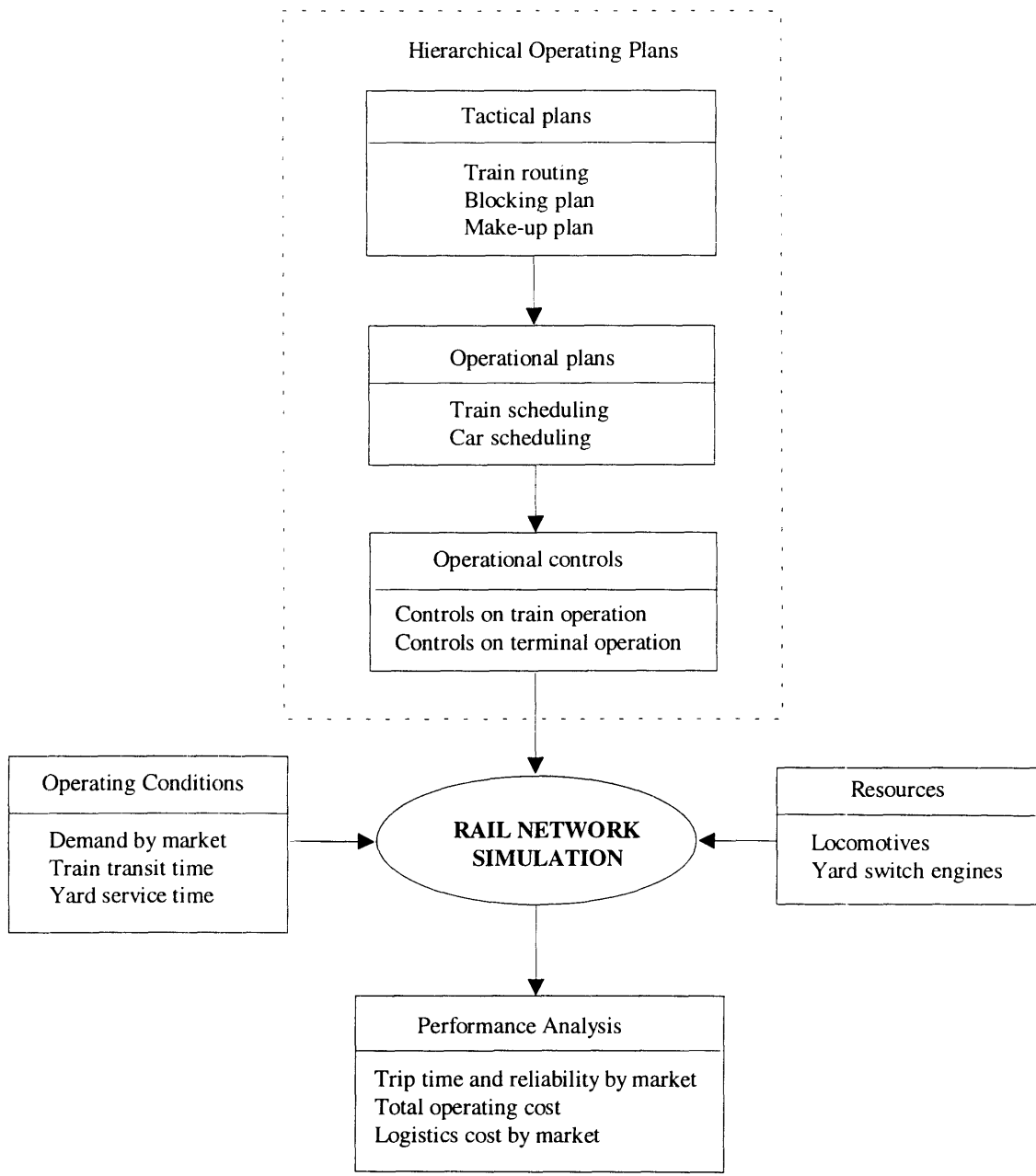


Figure A.1 : Structure of simulation model

In the line sections, it can control train moves and cars on those trains as a whole, but it cannot effectively control individual car moves. On the other hand, in the terminal, it can control both train moves and individual car moves. For example, receiving and dispatching policies effect train moves. Classification and make-up plans and possible control actions have direct effects on individual car moves.

In the model, car moves through terminal activities are simulated in a detailed manner but train operations in a line section are not simulated. Instead, the train transit time distributions are given as an input to the model. The major functions of a classification terminal are to receive inbound trains, inspect and repair cars, classify cars into blocks based upon blocking plan, and assemble blocks for outbound trains based upon a make-up plan. Figure A.2 shows the car movement through terminal and related operating plans and controls.

Inbound trains arrive at the receiving area of a terminal where they are inspected and join the queue for classification activity. Train priority is considered in receiving and inspection. When several inbound trains are waiting for receiving and inspection, the terminal manager determines an order of trains to be handled. If trains have same priority, they are handled on the "first come-first serve" basis. In the simulation, inbound trains are handled on the FCFS basis and a constant time of T_i hour is assumed for train arrival to completion of inspection.

The goal of classification process is to sort cars of the inbound train into blocks that will be moved together to the next common destination on their routes to final destinations. The operating plan for this process is the blocking plan. In the model, service time for classification of an inbound train is computed as

$$T_C = t_F + \sum_{i=1}^{X_t} t_C^i$$

where, T_C is total time for classification of inbound train, t_F is fixed set-up time for classification, t_C^i is service time for i -th car, and X_j is total number of cars on inbound train. Service time for each car is sampled from a classification time distribution of a given terminal.

After the classification, cars are usually waiting for train make-up. The waiting time from the point a car is classified to the point the appropriate train is to be made up depends on the frequency and schedule of train departures that a car can be connected. As the frequency of feasible train departure is increased, waiting time of a car will be decreased. It also depends on the policy on dispatching trains from the terminal. When a railroad implements a flexible dispatching policy that delays train departure to maintain certain minimum train length, cars that will be connected to this train may have additional delay. On the other hand, a benefit of this type of policy is to increase train capacity utilization and to reduce delays due to missed connection for some cars that arrived later than scheduled and could not make planned train connections. However, this type of policy influences car moves through downstream terminals and system-wide effects should be considered.

The next stage of major terminal operation is the train make-up process. Any outbound train has a "take-list" that specifies, in the order of preference, the blocks of cars it may pick up from the classification tracks. In the model, outbound train length is determined based on designed train length and the number of currently available locomotives in the terminal. Then the yard switch engine pulls blocks sequentially from the classification tracks to departure track based on the order specified in the take list. If there are not enough locomotives to dispatch a designed train capacity, or the number of cars that can be connected to a train is longer than dispatchable train capacity, some blocks of cars will miss their appropriate train connection and will be delayed until the next available train departure. In the state-of-the-practice, train make-up process is mostly

done in a static manner. It means that the yard switch engine simply pulls blocks sequentially based on the order specified in the take-list.

It is necessary to start make up of outbound train some time in advance of its scheduled departure time. In the model, when the time for train make-up of an outbound train is reached, the model checks the availability of locomotives and determines the train length to be dispatched. Then the model checks the availability of switch engine. If the switch engine is available, train is assembled up to dispatchable train length. Otherwise, train make-up is delayed until yard switch engine is available. Service time for this process is computed as

$$T_M = t_F + \sum_{i=1}^{X_O} t_M^i$$

where, T_M is total time for make-up outbound train, t_F is fixed set-up time for make-up, t_M^i is service time for i -th block, and X_O is total number of blocks connected for outbound train. Service time for each block is sampled from a make-up time distribution of given terminal. We need to note that the basic handling unit in this classification process is the car, but it is the block in the make-up process.

After the train make-up process and outbound inspection, the train departs from the terminal. Train transit time to the next terminal is sampled from input train transit time distribution of the corresponding line section. Train arrival time at the next terminal is then determined. For the next terminal, the same process is applied to simulate car moves through the terminal.

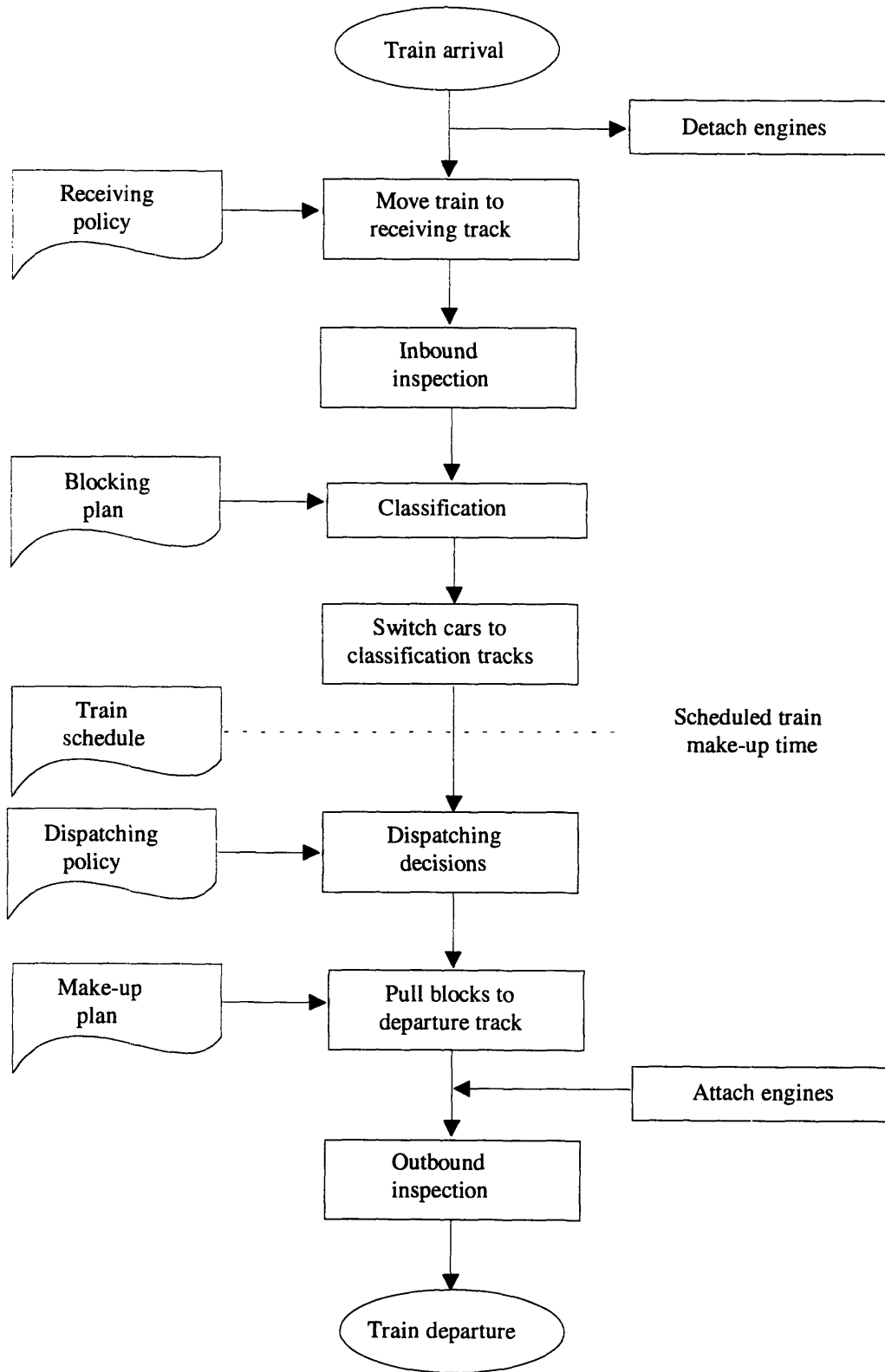


Figure A.2 : Car movement through terminal and related operating plans

A.3 Model input

Hierarchical operating plans

The hierarchical operating plans for simulation either can be supplied from the existing plans of a railroad or can be designed by various decision support models depending upon the type of study and data availability. The simulation model considers the following operating plans.

-
- Train schedule
 - Blocking plan
 - Train make-up plan
 - Car schedule
 - Controls in terminal operations
 - Controls on train make-up
 - Train and power dispatching policies at yards
-

Fixed resources

There are a number of resources required for rail operations. In the simulation model, some important resources such as number of locomotives and yard engines in terminals are considered.

Locomotives are the most important resource to dispatching designed train capacity. If there is not enough power, designed train capacity cannot be dispatched and some fraction of cars miss their appropriate train connections. A recent analysis of major railroads' data showed that among the 34 reasons given for cars failing to meet standard, 25% were due to power shortage.

Although this is a limited analysis, this figure indicates the importance of power resource in providing reliable service (see Little and Martland [1993]).

Yard engines are also important resources to determine the car handling capacity in the terminal. A sufficient number of yard engines and crews are needed to maintain car handling efficient. If there are not enough yard engines, service times for classification (especially in the flat yard) and train make-up are likely to be increased.

Stochastic demand and operating conditions

Stochastic demand and operating conditions considered are conditions on demand, train transit time in line sections, and service time for various operations in terminals.

Demand is the most important exogenous condition to be considered in designing tactical operating plans. In designing tactical operating plans, average demand pattern is usually considered (see Chapter 5). In daily operations, however, the variability of demand often causes capacity-related delays in car moves. In the simulation, demand is considered by including the daily demand distribution and weekly demand pattern of each market segmented by origin, destination and traffic class.

Furthermore, since the model does not attempt to simulate train operations in line, train transit time distributions for line sections are simply included as an input to the model. Service time distributions for classification and make-up servers (yard engine and crew) are also included as inputs to the model.

A.4 Model output

Trip time, reliability and operating costs are measured and evaluated as a result of simulation.

These performance measures are evaluated by market as well as for overall system.

OD trip time and reliability performance

Following trip time and reliability performance are measured by market.

Average actual OD trip time
Scheduled OD trip time
Standard deviation of OD trip time
Percent arrived on-time
Percent arrived N-hour late

Train transit time and utilization performance

Following train operation performance are measured for each train leg along with train route.

Scheduled departure time from departure terminal
Average actual departure time from departure terminal
Standard deviation of departure time from departure terminal
Average train length
Average train capacity utilization

Yard time and equipment utilization performance

Following performance are measured for each terminal.

Scheduled yard time
Average yard time
Standard deviation of yard time
Average daily hump utilization
Average daily switch engine utilization

Total operating costs

To estimate the operating costs, we need both unit cost and operating performance information.

Following performance are measured to estimate the operating costs.

T_m^c : total car-hours during analysis period for market m

Y_{pq}^t : total train miles during analysis period for train leg $q \in s_p$

Z_{pq}^l : total number of locomotives used during analysis period for train leg $q \in s_p$

Y_{pq}^l : total locomotive-miles during analysis period for train leg $q \in s_p$

T_{pq}^l : total locomotive-hours during analysis period for train leg $q \in s_p$

T_k^h : total switch-engine-hours during analysis period for terminal k

Following unit cost information are needed as the input to the model.

c_1 : crew cost per train-mile

c_2 : locomotive ownership cost per locomotive-day

c_3 : locomotive maintenance cost per locomotive-mile

c_4 : locomotive maintenance cost per locomotive-hour

c_5 : locomotive fuel cost per locomotive-mile

c_6 : locomotive fuel cost per locomotive-hour

c_7 : car-time cost per car-hour

c_8 : cost per switch-engine-hour

The unit cost information used for the study is obtained from the previous studies by Morgenbesser and Martland [1979] and by ALK Associates [1986].

Train crew cost : \$3 per train-mile

Locomotive ownership cost : \$100 per locomotive-day

Locomotive maintenance cost : \$2.03 per locomotive-mile and \$0.32 per locomotive-hour

Locomotive fuel cost : \$2.28 per locomotive-mile and \$0.78 per locomotive-hour

Car-time cost : \$18 per car-day

Unit switch engine hour cost : \$96.44 per switch-engine-hour

Based on this information, total train costs, total car-time costs and total car switching costs are computed. Total operating cost is computed as the sum of total train cost, total car hour cost and total car switching cost.

Total train cost

$$TC_1 = c_1 \sum_p \sum_{q \in s_p} Y_{pq}^l + c_2 \sum_p \sum_{q \in s_p} Z_{pq}^l + (c_3 + c_5) \sum_p \sum_{q \in s_p} Y_{pq}^l + (c_4 + c_6) \sum_p \sum_{q \in s_p} T_{pq}^l$$

Total car-time cost

$$TC_2 = c_7 \sum_m T_m^c$$

Total car switching cost

$$TC_3 = c_8 \sum_k T_k^h$$

Total operating cost

$$TC = TC_1 + TC_2 + TC_3$$

Appendix B

Input Data for Dynamic Freight Car Routing and Scheduling Model

This Appendix describes the train schedule, blocking plan, make-up plan, and O-D demand information that are major inputs to the model.

B.1 Train schedule

Train number	Train capacity	Terminal	Arrival	Departure	Day
162	120 cars	5	-	04:00	1
		4	08:00	08:50	1
		3	10:45	-	1
227	80	3	-	10:05	1
		6	12:35	12:55	1
		7	14:10	14:40	1
		9	19:00	-	1
249	80	3	-	08:00	1
		2	10:40	11:10	1
		1	15:15	-	1
261	160	3	-	18:00	1
		4	20:00	20:45	1
		5	01:00	-	2
307	120	8	-	07:00	1
		7	09:30	10:50	1
		9	16:40	-	1
703	120	9	-	19:00	1
		7	23:15	24:00	1
		8	02:05	-	2
718	160	12	-	17:00	1
		9	20:20	20:50	1

		11	22:30	22:55	1
		10	00:30	-	2
722	80	9	-	01:00	1
		7	05:20	05:55	1
		6	06:10	06:50	1
		3	08:10	-	1
729	160	9	-	09:30	1
		7	12:30	13:00	1
		6	13:40	13:45	1
		3	15:00	15:35	1
		2	16:55	17:00	1
		1	19:15	-	1
747	160	9	-	20:50	1
		12	24:00	-	1
757	160	12	-	05:00	1
		9	08:20	-	1
809	200	10	-	02:00	1
		11	03:40	04:10	1
		9	05:40	06:30	1
		7	08:50	09:40	1
		6	10:20	10:25	1
		3	11:30	12:00	1
817	160	10	-	17:00	1
		11	18:40	19:10	1
		9	21:00	21:45	1
		12	01:00	-	2
908	200	1	-	02:00	1
		2	02:30	03:35	1
		3	04:40	05:00	1
		6	06:00	06:05	1
		7	06:40	06:55	1
		9	09:20	-	1
927	160	1	-	09:30	1
		2	10:30	10:35	1
		3	11:55	12:25	1
		6	13:30	13:35	1
		7	14:15	14:45	1
		9	18:05	-	1
942	80	1	-	20:00	1
		2	22:30	23:30	1
		3	02:00	-	2

B.2 Block definition and train make-up plan

Block number	Origin	Destination	Trains carrying block
1	1	2	942
2	2	1	908, 927, 942
3	1	7	908
4	1	9	908, 927
5	1	10	908
6	1	11	908
7	1	12	908, 927
8	2	1	249
9	2	3	942
10	3	1	249, 729, 809
11	3	2	249
12	3	4	261
13	3	5	261
14	3	6	227
15	3	7	227
16	3	9	227, 908, 927
17	3	10	908
18	3	11	908
19	3	12	908, 927
20	4	3	162
21	5	3	162
22	5	4	162
23	6	3	722
24	6	7	227
25	6	9	227
26	7	1	729
27	7	3	722, 729
28	7	6	722
29	7	8	703
30	7	9	227, 908, 927
31	7	10	908
32	8	7	307
33	8	9	307
34	9	1	729, 809
35	9	3	722, 729, 809
36	9	6	722
37	9	6	722
38	9	7	703, 722
39	9	8	703
40	9	10	718, 908
41	9	11	718, 908

42	9	12	747, 817
43	10	1	809
44	10	3	809
45	10	7	809
46	10	9	809, 817
47	10	11	809, 817
48	10	12	817
49	11	1	809
50	11	3	809
51	11	9	817, 809
52	11	10	718, 908
53	11	12	817
54	12	1	718, 757
55	12	3	718, 757
56	12	7	757
57	12	9	718, 757
58	12	10	718
59	12	11	718

B.3 O-D demand by departure time during one week period

Origin	Dest	Priority	Departure time (hr)	Day (cars)						
				1	2	3	4	5	6	7
1	2	1	2.0	6	6	8	10	8	6	6
			9.5	6	6	8	10	8	6	6
			20.0	6	6	9	10	9	6	6
1	3	1	2.0	6	6	8	10	8	6	6
			9.5	6	6	8	10	8	6	6
			20.0	6	6	9	10	9	6	6
1	9	1	2.0	6	6	8	10	8	6	6
			9.5	6	6	8	10	8	6	6
			20.0	6	6	9	10	9	6	6
2	3	2	11.2	16	16	24	28	24	16	16
			23.5	16	16	24	28	24	16	16
2	4	2	11.2	4	4	6	7	6	4	4
			23.5	4	4	6	7	6	4	4
2	10	1	11.2	8	8	12	14	12	8	8
			23.5	8	8	12	14	12	8	8
2	12	1	11.2	8	8	12	14	12	8	8
			23.5	8	8	12	14	12	8	8
3	1	1	5.0	3	3	4	4	4	3	3
			8.0	3	3	4	4	4	3	3
			10.1	3	3	4	4	4	3	3
			12.0	3	3	4	4	4	3	3
			12.4	3	3	4	4	4	3	3
			15.6	3	3	4	4	4	3	3
			18.0	3	3	4	4	4	3	3
3	2	2	5.0	3	3	4	4	4	3	3
			8.0	3	3	4	4	4	3	3
			10.1	3	3	4	4	4	3	3
			12.0	3	3	4	4	4	3	3
			12.4	3	3	4	4	4	3	3
			15.6	3	3	4	4	4	3	3
			18.0	3	3	4	4	4	3	3
3	4	2	5.0	2	2	2	2	2	2	2
			8.0	2	2	2	2	2	2	2
			10.1	2	2	2	2	2	2	2

			12.0	2	2	2	2	2	2	2
			12.4	2	2	2	2	2	2	2
			15.6	2	2	2	2	2	2	2
			18.0	2	2	2	2	2	2	2
3	6	2	5.0	5	5	7	8	7	5	5
			8.0	5	5	7	8	7	5	5
			10.1	5	5	7	8	7	5	5
			12.0	5	5	7	8	7	5	5
			12.4	5	5	7	8	7	5	5
			15.6	5	5	7	8	7	5	5
			18.0	5	5	7	8	7	5	5
3	11	2	5.0	3	3	4	4	4	3	3
			8.0	3	3	4	4	4	3	3
			10.1	3	3	4	4	4	3	3
			12.0	3	3	4	4	4	3	3
			12.4	3	3	4	4	4	3	3
			15.6	3	3	4	4	4	3	3
			18.0	3	3	4	4	4	3	3
4	2	2	8.8	16	16	24	28	24	16	16
			20.8	16	16	24	28	24	16	16
4	7	2	8.8	8	8	12	14	12	8	8
			20.8	8	8	12	14	12	8	8
4	11	2	8.8	8	8	12	14	12	8	8
			20.8	8	8	12	14	12	8	8
5	1	2	4.0	16	16	24	28	24	16	16
5	3	1	4.0	16	16	24	28	24	16	16
5	7	2	4.0	16	16	24	28	24	16	16
5	8	2	4.0	16	16	24	28	24	16	16
6	2	2	6.8	8	8	12	14	12	8	8
			12.9	8	8	12	14	12	8	8
6	7	2	6.8	16	16	24	28	24	16	16
			12.9	16	16	24	28	24	16	16
6	9	2	6.8	16	16	24	28	24	16	16
			12.9	16	16	24	28	24	16	16
6	12	2	6.8	8	8	12	14	12	8	8
			12.9	8	8	12	14	12	8	8
7	1	1	5.9	4	4	5	6	5	4	4
			6.9	4	4	5	6	5	4	4
			13.0	4	4	5	6	5	4	4
			14.8	4	4	5	6	5	4	4
			24.0	4	4	5	6	5	4	4
7	1	2	5.9	7	7	10	12	10	7	7

			6.9	7	7	10	12	10	7	7
			13.0	7	7	10	12	10	7	7
			14.8	7	7	10	12	10	7	7
			24.0	7	7	10	12	10	7	7
7	6	2	5.9	4	4	5	6	5	4	4
			6.9	4	4	5	6	5	4	4
			13.0	4	4	5	6	5	4	4
			14.8	4	4	5	6	5	4	4
			24.0	4	4	5	6	5	4	4
7	10	2	5.9	7	7	10	12	10	7	7
			6.9	7	7	10	12	10	7	7
			13.0	7	7	10	12	10	7	7
			14.8	7	7	10	12	10	7	7
			24.0	7	7	10	12	10	7	7
7	11	1	5.9	4	4	5	6	5	4	4
			6.9	4	4	5	6	5	4	4
			13.0	4	4	5	6	5	4	4
			14.8	4	4	5	6	5	4	4
			24.0	4	4	5	6	5	4	4
7	11	2	5.9	4	4	5	6	5	4	4
			6.9	4	4	5	6	5	4	4
			13.0	4	4	5	6	5	4	4
			14.8	4	4	5	6	5	4	4
			24.0	4	4	5	6	5	4	4
8	1	1	7.0	16	16	24	28	24	16	16
8	2	1	7.0	16	16	24	28	24	16	16
8	3	2	7.0	16	16	24	28	24	16	16
8	7	2	7.0	32	32	48	56	48	32	32
8	9	1	7.0	16	16	24	28	24	16	16
8	10	2	7.0	16	16	24	28	24	16	16
8	12	1	7.0	16	16	24	28	24	16	16
9	1	1	1.0	3	3	4	4	4	3	3
			6.5	3	3	4	4	4	3	3
			9.5	3	3	4	4	4	3	3
			10.0	3	3	4	4	4	3	3
			19.0	3	3	4	4	4	3	3
			20.8	3	3	4	4	4	3	3
			21.8	3	3	4	4	4	3	3
9	3	2	1.0	5	5	7	8	7	5	5
			6.5	5	5	7	8	7	5	5
			9.5	5	5	7	8	7	5	5
			10.0	5	5	7	8	7	5	5

			19.0	5	5	7	8	7	5	5
			20.8	5	5	7	8	7	5	5
			21.8	6	5	7	8	7	5	5
9	6	2	1.0	5	5	7	8	7	5	5
			6.5	5	5	7	8	7	5	5
			9.5	5	5	7	8	7	5	5
			10.0	5	5	7	8	7	5	5
			19.0	5	5	7	8	7	5	5
			20.8	5	5	7	8	7	5	5
			21.8	6	6	8	9	8	6	6
9	7	2	1.0	3	3	4	4	4	3	3
			6.5	3	3	4	4	4	3	3
			9.5	3	3	4	4	4	3	3
			10.0	3	3	4	4	4	3	3
			19.0	3	3	4	4	4	3	3
			20.8	3	3	4	4	4	3	3
			21.8	3	3	4	4	4	3	3
9	11	2	1.0	3	3	4	4	4	3	3
			6.5	3	3	4	4	4	3	3
			9.5	3	3	4	4	4	3	3
			10.0	3	3	4	4	4	3	3
			19.0	3	3	4	4	4	3	3
			20.8	3	3	4	4	4	3	3
			21.8	3	3	4	4	4	3	3
10	1	2	2.0	16	16	24	28	24	16	16
			17.0	16	16	24	28	24	16	16
10	2	1	2.0	8	8	12	14	12	8	8
			17.0	8	8	12	14	12	8	8
10	6	1	2.0	8	8	12	14	12	8	8
			17.0	8	8	12	14	12	8	8
10	7	2	2.0	16	16	24	28	24	16	16
			17.0	16	16	24	28	24	16	16
10	8	2	2.0	8	8	12	14	12	8	8
			17.0	8	8	12	14	12	8	8
10	9	1	2.0	8	8	12	14	12	8	8
			17.0	8	8	12	14	12	8	8
10	11	1	2.0	8	8	12	14	12	8	8
			17.0	8	8	12	14	12	8	8
11	3	2	4.2	4	4	6	7	6	4	4
			12.2	4	4	6	7	6	4	4
			19.2	4	4	6	7	6	4	4
			22.9	4	4	6	7	6	4	4

11	8	1	4.2	4	4	6	7	6	4	4
			12.2	4	4	6	7	6	4	4
			19.2	4	4	6	7	6	4	4
			22.9	4	4	6	7	6	4	4
11	9	2	4.2	8	8	12	14	12	8	8
			12.2	8	8	12	14	12	8	8
			19.2	8	8	12	14	12	8	8
			22.9	8	8	12	14	12	8	8
11	10	1	4.2	4	4	6	7	6	4	4
			12.2	4	4	6	7	6	4	4
			19.2	4	4	6	7	6	4	4
			22.9	4	4	6	7	6	4	4
11	12	2	4.2	4	4	6	7	6	4	4
			12.2	4	4	6	7	6	4	4
			19.2	4	4	6	7	6	4	4
			22.9	4	4	6	7	6	4	4
12	6	2	5.0	16	16	24	28	24	16	16
			17.0	16	16	24	28	24	16	16
12	8	1	5.0	8	8	12	14	12	8	8
			17.0	8	8	12	14	12	8	8
12	9	1	5.0	8	8	12	14	12	8	8
			17.0	8	8	12	14	12	8	8

Appendix C

Output of Dynamic Freight Car Routing and Scheduling Model : Train Capacity Utilization during the Planning Period

Train segment	Train number	Train cap. (cars)	From terminal	To terminal	Min. util. (%)	Max. util. (%)
1	162	120	5	4	53.3	80.0
2			4	3	80.0	100.0
3	227	80	3	6	18.8	100.0
4			6	7	100.0	100.0
5			7	9	30.0	92.5
6	249	80	3	2	7.5	100.0
7			2	1	10.0	47.5
8	261	160	3	4	0.0	97.5
9			4	5	5.0	40.0
10	307	120	8	7	100.0	100.0
11			7	9	53.3	100.0
12	703	120	9	7	20.0	100.0
13			7	8	10.0	75.8
14	718	160	12	9	0.0	46.9
15			9	11	16.3	50.6
16			11	10	7.5	49.4
17	722	80	9	7	66.3	100.0
18			7	6	22.5	100.0
19			6	3	30.0	100.0
20	729	160	9	7	5.0	53.1
21			7	6	38.1	100.0
22			6	3	38.1	100.0
23			3	2	26.3	63.8
24			2	1	26.3	63.8
25	747	160	9	12	0.0	51.9

26	757	160	12	9	17.5	29.4
27	809	200	10	11	65.5	74.5
28			11	9	69.5	100.0
29			9	7	65.0	100.0
30			7	6	53.0	86.5
31			6	3	53.0	86.5
32			3	2	20.5	83.5
33			2	1	20.5	83.5
34	817	160	10	11	20.0	70.0
35			11	9	32.5	100.0
36			9	12	7.5	21.9
37	908	200	1	2	6.0	35.5
38			2	3	6.0	35.5
39			3	6	4.5	65.0
40			6	7	4.5	65.0
41			7	9	17.5	100.0
42			9	11	18.5	96.5
43			11	10	14.5	59.5
44	927	160	1	2	7.5	13.1
45			2	3	7.5	12.5
46			3	6	3.8	63.8
47			6	7	3.8	63.8
48			7	9	13.8	73.1
49	942	80	1	2	30.0	100.0
50			2	3	60.0	100.0

Appendix D

Output of Dynamic Freight Car Routing and Scheduling Model :

Car schedule by O-D, Departure Day and Departure Time

This Appendix describes examples of car schedules that are generated from the model.

Origin : 5
Destination : 3
Priority : high

Departure day : 1
Departure time : 04:00
Volume : 16 cars
Scheduled O-D trip time : 6.8 hours

From	To	Block number	Train number	Departure day	Departure time	Arrival day	Arrival time
5	3	22	162	1	04:00	1	10:45

This car schedule is the same for all days during the planning period

Origin : 5
Destination : 7
Priority : low

Departure day : 1
Departure time : 04:00
Volume : 16 cars
Scheduled O-D trip time : 34.2 hours

From	To	Block number	Train number	Departure day	Departure time	Arrival day	Arrival time
5	3	22	162	1	04:00	1	10:45
3	7	15	227	2	10:05	2	14:10

Departure day : 2
 Departure time : 04:00
 Volume : 3 cars
 Scheduled O-D trip time : 34.2 hours

From	To	Block number	Train number	Departure day	Departure time	Arrival day	Arrival time
5	3	22	162	2	04:00	2	10:45
3	7	15	227	3	10:05	3	14:10

Departure day : 2
 Departure time : 04:00
 Volume : 13 cars
 Scheduled O-D trip time : 67.2 hours

From	To	Block number	Train number	Departure day	Departure time	Arrival day	Arrival time
5	3	22	162	2	04:00	2	10:45
3	9	16	927	3	12:25	3	18:05
9	7	38	703	4	19:00	4	23:15

Departure day : 3
 Departure time : 04:00
 Volume : 10 cars
 Scheduled O-D trip time : 91.2 hours

From	To	Block number	Train number	Departure day	Departure time	Arrival day	Arrival time
5	3	22	162	4	04:00	4	10:45
3	9	16	927	5	12:25	5	18:05
9	7	38	703	6	19:00	6	23:15

Departure day : 3
 Departure time : 04:00
 Volume : 14
 Scheduled O-D trip time : 178.2 hours

From	To	Block number	Train number	Departure day	Departure time	Arrival day	Arrival time
5	3	22	162	9	04:00	9	10:45
3	7	15	227	10	10:05	10	14:10

Departure day : 4
 Departure time : 04:00
 Volume : 28 cars
 Scheduled O-D trip time : 154.2 hours

From	To	Block number	Train number	Departure day	Departure time	Arrival day	Arrival time
5	4	23	162	8	04:00	8	08:00
4	3	20	162	9	08:50	9	10:45
3	7	15	227	10	10:05	10	14:10

Departure day : 5
 Departure time : 04:00
 Volume : 24 cars
 Scheduled O-D trip time : 82.2 hours

From	To	Block number	Train number	Departure day	Departure time	Arrival day	Arrival time
5	3	22	162	6	04:00	6	10:45
3	6	14	227	7	10:05	7	12:35
6	7	25	227	8	12:55	8	14:10

Departure day : 6
 Departure time : 04:00
 Volume : 16 cars
 Scheduled O-D trip time : 58.2 hours

From	To	Block number	Train number	Departure day	Departure time	Arrival day	Arrival time
5	3	22	162	6	04:00	6	10:45
3	6	14	227	7	10:05	7	12:35
6	7	25	227	8	12:55	8	14:10

Departure day : 7
 Departure time : 04:00
 Volume : 16 cars
 Scheduled O-D trip time : 58.2 hours

From	To	Block number	Train number	Departure day	Departure time	Arrival day	Arrival time
5	3	22	162	8	04:00	8	10:45
3	7	15	227	9	10:05	9	14:10