ABSTRACT

Title of Document:	MULTI-CARRIER TRACK CAPACITY ALLOCATION IN FORWARD AND SPOT MARKETS OF FREIGHT TRANSPORT
	April Yi-Wen Kuo, Doctor of Philosophy, 2008
Directed By:	Associate Professor, Elise Miller-Hooks, Department of Civil and Environmental Engineering

This dissertation addresses multi-carrier, multi-line train scheduling problems for forward and spot markets. Schedules resulting from solution of these train scheduling problems enable carriers to make customer commitments for serving forward contracts and to transport one-off-loads arising on spot markets.

A train slot selection model based on multicommodity network flow concepts is developed for determining timetables for use in forward markets and a column generation methodology is proposed for its solution. The model considers needs of both shippers and carriers. The solution approach is embedded in a simulation-based iterative framework, where demand elasticity is explicitly treated.

A combinatorial auction-based track capacity allocation framework is introduced to allocate residual track capacity among competing carriers seeking to run additional trains on an as-needed basis. Bid set construction techniques are proposed that allow carriers to express their preferences for track usage. A winner determination problem is formulated, solution of which provides the optimal allocation of track capacity among carriers.

The potential benefits of collaborative operation among carriers in both markets were recognized. Collaborative decision-making (CDM) strategies are designed for scheduling trains to serve forward markets. Performances of these strategies are assessed in a carrier collaboration simulation-assignment framework. A train space leasing technique and a CA-based train slot creation approach are proposed to accommodate one-off-loads on previously scheduled and newly formed trains, respectively. Required techniques for bid set construction are developed. A WDP is formulated that seeks the optimal allocation of track access rights to bidders for the given bid sets.

Implementation of the resulting train schedules will aid in creating efficient and cost-effective rail transport, resulting in a competitive and green alternative to truck transportation. Additionally, collaboration among competing carriers can lead to the formation of profitable trains that might otherwise have been underutilized and an ability of each carrier to serve a greater share of the freight market. The methodologies were specifically intended for solving large, real-world train scheduling problems.

MULTI-CARRIER TRACK CAPACITY ALLOCATION IN FORWARD AND SPOT

MARKETS OF FREIGHT TRANSPORT

By

April Yi-Wen Kuo

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2008

Advisory Committee:
Professor Elise Miller-Hooks, Chair
Michael Ball
Peter Cramton
Hani Mahmassani
S. Raghavan

Acknowledgements

I would like to express my deepest gratitude to my advisor, Dr. Elise Miller-Hooks, for her guidance, support, and encouragement throughout my graduate study at the University of Maryland. Her exemplary work ethic and attitude and her commitment to quality and excellence have always been sources of inspiration for me to become the person who I am. She is not only a mentor on my research and professional development over these years, but also a dear friend and family to me. There is no doubt that her impacts will continue to benefit me in my future career and life.

I am indebted to my committee members, Dr. Michael Ball, Dr. Peter Cramton, Dr. Hani Mahmassani, and Dr. S. Raghavan for their participation and help. Their insightful comments and constructive criticisms have been valuable inputs to my research and dissertation. My appreciation is also extended to the faculty of the Department of Civil Engineering at the University of Maryland for providing excellent academic environment and numerous learning opportunities.

I want to thank Roger for his friendship and advice on my dissertation. I am grateful to my officemates and friends, Alex, Chang-Han, Frank, Jason, Kaiwei, Lisa, Ning, Rahul, Shiaaulir, Spencer and Wenxin. Special thanks go to my young friends Daniel, Eddie, and Mark for the happy moments they brought to me.

Finally, my sincerest gratitude goes to my parents for their generous love. I dedicate this dissertation to them. I would also want to thank my brother, Maggie, Pengcheng for support and encouragement.

Table of Contents

List of Tables

List of Figures

Chapter 1   Introduction

1.1 Motivation and Research Objectives

This dissertation addresses multi-carrier train scheduling problems in rail-based intermodal (IM) freight networks. Schedules that result from solution of these train scheduling problems enable carriers to make customer commitments for serving forward contracts and to transport shipments arising on spot markets. Despite increased desire in many societies for green transport alternatives, over the past two decades, shipper preference for road-based transport has steadily increased, thereby increasing environmental pollution and congestion (EUROSTAT, 2007). Rail as a mode of transport offers the lowest negative socio-environmental external costs amongst all land-based modes (Nair et al., 2007). Thus, the goal of this dissertation is to develop scheduling techniques or strategies that support the efficient and profitable operation of rail-based IM transport services, aiding in the creation of a competitive alternative to truck transport.

The scheduling techniques or strategies adopted for the operation of the rail-based freight transport system are critical to the success of the system. That is, how the system is operated will affect the time required to complete deliveries and cost of operating the system and, in turn, will affect demand for the system. If enough business can be generated such that revenue exceeds costs, the system can be self-sustaining. Revenue can be gained by serving both forward contracts and one-off loads arriving on the spot market. Forward contracts are assumed to require repeated service, and thus, a periodic timetable is developed for this purpose. For shippers seeking transport of shipments with highly irregular frequency (i.e. one-off loads), carriers can schedule trains on a one-time delivery basis within the track capacity that has not been committed for use by trains serving the forward market, i.e.

employing residual track capacity. Train schedules that permit relatively quick delivery of goods at competitive prices, as might be developed through an optimization-based approach, can meet the needs of shippers of both forward and spot markets, satisfy existing customers and produce new demand for the IM system.

Profitability and efficiency of the freight transport system can further be improved through carrier collaboration. Collaborative methods can provide opportunities for multiple carriers to work together for mutual benefit, increasing their flexibility in transporting one-off loads and shipments associated with forward contracts, and achieving on-time performance. Strategies for carrier collaboration can enhance competitiveness of the freight system, resulting in benefits for all carriers.

The problems addressed in this dissertation were motivated by a particular European region, where it is hoped that the European railways will be transformed from nationally fragmented rail-based IM freight systems into an internationally integrated network. Within this EU region, track capacities are not efficiently utilized and train schedules primarily serve national interests. The majority of international freight is transported by truck, resulting in significant roadway congestion and vehicle emissions.

Fig. 1-1 provides a conceptual framework for the train scheduling problems addressed in this dissertation and an outline of the dissertation structure. Scheduling techniques or frameworks proposed herein to address train scheduling problems are developed based on concepts of network optimization, combinatorial auctions (CAs), collaborative decision-making (CDM), and combinations of CA and CDM. These techniques, once implemented, will lead to improved rail service with greater efficiency, transparency and profitability. Furthermore, these methodologies will lead to improvements in the operations of the IM

system, producing a competitive alternative to truck transportation. In turn, this will result in reduced roadway congestion and vehicle emissions. The conceptual frameworks and specific methodologies developed within this dissertation were specifically intended for use in solving large, real-world train scheduling problems.

Figure 1-1 The conceptual framework of this dissertation

1.2 Problem Addressed in this Dissertation

A brief overview of each of the problems addressed in this dissertation, conceptual frameworks and techniques developed for their solution, and associated intellectual and social contributions are provided in this section.

1.2.1 Freight train scheduling with elastic demand

3

system, producing a competitive alternative to truck transportation. In turn, this will result in reduced roadway congestion and vehicle emissions. The conceptual frameworks and specific methodologies developed within this dissertation were specifically intended for use in solving large, real-world train scheduling problems.

Figure 1-1 The conceptual framework of this dissertation

1.2 Problem Addressed in this Dissertation

A brief overview of each of the problems addressed in this dissertation, conceptual frameworks and techniques developed for their solution, and associated intellectual and social contributions are provided in this section.

1.2.1 Freight train scheduling with elastic demand

3

The assignment of freight trains to time slots in a timetable over the railway infrastructure is viewed as a track capacity allocation problem, where carriers wishing to deliver shipments by train compete for track capacity. While the timetable may be set to meet expected demand, this demand depends on the services (supply) that are offered as a function of the timetable. A scheduling tool that explicitly considers the simultaneity between setting the schedule, which determines service levels experienced by shippers, and the associated demand, i.e. the elasticity in demand, is required. Few works in the literature address problems with multi-line operations. No existing work considers the objectives of both the shipper and carrier or demand elasticity.

A set of techniques, consisting of the track capacity allocation technique, track capacity modification technique and train slot selection model, are presented in Chapter 4 that together create a timetable for the multiple decision-maker freight train scheduling problem over multiple lines with given (constant) demand. The timetable generated (or modified) through the use of the track capacity allocation technique (or track capacity modification technique) is employed as input to the train slot selection model. The train slot selection model based on multicommodity network flow concepts is developed for determining freight train timetables used in scheduling international rail services along multiple interconnected routes. The model seeks to minimize operational costs incurred by carriers, and delays incurred by shippers relative to scheduled arrival times, while ensuring that the schedules and demand levels are mutually consistent. A column generation-based methodology is proposed for solution of the model. The set of techniques is embedded in a simulation-based iterative framework (REORIENT Consortium, 2007), where demand for

rail services is re-computed in accordance with the train schedule obtained by solving the freight train scheduling problem.

The proposed scheduling tool was employed over a pan-European rail-based IM network. Results of computational experiments employing the simulation-based framework with embedded optimization-based techniques for obtaining optimal or near-optimal timetables indicate that the introduction of new, integrated and optimizaed rail services can lead to significant increases in IM market share for the rail-based IM services in the region of study.

1.2.2 Collaborative decision-making in train scheduling

The collaborative operation of international rail-based IM freight services by multiple carriers is addressed in Chapter 5. In a low traffic environment, train slots (i.e. a combination of segments of track-time pairs that form a route) may not be scarce resources, and some inefficiency in its allocation may be tolerated. However, if the rail network or portions thereof have high levels of utilization, train slots will be limited.

Three strategies for cooperation, i.e. CDM schemes, are proposed in Chapter 5: (1) train slot cooperation, (2) train slot swapping and (3) train space leasing. In the train slot cooperation approach, two or more carriers can jointly operate a train slot. Thus, through collaboration with other carriers, carriers can transport shipments with origins or destinations (O-Ds) that are not covered by the carrier's own service routes. The train slot swapping approach allows two carriers, each of which owns train slots, to exchange capacity rights for the slots. This can facilitate cooperation when one carrier has excess capacity in a slot and the other has newly arising need for transport along the other carrier's route. The train space leasing approach proposed herein allows the carrier to lease a portion of the train capacity to

other carriers. Through such an approach, the carrier that owns the slot can increase its revenue by selling the residual train capacity to other carriers. While numerous works in the literature consider collaboration in other contexts, no work has applied these concepts on track capacity allocation.

The benefits of the proposed strategies were assessed using a carrier collaboration simulation-assignment framework on a real-world, international network. Results of numerical experiments indicate that the proposed strategies will lead to significant increases in the number of shipments attracted to the proposed services. In addition to attracting more demand, cost savings in terms of rolling stock and labor and reduced shipment delays can be achieved. Such CDM strategies result in a win-win situation for all parties.

1.2.3 Track capacity allocation among multiple carriers competing in spot markets

In Chapter 6, the problem of allocating residual track capacity among multiple competing carriers for use in accommodating shipments arising on the spot market is addressed. To accommodate shipments upon their arrival on the spot market, new trains are scheduled given track capacity usage of trains serving forward contracts. In rail networks, where excess track capacity is abundant, inefficiencies in constructing slots for new train lines as required can be tolerated. Where such excess track capacity is scarce, as may be the case in highly congested systems, efficiency and flexibility in residual capacity utilization is critical to achieving desired service levels and capturing and serving demand in this market.

A CA-based track capacity allocation framework, consisting of bid construction and bid allocation (i.e. the selection of bids to which track access rights will be awarded) procedures, is proposed. Two bid set construction approaches are proposed that allow carriers to fully express their preferences for the track capacity. The first is an all-or-nothing bid set

construction approach, which leads to an XOR bid. The second is a train slot-based bid set construction approach that leads to an OR-of-XOR bid. This technique relies on proposed carrier and revised carrier models that permit quick generation of a manageable number of competitive bid combinations. Bids are evaluated through a winner determination problem (WDP) formulated as a multidimensional knapsack problem for this purpose. Solution of the WDP provides the optimal winning combination of segments of track-time pairs for the given set of bids.

Only two existing auction-based approaches have been proposed in the literature for track capacity allocation. Both use single-unit auction techniques, preventing the carriers from expressing their preferences for desired combinations of train slots. These works ignored path interdependencies and, thus, findings from these works do not extend to the more complex rail networks considered herein.

Results of experiments employing the proposed CA-based framework and both bid set construction approaches show that the train slot-based bid set construction approach performs best in terms of number of atomic bids in the bid sets, number of shipments that are accommodated and average shipment delay.

1.2.4 Collaboration among multiple carriers in the transport of goods arriving on spot markets

Chapter 7 exploits concepts of CDM and CAs employed within Chapters 5 and 6 for scheduling trains to accommodate shipments arriving on forward and spot markets, respectively. Two problems are addressed in this chapter. The first involves the transport of one-off loads on scheduled trains designed to accommodate forward contracts. For this purpose, a train space leasing approach is proposed that permits carriers to transport their shipments on existing trains operated by competing carriers of an alliance. The second

problem considers the transport of one-off loads on new train lines, where carriers are permitted to transport shipments of competing carriers in the alliance. A CA-based train slot creation approach is developed for allocating track capacity among the carriers that seek to operate additional trains to ship the one-off loads that cannot be accommodated through train space leasing. Required techniques for bid set construction are proposed. A WDP is formulated that seeks the optimal allocation of track access rights to bidders for the given bid sets.

No prior work has simultaneously applied both CA and CDM concepts on track capacity allocation. Both train space leasing and CA-based train slot creation approaches aid in forming profitable trains that might otherwise have been underutilized. Moreover, carriers will be able to serve a greater share of the spot market. Such approaches to utilizing limited residual track capacity is essential for providing desired service levels at an acceptable cost. Experimental results indicate that employing the proposed alliance-based framework, as compare with a comparable non-cooperative CA, results in a greater number of shipments that can be accommodated by awarded train slots with fewer trains, leading to higher train capacity utilization and more efficient use of limited public and private resources.

Chapter 2   An International, Rail Freight Network

This dissertation work was motivated by interoperability legislation imposed by the European Commission (EC) on all member states that was designed to aid in transforming the European railways from currently existing nationally fragmented systems into an internationally integrated IM freight transport system. Numerous obstacles exist to the creation of an internationally integrated IM freight transport system in this region depicted in Fig. 2-1. Consequently, nearly all international freight traffic is delivered using trucks, creating significant externalities, such as environmental pollution and congestion. The harmonization of these fragmented systems, however, is crucial to creating a competitive and profitable international IM system. Whether or not this system will be able to provide competitive service as compared with other modes, e.g. transport by truck, depends on its efficiency as measured by both carriers and shippers.

All approaches were designed to be capable of efficiently allocating the track capacity. Techniques proposed in Chapter 4 and 5 are applied on a European network, the REORIENT network, which spans 11 countries from the Nordic Region of Europe through Poland to Romania and Greece.

It is anticipated that the establishment of a multinational IM freight network will contribute to opening new markets and increasing trade along the supply corridor in northern, central and eastern Europe, thereby fostering economic growth in the region.

Figure 2-1 The REORIENT study area

Four southbound service design options described in Table 2-1, developed in consultation with rail operators from the region and founded on market-based research, have been proposed for the REORIENT network depicted in Fig. 2-2. The REORIENT network is designed as a connected hub and spoke network of rail, road and sea in which trains, trucks and ferries are operated. Such a design can support efficient and cost-effective transport, thus, promoting the European railways and inland waterways as a competitive freight transport alternative. The northern portion of the corridor consists of sea-land connections from Sweden, Finland and Norway to an IM hub in Poland. From there, the network is connected via regularly scheduled trains to IM terminals in cities, including Vienna and Budapest. The southern portion of the corridor will involve the utilization of rail and other existing land connections to destinations in central and southeastern Europe. Thus, the network involves

Sweden, Finland, Norway, Poland, Slovakia, Czech Republic, Austria, Hungary, Italy, Bulgaria, Romania, and Greece.

Table 2-1 Route details for newly proposed services

Route ID	Route design	Types of flows carried	Distance (km)	Travel time (hr)
T1	Halsberg-Trelleborg-Swinoujscie-Vienna/Bratislava-Budapest	Bulk	1,182	13.5
T2	Trelleborg-Swinoujscie-Bratislava/Vienna	Unitized	934	11
T3	Gdansk/Gdynia-Bratislava/Vienna-Budapest-Beograd-Thessalonica	Unitized	2,203	27
T4	Bratislava-Budapest-Bucharest-Constantia	Bulk and unitized	1,268	15

Figure 2-2 Four expert-generated service routes

Chapter 3   Comprehensive Review of the Literature

The timetables of the various modes of an intermodal (IM) freight transport system, e.g. rail and ferry, will affect the time required to complete deliveries and the cost of operating the IM system. In this dissertation, a primary objective is to determine the optimal or near-optimal timetables for trains operating on the rail within a freight transport system. Collaborative and non-collaborative techniques, including game-theoretic and network optimization approaches, for determination of these schedules, and for creation of additional train lines to handle one-off-loads, will be developed. In this chapter, the state-of-the-art in IM freight transport scheduling is reviewed. Portions of this chapter may be repeated in other sections of the dissertation where necessary.

In section 3.1, basic background on freight transport in the railway industry is provided and the space-time network often used to represent train scheduling problems is described. The majority of works in the literature that address the scheduling of rail or IM systems for freight transport employ one of two approaches to modeling the problem: game-theoretic modeling techniques or mixed integer programming techniques. These approaches are described in section 3.2. In section 3.3, the contributions of this dissertation in light of the existing state-of-the-art in IM freight scheduling and related areas are delineated. Before proceeding with the review of the state-of-the-art, preliminary background on IM transport and related terminology is provided.

IM transportation can be defined as the use of at least two different modes, such as rail, road, air or sea, to transport goods loaded in a transport unit (e.g. container, trailer, railcar or similar) from an origin to a destination without either unpacking and repacking or

unloading and reloading of goods when changing modes (Arnold et al., 2004; Macharis and Bontekoning, 2004; Corry and Kozan, 2006). The main transport unit considered in the literature is the container. Trailers are equivalent to containers with the exception that the trailer has road wheels. IM shipments are sometimes referred to as COFC (container on flatcar) or TOFC (trailer on flatcar). Railcars are often used for bulk transport, where such trailers or containers cannot be used. They can also be used over multiple modes, such as from rail to ferry; however, they are not as versatile as containers or trailers in that they cannot be transported by truck. For simplicity, and because the techniques reviewed do not rely on a specific definition of the transport unit, it will be assumed that the transport unit is a container throughout this dissertation.

Transfer of containers between modes must be done in a manner that does not decrease the competitiveness of the IM system. Such transfers are completed in IM terminals. A simple example of road–rail IM freight transport from origin to destination is depicted in Fig. 3-1. A container is picked up from a shipper at an origin and is shipped by truck to an IM terminal. After arriving at this IM terminal, the container is placed on a railcar (i.e. flatcar). Railcars are then blocked into trains at a classification yard and delivered by rail carrier to another IM terminal. At the terminal, the container is removed from the railcar and is delivered to its final destination by truck carrier to the receiver.
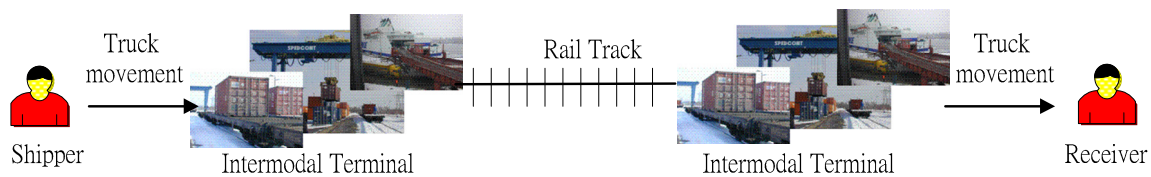
Figure 3-1 A simple representation of road–rail IM freight transport

3.1 Freight Transport in the Railway Industry

In this section, a brief description of freight delivery processes and the terminology employed within the railway industry is given. The space-time representation technique that is commonly used to describe the freight transport movement is also explained.

3.1.1 Freight delivery process in the railway industry

Before goods are shipped, they must be loaded into a container. A container is then loaded onto a railcar. Each railcar has its own origin and destination. Railcars are grouped into blocks and blocks are joined into trains. During the classification process, railcars with different final destinations may be assembled into the same block and railcars in this block will not be disassembled or reassembled until reaching the block's destination. After blocks arrive at their destinations, new blocks are generated by reassembling the railcars within the blocks that have not yet reached their destinations within a classification yard. The blocking policy describes which block should be constructed in each classification yard and which railcars comprise each block. Once blocks are constructed, they must wait in the classification yard for an outbound train. The arrangement of blocks into trains is referred to as the train make-up plan.

The freight shipping process involves many operations and decisions, including classification yard operations, locomotive assignment, train scheduling, empty railcar redistribution, and freight network design (e.g. yard location decisions). These operations and decisions can be classified from strategic, tactical, and operational planning perspectives (Ferreira, 1997; Cordeau et al., 1998). Network design issues considered in this review of the state-of-the-art, for example, are viewed primarily from a strategic perspective, while train

scheduling is primarily a tactical concern. Operational-level decisions are concerned with daily activities. Such models can respond to dynamic changes in the operating environment.

3.1.2 Space-time network representation

The space-time network representation technique has been used in many problems, including railroad scheduling (Nemhauser, 1969; Chih, 1990, Mees, 1991; Huntley et al, 1995; Kwon et al., 1998), empty container allocation (Crainic et al., 1990; Choong et al., 2000), truck fleet management (Powell, 1987). The space-time network representation technique will be used in Chapter 4 to analyze the temporal and spatial train movements over the IM network and to find the optimal routes and schedules for tactical and operational decision-making.

Space-time network representation as it is applied to train scheduling is illustrated in Fig. 3-2. Two separate nodes, IN and OUT, are used to represent processing activities within a terminal. An IN node is used to represent the train arrival at the terminal. An OUT node is used to represent both the train departure from terminal and the time when blocks finish terminal processing, i.e. railcar blocking or train make-up plan. Railcar blocking, train make-up and train dispatching along a network are illustrated in a space-time network for a given period of time in Fig. 3-2. Along the space-time network, the railcar movement, including railcar-to-block assignment, block-to-train assignment, railcar processing, and railcar holding, can be represented by four different types of links: processing links, movement links, holding links and dummy links. Each link is described hereafter.
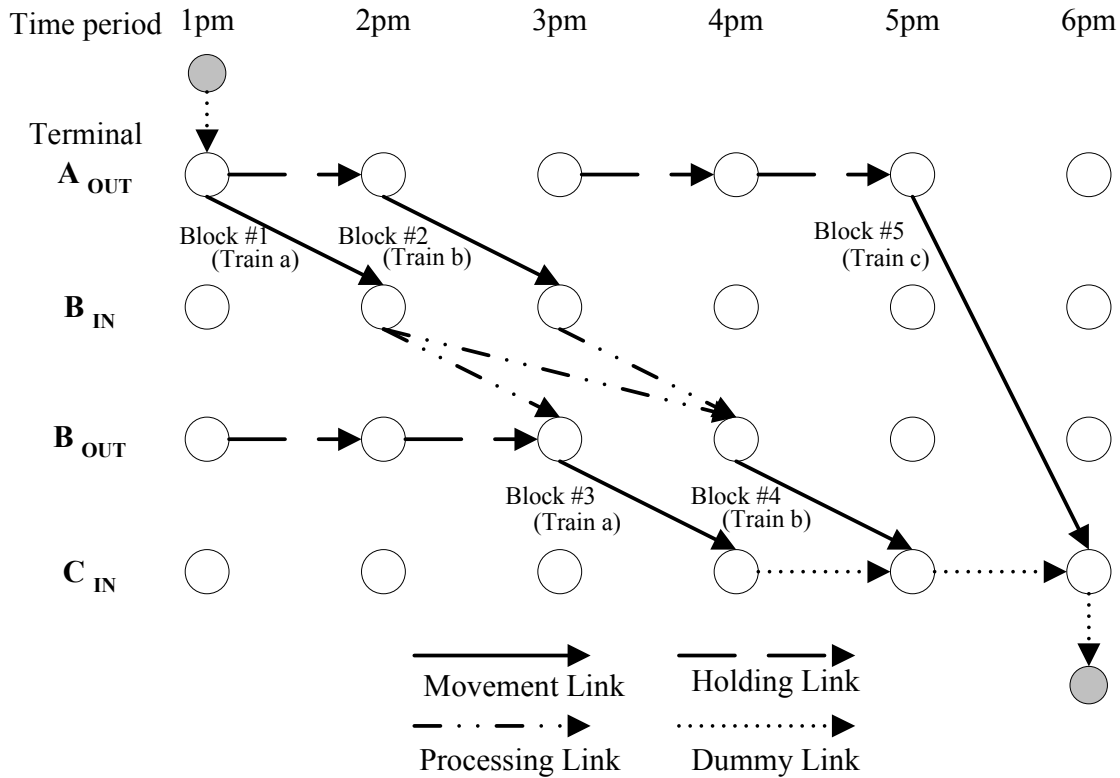
Figure 3-2 Space-time network representation for scheduling

A single shipment, defined in this example, as a number of railcars with the same scheduled departure and arrival times, is shown in Fig. 2-2. In the figure, the shipment's arrival at the network is represented by a sink node and the shipment's arrival at the destination is represented by a dummy link. Hence, dummy links are generated to connect all nodes for actual train arrivals at the destination within the delivery time window. The flow on each dummy link represents the number of railcars arriving within a time window at the destination terminal. The sum of flows on all dummy links to a specific sink node is equal to the volume of the total shipment arriving at that sink node. The cost of a dummy link could represent a penalty cost for late arrival at the destination.

A processing link represents a railcar processing activity at the terminal, such as railcar-to-block and block-to-train assignment, before a railcar is carried by train. In Fig. 3-2,

for example, two processing links at terminal B from 2:00 p.m. to 4:00 p.m. and 2:00 p.m. to 3:00 p.m. are used to show the activity of railcars assembled into blocks #3 and #4 loaded onto trains a and b, respectively, in terminal B.

A movement link represents a block carried by a train with a certain schedule from one train station to another. The flow on this link denotes the number of railcar assigned to blocks that are shipped along this link. It is sometimes necessary to disassemble and assign a block to several trains. Separate movement arcs are thus generated for each part of the original block corresponding to the assigned train. A situation in which a block must be assigned to two trains is illustrated in Fig. 3-2. For example, block #1 at terminal B is split into two blocks: block #3 and block #4. The blocks are assigned to two trains (a and b). In this case, two movement links are generated (one for block #3 on train a and the other for block #4 on train b). Some blocks need not be reclassified at the terminal. They may be carried either by the same train on which they arrived or by a different train to the next terminal. For example, block # 5 uses two train segments of train c (one segment from terminal A to B and the other from terminal B to C) bypassing classification at terminal B. Note that no time is recorded at terminal B.

A holding link represents the railcar held at the terminal until the next available train departure from that terminal. This link may be required when the number of railcars expected to be shipped by a specific train is more than the train's capacity. In such a situation, some of the railcars, consequently, must be held at the terminal until the next available train. The flow on the holding arc denotes the number of railcars held in the terminal that are waiting for the next available train connection. In Fig. 3-2, for example, the number of railcars in terminal A that arrived and finished the necessary terminal processing before departure time of train a

(1:00 pm) is more than the capacity of train a. Thus, some railcars must be held at the terminal to connect to train b (which departs at 2:00 pm) or another train scheduled to depart later.

3.2 Modeling Approaches for Freight Scheduling

A significant problem faced in running rail for tactical and operational planning horizons is how to generate the train timetable over a network. Generating the timetable involves specifying the departure and arrival times of the train at intermediate and final destinations within the network. Train scheduling can be treated as giving trains the right to use railway infrastructure according to a specified timetable. That is, the train scheduling problem can be viewed as a problem of allocating railway track capacity to various trains. Consequently, the train scheduling problem is considered herein as a track capacity allocation problem.

In the freight train scheduling problem, reliability maximization, schedule deviation minimization, and operation cost minimization are typical objectives employed in generating the train timetable (Assad, 1980; Cordeau et al., 1998). In this section, two modeling approaches, game-theoretic and integer programming approaches, used to formulate the freight train scheduling problem are discussed. Basic concepts of game theory, auction theory are provided in section 3.2.1. Works employing auction theory for freight scheduling are also reviewed. Mixed integer programming formulations of this problem are described in section 3.2.2.

3.2.1 Game-theoretic approach with collaborative decision-making

The management of complex spatial dynamic systems, such as an international IM freight transport system, requires the cooperation of various decision agents. Such agents

18

may be organizations and agencies, each of which may be in charge of different portions of interacting systems. Each of these agents may benefit from computer-based decision support systems, including advanced models of underlying processes and algorithms that determine actions to optimize these agents' individual objectives. The resulting outcomes, however, depend on the interaction among these various decisions.

In this section, basic concepts or works with potential to contribute to the conceptual framework or specific algorithmic steps for addressing freight scheduling are reviewed. For more information of auction theory, see Krishna (2002), Fundenberg and Tirole (2002), Parkes (2001), and Cramton et al. (2006). Works employing auction theory for use in track capacity allocation are discussed thereafter. Concepts of collaborative decision-making (CDM) and related works in the literature are also described.

3.2.1.1 Combinatorial auctions

Combinatorial auctions (CAs) are those auctions in which bidders can place bids on combinations of objects, called "packages", rather than only individual objects. CAs have been employed in a variety of industries, including truckload transportation, bus routing, industrial procurement, airport arrival and departure slots arrangement, and in allocating radio spectrum for wireless communications services (Hogan, 1992; Brunekreeft, 2001; Jain, 2001; Melody, 2001; Klemperer, 2002; Milgrom, 2004; Plott et al., 2004; Bourbeau et al., 2005; Elmaghraby , 2005; Strandenes et al., 2005; Triki et al., 2005). The advantage of CAs is that the bidders can more fully express their preferences, leading to improved economic efficiency (allocating the items to those who value them the most) and greater auction revenues. This is particularly important when objects are complements (Xia, et al., 2004; Xia,

19

et al., 2005). Objects are complements when a set of items has greater utility than the sum of the utilities of the individual objects.

In the problem of train scheduling, complementarities may exist, because access rights of track segments exhibit strong interdependencies. That is, the utility of serving one arc depends on the opportunity to serve other specific arcs, where the arcs form a path between an origin and destination. Suppose carriers do not cooperate with each other. The utility of serving a set of arcs together by one carrier may be greater than or equal to the utility of serving them by separate carriers. For example, in Fig. 3-3, a route from A to E and a route from B to F are not complementary, since these two routes have different pairs of origins and destinations. Consequently, it does not make a significant difference if they are contracted to two carriers separately or to a single carrier. However, routes from A to C, C to D, and D to E are complementary to a route between A and E. Combined operations by one carrier will be more cost-effective.
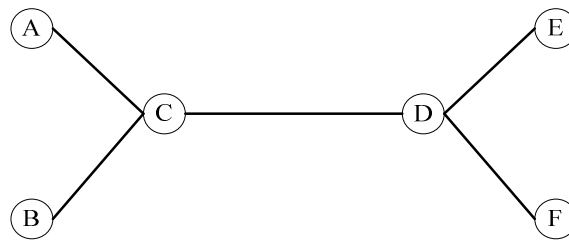
Figure 3-3 A simple network

The most well-known combinatorial auction model is the Vickrey-Clarke-Groves (VCG) auction. In a VCG auction, packages of objects are allocated efficiently to maximize the total value after bidders report their valuations for all packages. Each winner pays the opportunity cost of his/her winnings, i.e. the incremental value that would be derived by assigning the bidder's objects according to his/her next best bid among the other bidders. In

this way, a winning bidder achieves a profit equal to his/her incremental contribution to the total value, and it is a optimal strategy for the bidder to truthfully report the value of the package they are interested in.

In CAs, the problem of identifying which set of bids to accept is called the winner-determination problem. The determination of an optimal winner combination in combinatorial auction is an NP-hard problem (Rothkopf et al., 1995). It has been shown that this problem can be expressed as a mixed integer programming problem (Nisan, 2000) and can be solved by integer programming methods.

3.2.1.2 Existing auction approaches for freight scheduling

In most operations, train schedules are generated by rules and train conflicts are solved through an administrative committee if they occur. Different from most operations, two auction approaches (Brewer and Plott, 1996; Nilsson,1999) for freight scheduling have been found in the literature and are described next.

Brewer and Plott (1996) proposed a game-theoretic approach to address the freight train scheduling problem. A binary conflict ascending price mechanism, a decentralized mechanism based on a first-price auction was proposed. Through numerical experiments, it is shown that this mechanism could solve a track capacity access right allocation problem. In their experiments, ten study participants represented ten bidders and sent the bids for their preferred path through the internet. The bidder who had the highest bid had the right to operate that path. Nine feasible paths were available on a single track line connecting three stations. All of the paths had the same origin and destination with different departure and arrival times. Each path was the subject of seven bidding rounds in which bidders could send

their bidding prices. Three identical experimental sessions with the same bidders were conducted.

Nilsson (1999) developed a second price iterative auction-based procedure to deal with the train scheduling problem. Since in a second price auction it is beneficial for bidders to truthfully report their valuation of the track capacity, a better allocation could be generated compared with using a first-price auction. This approach was implemented by asking rail carriers to submit bids for their primary and secondary preferred paths. The price which is paid by the winner is equivalent to the foregone benefits. There were nine paths and three train operators who were treated as bidders. Ten bidding rounds were held in their experiment. Before the bidding process started, if there were any arrival or departure time conflicts between any pairs of paths, the paths' departure or arrival time was moved backward or forward. Bidders can choose their top two preferred paths from all available paths. Path interdependency over multiple stations was ignored in this work.

The two track capacity allocation mechanisms proposed by Brewer and Plott (1996) and Nilsson (1999) provide a basis for investigating how to design a game-theoretic mechanism to handle the track capacity allocation problem with multiple O-D pairs, multiple tracks and multiple stations, features of the REORIENT network. The network configuration in these prior works contains only two stations and one track and therefore, has limited connection to real-world rail networks. Few conflicts exist in such simple networks. Thus, findings from these works may not extent to more realistic networks. Moreover, consideration of other types of conflicts that may arise when paths over multiple stations are addressed, for example, might limit the applicability of the proposed technique.

Since the network size used by these two works was very small, a single path bidding process was implemented in both works. The goal of this dissertation, however, is to address large, real-world freight networks, which are more complicated than the example networks tested in these works. Moreover, the single object bidding process will not necessarily result in efficient allocation of the bidding object and revenues may not be maximized due to path or arc interdependencies and complementary attributes of the objects. In this dissertation, CAs will be applied in allocating track capacity.

3.2.1.3 Collaborative decision-making

CDM involves teamwork through communication, cooperation and coordination among each of the agents in the team (Berzins, et al., 1988). CDM is usually performed through debates and negotiations among a group of people. Conflicts of interest are inevitable and support for achieving consensus and compromise is required. For problems in which agents compete, but where there is an opportunity to cooperate, an improved solution for each agent might be achieved by incorporating CDM. CDM has been applied in many works addressing, for example, air traffic flow management, supply-chain systems, submarine command and control, engineering design projects, and homeland security problems (Ball et al., 2000; Carlson, 2000; Chang, et al., 2001; Sherali et al., 2003; Groothedde et al., 2005). Among those works, the works in air traffic flow management (Ball et al., 2000; Vossena nd Ball, 2005; Vossen and Ball, 2006) are the most relevant to this dissertation work, especially the aircraft arrival and departure slot arrangement is an example of a capacity allocation problem. The goal of aircraft arrival or departure slot arrangement is to minimize delays incurred at congested airports. Through a procedure based on CDM, arrival or departure slots to an appropriate aircraft are assigned to minimize the total delay of

airlines, thus arranging slots more efficiently. Airlines can benefit from cooperating with each other despite that they are inherently competitive.

CDM can be applied to the allocation of track capacity for freight train scheduling, allowing carriers to cooperate to achieve better performance individually. It may be possible for carriers to save on costs related to rolling stock and labor and to create improved individual carrier train schedules. Such techniques are proposed in Chapter 5 and 7.

3.2.2 Mixed integer programming model

Freight train scheduling works that employ techniques of integer programming are discussed in this section. The majority of the rail scheduling literature has focused on the problem of modeling single-line operations (for example, Szpigel, 1973; Assad, 1980; Petersen et al, 1986; Kraft, 1987; Carey and Lockwood, 1995; Brannlund et al., 1998; Higgins et al, 1996; Nou, 1997; Cordeau et al., 1998). Single-line operations may involve single or double tracks between two yards, junctions or other significant points. The network over which the trains are operated is very simple and form a directed, simple and elementary path. Few works address problems with multi-line operations (for example, Petersen and Taylor, 1982; Jovanvic and Harker, 1991; Odijk, 1996; Kwon et al., 1998; Newman and Yano 2000). Two techniques, multicommodity network flow–based techniques and techniques based on mixed integer programming, have been proposed in the literature to address both single-line and multi-line operations. These works are discussed next.

3.2.2.1 Multicommodity network optimization model

The capacitated multicommodity network flow problem (CMCNP) seeks to ship multiple commodities along capacitated arcs of a network between multiple sources and multiple sinks such that the cost of completing shipments is minimum. Commodities can be

defined as different physical goods that, in the CMCNP, are shipped over a common network, sharing arc capacities, or as distinct pairs of origins and destinations (see Ahuja et al., 1993 for additional detail). In the latter case, flows are to arrive at pre-selected destinations from particular locations of origin. The integral CMCNP has application in a vast number of arenas including, for example, train scheduling, airline crew scheduling, resource-allocation, production-inventory, transportation, and communication network problems (Kennington, 1978; Vemuganti et al., 1989; Bazaraa et al., 1990; Ahuja et al., 1993).

Let x_{ij}^k denote the number of flow units of commodity k on arc (i, j) and c_{ij}^k is the unit flow cost of commodity k on arc (i, j). b^k is the demand for each commodity. u_{ij} is the capacity of arc (i, j). b^k and u_{ij} are integral-valued. The capacitated multicommodity network flow problem can be formulated as follows:

$$\text{Minimize} \sum_{(i,j) \in A} \sum_{1 \leq k \leq K} c_{ij}^k x_{ij}^k \tag{1}$$

Subject to

$$\sum_{(i,j) \in A} x_{ij}^k = b^k, \qquad \forall k \tag{2}$$

$$\sum_{1 \leq k \leq K} x_{ij}^k \leq u_{ij}, \qquad \forall (i, j) \in A \tag{3}$$

Objective (1) seeks to minimize the total shipping cost. Constraints (2) forces the number of flow units must be delivered for commodity k. Constraints (3), often referred to as bundle constraints, impose a flow bound u_{ij} on the capacity of each arc (i, j). Every unit of flow of each commodity is assumed to use one unit of capacity on each arc.

As will be described in Chapter 4, a CMCNP formulation will be proposed for the train scheduling problem. This formulation will require an integer-valued solution. It is known that the single-commodity network flow problem will have only integer-valued

solutions if the supply, demand and capacity data are integer-valued. This is sometimes referred to as the integrality property. Thus, one can employ linear programming solution techniques to solve the single-commodity network flow problem. In addition, specially designed computational techniques (e.g. the network simplex algorithm) with improved efficiency have been developed to determine the optimal solution for these problems. Likewise, exact algorithms exist that solve the CMCNP with no integrality requirements (see, for example, Ahuja et al., 1993 or Assad, 1980), referred to herein simply as the CMCNP as opposed to the integral CMCNP. These can be categorized as price-directive decomposition, resource directive decomposition, and partitioning techniques that are typically more efficient than the simplex method (Liu, 1997). These techniques require repeated solution of large-scale linear programs. While techniques exist for solving the CMCNP, such methods cannot, in general, be used to solve integral CMCNPs, because fractional solutions are not permitted.

The price-directive, resource-directive and partitioning techniques proposed in the literature for the CMCNP can be adapted to solve integral CMCNPs (see Evans, 1980; Liu, 1997; Barnhart et. al., 2000; Aggarwal et al.,1995; Crainic et al., 2000). Price-directive methods, such as Lagrangian relaxation and column generation, decompose multicommodity network flow problems into single-commodity flow problems by placing tolls or prices on bundle constraints. Among other advantages, column generation has one important advantage over Lagrangian-based algorithms. In column generation, a feasible solution is maintained throughout. However, while typical problems require few iterations before termination, each iteration is likely to be time consuming. This is because, at each iteration, a linear program

containing bundle constraints must be solved to determine updated dual prices. Consequently, column generation has not, in general, proven to be an efficient method.

Instead of decomposing the CMCNP into a set of single commodity problems with bundle constraints, as in price-directive methods, resource-directive methods allocate a portion of the bundle capacity to each of the individual commodities. This method is easy to implement, but does not ensure convergence to an optimal solution unless an ideal step size is used. Unfortunately, the ideal step size can only be determined through trial-and-error.

Partitioning techniques employ an arc-based formulation and exploit the spanning tree structure of the linear programming basis (Ahuja et al., 1993) to speed up the computations. Such arc-based formulations may contain significantly more constraints than other comparable path-based formulations.

The integral CMCNP has been shown to be NP-hard (Brunetta, 2000). Thus, it is not surprising that these conventional LP-based techniques have nonpolynomial worst-case complexity. As the problem size grows, the required computational effort increases exponentially and slow convergence in the tail-end of computer runs is common. While some works have sought to extend these techniques for use in solving the integral CMCNP, such extensions have even worse average computational performance due to the integrality requirements. Thus, heuristic techniques designed to address the integral CMCNP are warranted.

The majority of the proposed heuristics for the integral CMCNP employ concepts of mathematical programming. Some heuristics build on the price-directive method (Barnhart et. al., 2000; Liu, 1997). These heuristics employ concepts of dual feasibility and complementary slackness. They decompose the multicommodity problem into a set of single

commodity problems, called pricing problems. A separation problem for the dual linear program (LP) is solved to identify which commodity should enter the basis next. Several heuristics have been developed that employ Lagrangian relaxation within a branch and bound technique (see Schneur and Orlin, 1998; Holmberg and Yuan, 2000; and Mamer and McBride, 2000). In Lagrangian based heuristics, however, it is more difficult to ensure feasibility at each iteration than in comparable column-generation based heuristics.

Some heuristics build on the resource-directive method (Aggarwal et al.,1995; Shetty and Muthukrishnan, 1990). These heuristics first determine an initial integer-valued solution by dividing the capacity on each arc among the commodities and then decomposing the problem into single commodity problems. The capacity is reallocated one arc at a time using dual information. Also, the performance of such resource-directive heuristics depends heavily on the selection of an appropriate step size. Costa et al. (2002) proposed a heuristic based on concepts of decomposition that combines local search with systematic neighborhood changes in the descent phases to avoid getting caught in a local optimum.

Freight train scheduling could be viewed as multicommodity network flow problem if shipments with the same O-D pair are treated as a commodity. Due to the characteristic of CMCNP structure, some freight scheduling works have been formulated as either the integral CMCNP or the CMCNP with side constraints. Kwon et al. (1998) formulated a compound routing and scheduling problem as an integral CMCNP that considers the objective of minimizing the penalty cost. This approach uses a path flow formulation and has a very simple constraint structure. Two types of constraints are emphasized. The first type of the constraint is the capacity constraint, requiring that the sum of railcar flows on a specific train segment cannot exceed the train capacity. The second type of the constraint requires that the

total railcar flow on all feasible train paths must satisfy the demand of each commodity. Inputs to this model generally include the train capacity, the blocking plan, the train make-up plan, estimated terminal processing time, delivery time windows, and demand by shippers. Output from the model is train timetables and the train capacity utilization. A column generation technique was tested on a hypothetical T-shape network in which there are 12 train stations and 16 trains.

Newman and Yano (2000) addressed a problem of scheduling trains on an IM network to achieve on-time delivery at a minimum cost. This problem was formulated as a multiple-fixed-cost integral CMCNP, which can be seen as an integral CMCNP with side constraints. The solution technique developed in this work employed Lagrangian relaxation and Benders' decomposition. Small networks with three to six O-D pairs were tested.

3.2.2.2 Other mixed integer programming methods for train scheduling

Mixed integer programs have been used to formulate the freight scheduling problem with different objectives and planning horizons. Works in which such formulations have been proposed seek to optimize the train timetable considering a certain goal, such as delay minimization, reliability maximization or operation cost minimization (profit maximization) during the planned schedule while satisfying a set of operational constraints. The following works are discussed with respect to these different objectives.

The competitiveness of a rail system depends on, in part, train schedule reliability. Jovanvic and Harker (1991) defined the reliability as the probability that a given set of timetables will still work feasible under the multi-line operating conditions arising from bad weather conditions and train break-down conditions. A decision-support system (DSS) was proposed for tactical scheduling of freight transport on single and double track for multi-line

operations. The DSS seeks to maximize the schedule reliability within a single day. Constraints include on-time arrival at the destination, train conflict prevention, train overtaking, and operational considerations (e.q. minimum headway enforcement). A branch and bound technique in which the decision variable is the ordering of the trains is suggested to solve this problem within the DSS.

The reliability of train arrivals is critical to the performance of the railway industry. Thus, to enhance customer satisfaction, the objective of many works on freight transportation by rail has been to minimize delay from the planned schedule. Some of these works focus on single line operations (Szpigel, 1973; Petersen et al, 1986; Kraft, 1987; Higgins et al, 1996) and others on multi-line operations (Petersen and Taylor, 1982; Kraay et al., 1991; Karry and Harker, 1995; Hallowell and Harker, 1996).

Whether or not a train arrives close to or on schedule depends on many factors, such as train conflicts, terminal congestion, and train speeds. Given the train departure time and train speed, Szpigel (1973) sought to minimize the train delay by determining the best train overtaking and crossing positions. Kraft (1987) presented a branch-and-bound based technique to determine the train siding location such that delay is minimized. Hallowell and Harker (1996) addressed the performance of railway operations by discrete-event simulation in which each simulation run considers the train delay. Higgins et al. (1996) attempted to minimize total delay by calculating the number and position of the sidings for each train on a single line track. A decomposition based heuristic was developed in which two sub-problems were solved iteratively until no improvement could be made. The positions of the sidings are determined in the first sub-problem and the train schedule is decided in the second sub-problem. Rather than calculating sidings, Kraay et al. (1991) developed a branch-and-bound

based technique to determine the train schedule that minimizes delay given train crossing and meeting positions, as well as train speeds, while satisfying departure and arrival time requirements of each train. Petersen and Taylor (1982) developed a model to simulate train movement over a line to minimize the total delay. The model permits an arbitrary number of trains with different speeds and priorities to be dispatched over any line configuration, including single or multiple tracks and sidings with restricted switching or crossovers.

Similar to minimizing the total train delay, many works seek to minimize train schedule deviation for either the single line (Carey and Lockwood, 1995; Brannlund et al., 1996) or multi-line operation. (Carey, 1994; Karry and Harker, 1995). Minimization of deviation from train schedules differs from minimization of train delay. In the former, it is necessary to consider both late and early arrival; whereas, in the latter only late arrival is considered. Carey (1994) presented a decomposition technique to minimize the train schedule deviation of trains running on multiple lines in each direction by choosing the departure and arrival time of each train at each station. Carey and Lockwood (1995) solved the same problem using a branch-and-bound technique in which the decision variables represent the order of the train. Kraay and Harker (1995) proposed a heuristic with a local search method to solve the freight train scheduling. A heuristic was first proposed to quickly find a feasible solution and then simplified decomposition, developed by von Hohenbalken (1977), was applied to find a near optimal solution.

Profit maximization or minimization of costs has also been considered in the context of freight transport by rail. Li and Tayur (2005) developed a decomposition-based technique for solving a tactical IM freight scheduling problem for multi-line operation so their techniques determine the price of freight transport with the objective of profit maximization.

Several works proposed techniques that seek to minimize the total operating cost (Nozick and Morlok, 1997; Gorman, 1998). The schedule-related operating cost considered in these works includes a fixed charge for each train, variable transportation and handling costs for each container and a yard storage cost. Constraints in these models include: on-time delivery requirements, customer service expectations, link and node congestion, locomotives, and train-size limitation. Given an available schedule, Nozick and Morlok (1997) proposed a heuristic to address the freight movement and locomotive assignment problem for IM rail-truck service. This heuristic seeks to provide more reliable service while minimizing the operational cost. Gorman (1998) addressed the problem of minimizing the schedule-related operating costs for a one-week scheduling horizon. A hybrid tabu search/ genetic algorithm was proposed for solution of this problem. Computational times were greatly improved as compared with exact procedures without significant loss in solution quality.

3.3 Conclusions

In summary, none of the existing freight train scheduling works in the literature considers the objectives of both the shipper and carrier or demand elasticity, characteristics that are explicitly considered herein. Only two existing auction-based approaches have been proposed in the literature for track capacity allocation. Both use single-unit auction techniques, preventing the carriers from expressing their preferences for desired combinations of train slots. Use of single-unit auctions is time-consuming, particularly for large problem instances. Moreover, these works consider ignore path interferences and, thus, findings from these works do not extend to the more complex rail networks considered herein. Combinatorial auctions that take into consideration such interdependencies and need for fully expressing preferences for combinations of bidding items are proposed for this purpose in

this dissertation. Numerous works in the literature consider collaboration in other contexts; however, no prior work considers opportunities for collaboration among multiple carriers as it relates to track capacity allocation. In this dissertation, techniques that capitalize on the potential benefits of collaboration are proposed.

Chapter 4   Freight Train Scheduling with Elastic Demand

4.1 Introduction

In this chapter, a train scheduling optimization tool is proposed for forward markets that enables more efficient utilization of track capacity, improved delivery times and reduced operating costs over a rail-based intermodal (IM) freight transport network. Improved rail-based IM services will lead to increased competitiveness and, thus, greater market share, for this freight transport alternative, resulting in reduced roadway congestion and reduced vehicle emissions.

The proposed optimization tool is comprised of a train slot selection model and iterative simulation-based framework (REORIENT Consortium, 2007). The train slot selection model is a multicommodity network flow model and a column generation-based technique is proposed for its solution. The model and solution technique seek an optimal periodic (e.g. weekly) rail and ferry timetable based on given demand estimates. Without loss of generality, ferries are treated as trains in the scheduling procedure. Because demand for this IM alternative depends on the timetable and the optimal timetable depends on the demand for the services, the train slot selection model is iteratively solved within a simulation framework. Once a timetable is produced, it is fed to the simulation model and demand for the services is estimated. The new demand estimates are fed back to the train slot selection model and a new timetable is generated. This process is repeated until termination criteria are met or convergence is achieved.

The train slot selection model considers objectives of both the shippers and the carriers. The shipper, on the one hand, seeks a schedule that provides the quickest, most reliable service at the lowest cost and the carrier, on the other hand, seeks a schedule that

minimizes operating cost and captures the greatest market share. The objectives of these decision-makers are, therefore, conflicting in nature. Ideally, the resulting timetable will consider both decision-makers' objectives simultaneously. The objective employed in the model, thus, seeks to minimize a function of both operational costs (composed of both infrastructure charges and track access costs) and delays in delivery from scheduled arrival times. The frequency of service is set to capture as much of the demand as possible.

In addition to developing and refining the rail timetables, this optimization tool can be used to evaluate numerous changes to the IM system. For example, it can be used to assess the impact of: reductions in delay, changes in rail infrastructure charges, and introduction of improved technologies on shipment times, market share and expected revenue. It can also permit identification of system bottlenecks.

Previous works on freight train scheduling are discussed in section 4.2. Nomenclature and a space-time network representation used in this study to create the rail timetables are introduced in section 4.3. The process of creating a timetable for a fixed demand is first introduced (section 4.4). This process begins with an initial timetable. The procedure proposed for developing the initial timetable, an *initial track capacity allocation technique*, is described in section 4.4.1. The initial timetable is provided as input to a *track capacity modification technique* proposed in section 4.4.2. This technique makes adjustments to the timetable in an effort to reduce incurred delays at terminals and resolve conflicts that might arise between trains. The output of this technique, a modified timetable, is employed by the *train slot selection model*, presented in section 4.4.3, produces a reduced timetable that meets frequency requirements.

In section 4.5 an iterative simulation-based framework is described that, employing techniques from section 4.4, produces a timetable for elastic demand, where the demand for the services depends on the services offered and services offered depend on the demand. An initial timetable, produced either by the *initial track capacity allocation technique* or improved through the additional *track capacity modification technique* and *train slot selection model* processes described in section 4.4, is implemented in a freight train shipment-assignment microscopic simulation platform. The simulation employs a mode choice model to estimate the number of zone-to-zone shipments that are attracted to the rail-based IM services, which are offered based on the given timetable. The simulation then assigns these shipments to the train slots of the timetable, tracks their progress through the rail-based IM system, and calculates resulting *shipment delays*. These shipment delays are then fed back to the *track capacity modification technique* and *train slot selection model* processes to produce a new timetable that better meets the new estimates of demand. The resulting updated timetable is fed back to the simulation model and the process is repeated. This iterative simulation-based process for developing a timetable for elastic demand is presented in section 4.5. This process was applied on a real-world, pan-European network spanning 11 countries, results from which are discussed in section 4.6. Conclusions are given in section 4.7.

4.2 Previous Works on Freight Train Scheduling

The majority of the rail scheduling literature has focused on the problem of modeling single-line operations(for example, Szpigel, 1973; Assad, 1980; Petersen et al, 1986; Kraft, 1987; Carey and Lockwood, 1995; Brannlund et al., 1998; Higgins et al, 1996; Nou, 1997). Single-line operations may involve single or double tracks between two yards, junctions or

other significant points. The network over which the trains are operated is very simple. Few works address problems with multiple-line operations (for example, Petersen and Taylor, 1982; Crainic, 1984; Jovanvic and Harker, 1991; Odijk, 1996; Kwon et al., 1998; Newman and Yano, 2000). Two general techniques based on mixed integer programming or multicommodity network flow programming have been proposed in the literature to address both single-line and multi-line operations.

Mixed integer programs have been used to formulate the freight scheduling problem with different objectives and planning horizons. Various integer programming based techniques, such as branch and bound techniques (Jovanvic and Harker, 1991; Kraay et al., 1991; Kraft, 1987; Carey and Lockwood, 1995; Higgins et al., 1996), decomposition based heuristics (Carey, 1994; Li and Tayur, 2005; Nozick and Morlok, 1997; Kraay and Harker, 1995), and meta-heuristics (Gorman,1998), were proposed to solve the scheduling problem.

Freight train scheduling could be viewed as an integral multicommodity network flow problem if shipments with the same O-D pair are treated as a commodity. This approach is taken in several works. Kwon et al. (1998) proposed a combined routing and scheduling model to minimize the total delay. They tested their solution approach for their model in a T-shape network. Newman and Yano (2000) proposed an integral multicommodity network flow model with side constraints in which on-time delivery is required to minimize the operating costs. Small networks with three to six O-D pairs were tested. An extensive survey of optimization tools developed for train scheduling is provided by Cordeau et al. (1998).

Surveys by Assad (1980) and Haghani (1987) indicated that analytical optimization models for rail transportation were not widely employed in practice. Even more recent works (Kwon, et al., 1998; Gorman, 1998; Newman and Yano, 2000; Ghoseiri, et al., 2004)

consider small problem instances that do not reflect the characteristics of real-life applications. In addition, a number of works describe simulation techniques that aid in resolving train conflicts (Cheng, 1998) and assessing train delays (Petersen and Taylor, 1982; Dorfman and Medanic, 2004).

Freight train scheduling is inherently multi-objective due to the conflicting interests of the various stakeholders. However, most analytical optimization scheduling models consider only a single objective. These works take the perspective of either the carrier or the shipper. From the carrier's perspective, the objective might be to minimize deviation from the schedule, (Carey and Lockwood, 1995; Carey, 1994a; Carey, 1994b; Kraay and Harker, 1995; Brannlund et al., 1995; Nou, 1997) or operational cost (Li and Tayur, 2005; Nozick and Morlok, 1997; Gorman, 1998). From the perspective of the shipper, the objective might be to minimize shipment arrival times or differences from desired times (Kraay et al., 1991; Higgins et al., 1996). It appears that only one study in the literature explicitly models the multi-objective character of the scheduling problem (Higgins et al., 1996), where the objective is a function of both deviation from scheduled arrival time and fuel consumption costs. Higher priority is given to the cost of delay than to fuel consumption. These measures are of interest to the carrier only.

None of the existing works in the literature considers the objectives of both the shipper and the carrier. Nor does any of these works consider elasticity in demand in multi-line operations. In this dissertation, the multi-line and multi-objective freight train scheduling problem with elastic demand is addressed. To address elasticity of demand, the solution approach is embedded in an iterative simulation-based framework, where the demand is re-estimated in light of the train schedule (and resulting estimates of level-of-service) that

results from solution of this freight train scheduling problem. The scheduling problem is re-solved in response to changes in demand estimates and the procedure repeats until termination criteria are met.

4.3 Problem Representation

The representation of freight moved by rail and associated modes from origin to destination requires a model of the underlying network structure and processes. On the supply side, physical infrastructure of roads and rail links serve as a network on which services operate. This network connects zones where freight originates and terminates (representing the demand side). The *train slots* (i.e. the use of a route from shipment origin to shipment destination during a given period of time) constructed from residual track capacity. A space-time network representation over a period of time T, $G_T = (N_T, A_T)$, is exploited by the *train slot selection model* and the associated solution technique. The nodes (except for the pseudo source defined as α and pseudo sink defined as β) and arcs in the space-time network G_T have both a space and time component. Let the node set $N_T = \{U_T, V_T, \alpha, \beta\}$, where node $v^t \in V_T$ denotes node v at time $t \in T$ and $u^t \in U_T$ denotes node u at time $t \in T$. $(u^q, v^r) \in A_T$ is an arc representing a potential space-time pair for which a train can depart from terminal u at time q and arrive at terminal v at time r.

A train's itinerary along the route and its activities required at the terminal can be represented by four distinct arc types: (1) departure arcs, (α, u^q), each corresponding to a feasible departure for a train at its origin; (2) movement arcs, (u^q, v^r), representing a train denoted as i traveling from terminal u at time q to another terminal v at time r; (3) process/siding arcs, (v^r, u^q), representing train processes or siding decisions at terminals;

and (4) arrival arcs, (v^r, β), representing train arrival at its destination terminal. The time spent on process/siding arcs is at least the minimum time required for railcar processing activities, border crossing or train holding at the associated terminal if relevant. Border crossing operations are modeled as terminals. Note that infinite capacity of stations is assumed. A *train slot* consists of a departure arc, several process/siding and movement arcs, and an arrival arc along the space-time network. The space-time representation of three train itineraries are depicted in Fig. 4-1.
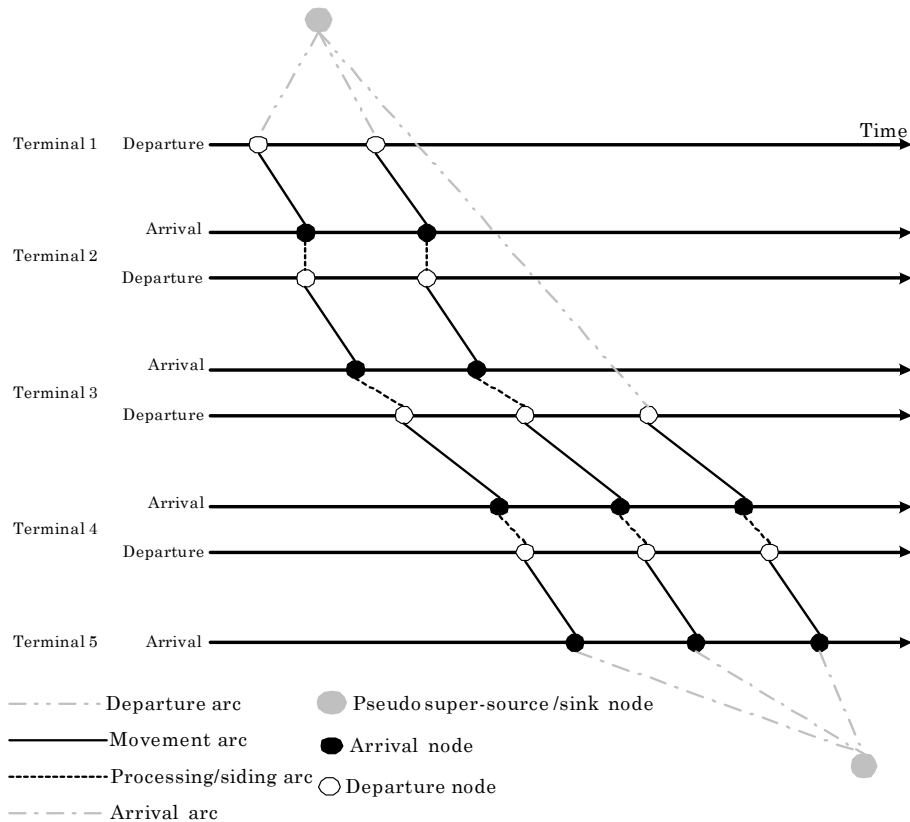
Figure 4-1 Space-time network representation of train itineraries

A train i's charge, c_i, is associated with each arc in the space-time network. This charge consists of operational costs associated with running a train. Such costs include locomotive, c_{lo}, track access, c_{ac}, and infrastructure c_{fr} charges. The locomotive charges

are applied to the departure arcs. The track access charges are applied to the movement arcs and the infrastructure charges are applied to the process/siding arcs. Note that there is no charge associated with arrival arcs.

Train slots must be constructed such that no train conflicts on the track exist. Fig. 4-2 shows possible conflicts that must be avoided on the track while scheduling the trains along the space-time network. The train cannot depart without keeping a minimum headway with previous scheduled train on the same track. Thus, while creating the train slot, only one arc on the space-time network shown on the Fig. 4-2(a) (or (b)) can be selected, since two consecutive trains must maintain a minimum headway. A track can only be accessed by one train during a given time interval to prevent two trains from meeting on the same track. Thus, the trains are not allowed to meet on the track as shown in Fig. 4-2(c).
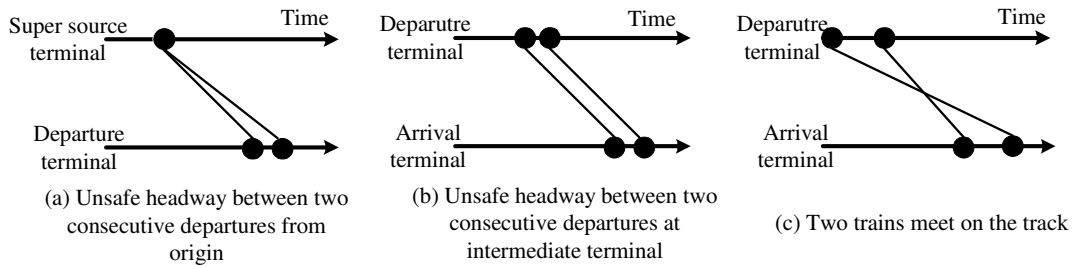
Figure 4-2 Possible train conflicts on the space-time network

To model *train delay* (i.e. actual arrival time minus preferred arrival time), it is necessary to know the arrival time at the destination terminal for each train i. The arrival time can be obtained from the arrival arc. Train i's arrival time is equal to r, since r is the time when the train departs from v for the pseudo-sink node β. If the train arrives later than the train's preferred arrival time, delay, μ_i, for train i is incurred that is set to the difference between the train's actual arrival time and the ideal arrival time.

Shipment delay at terminal (i.e. *shipment delay*) is incurred at the terminals due to waiting and/or processing times built into the current train timetable. The *shipment delay* at each terminal is determined from the shipment's arrival time at the terminal, which is affected mainly by the predetermined distribution of the shipment generation rate and the train/truck arrival time at this terminal, and the timetable. This delay is defined as the time period starting when the shipment arrives at the terminal and ending when the shipment is loaded on the train.

The *shipment delay* may result from either the shipment process time at the terminal or the shipment waiting time for the next available train. In this chapter, it is assumed that the cause for the delay is unknown. Thus, it is assumed that half of the delay is caused by the shipment process activity at the terminal. For shipments arriving approximately at the same time and waiting for the same train, these *shipment delays* form a *cluster delay*, which is defined as the sum of the shipment delays of these shipments.

4.4 Creating a Timetable for Inelastic Demand

In this chapter, a set of techniques are presented that together create a timetable for the multiple decision-maker freight train scheduling problem over multiple lines with inelastic demand. The process of developing the timetable begins with the development of a set of candidate train slots that form an initial timetable (the *initial track capacity allocation technique* presented in subsection 4.4.1). The initial timetable is adjusted (by the *track capacity modification technique* described in subsection 4.4.2) such that delays incurred at terminals are minimized and conflicts are resolved. A reduced timetable is constructed from the modified timetable (by the *train slot selection model* presented in subsection 4.4.3) that meets frequency requirements of the given demand.

42

4.4.1 Initial train timetable

The problem of developing an initial train timetable as input for the process of developing an optimal timetable for a given demand is described in this subsection. This problem is formulated as a linear, binary integer program (subsection 4.4.1.1) and a fast heuristic is presented for its solution (subsection 4.4.1.2).

4.4.1.1 Mathematical model formulation

Before proceeding to the presentation of the mathematical formulation of the initial train timetable problem, notation employed in the formulation is given.

Notation

K : set of routes

I^k : set of candidate *train slots* operating on each route $k \in K$

A^i : set of tracks passed by train $i \in I^k$

x_a : binary decision variable that indicates whether or not the arc $a \in A^i$ is used in the train timetable

c_a : operational cost of track $a \in A^i$

η^k : suggested train frequency for each route $k \in K$

$d_i^+(n)$: set of train i's outgoing arcs $a \in A^i$ at node $n^t \in N_T$

$d_i^-(n)$: set of train i's incoming arcs $a \in A^i$ at node $n^t \in N_T$

η : set of conflicts between any pair of trains

C : set of conflicts for a specific pair of trains

Model formulation

The model formulation is given in (1) through (7).

$$\text{Min} \quad z(x) = \sum_{k \in K} \sum_{i \in I^k} \sum_{a \in A^i} c_a x_a \qquad \qquad (1)$$

$$\text{subject to} \quad \sum_{i \in I^k} \sum_{a \in d_i^+(\alpha)} x_a = \eta^k, \qquad \forall k \in K. \qquad (2)$$

$$\sum_{a \in d_i^+(\alpha)} x_a \le 1, \qquad \forall i \in I^k \;\; \forall k \in K. \qquad (3)$$

$$\sum_{a \in d_i^-(n^t)} x_a = \sum_{a \in d_i^+(n^t)} x_a, \qquad \forall i \in I^k \;\; \forall k \in K \;\; \forall n^t \in N_T \setminus \{\alpha, \beta\} \;\; \forall t \in T. \qquad (4)$$

$$\sum_{a \in d_i^-(\beta)} x_a \le 1, \qquad \forall i \in I^k \;\; \forall k \in K. \qquad (5)$$

$$\sum_{a \in C} x_a \le 1, \qquad \forall C \in \eta. \qquad (6)$$

$$x_a \in \{0,1\}, \qquad \forall a \in A^i \;\; \forall i \in I^k. \qquad (7)$$

Binary decision variable x_a represents whether or not the track segment a associated with a certain period of time on G_T is included in the timetable. The objective given in equation (1) seeks to minimize the total operational cost of transporting shipments along the provided train route for the given schedule within the corridor and to minimize the sum of *train delays* incurred along the network. Constraints (2) ensure that the total number of train slots on a route that will be operated is equivalent to the number of suggested train slots that may be operated for transporting the required shipments for the route. Constraints (3) ensure that at most one track associated with a train is selected among those leaving the super-source terminal. Constraints (4) are the mass balance constraints for each terminal. These constraints impose equality on the number of selected arcs associated with a train entering and leaving each arrival or departure terminal. Constraints (5) ensure that at most one track associated with a train is selected among those arriving at the super-sink terminal.

Constraints (6) prevent two consecutive trains from running on the same track at the same time (or within the minimum headway), while simultaneously imposing the track capacity constraints. Binary integrality requirements for every arc are given in constraints (7).

The size of the formulation is a function of the number of the track segments in the space-time network (number of decision variables), the number of routes, the suggested train frequency for each route, the time length of the planning horizon, and the number of the train conflict constraints. One can expect the number of decision variables and constraints in the formulation to be quite large for a real-world problem instance. The number of train slots will increase exponentially with the number of arcs in A_T of G_T. In addition, this formulation, a binary multi-commodity network flow formulation with the side constraints, is an NP-complete problem. Thus, a formidable computing task would be expected if solution to optimality is required.

The purpose of this initial track capacity allocation is to generate an initial feasible train timetable, where there are no conflicts between trains and a minimal frequency along each route is upheld. The initial train timetable is further improved through the *track capacity modification technique* and *train slot selection model* described in succeeding subsections. Thus, it suffices to employ a quick and efficient technique for producing this initial feasible timetable. Such a technique is described next.

4.4.1.2 Solving the initial train timetable problem

An initial timetable is constructed from the residual track capacity (i.e. the entirety of existing track capacity or remaining capacity given existing services with fixed schedules) by assigning as many trains to a train slot as permissible or by assigning a preset number of

trains to train slots along a given route at a suggested frequency. This procedure of assigning trains to train slots is referred to as the *initial track capacity allocation technique*.

Before applying the *initial track capacity allocation technique*, the routes are ranked from longest to shortest according to their average travel times, assuming that delays at intermediate terminals are zero. Note that trains are assumed to travel at a constant average speed. In each iteration of the *initial track capacity allocation technique*, the route with the highest rank is selected (if not previously selected) and a train slot is constructed along that route. Each train slot, by definition, must connect the origin and destination nodes ordered in time. In addition, the train slot designates the order in which intermediate terminals are visited. To determine potential departure times (i.e. a departure arc in the space-time network), the minimum required headway between trains employing the same track and heading toward the same terminal must be considered. The process/siding arc must accommodate the minimum required terminal process time for the shipments. Given the earliest time that the train can leave from the terminal after the train finishes its required activity (e.g. shipments pick-up/drop-off or train classification/reclassification), an earliest feasible track capacity on the movement arc is sought between the departure terminal and the arrival terminal. Once track capacity on every track segment between all consecutive terminals is assigned, the *train slot* is constructed by the *initial track capacity allocation technique* described next.

Initial track capacity allocation technique

Step 1. (Initialization)

> *1.1* Let F^k denote the frequency of the *train slots* constructed on route k. $F^k = 0$,
>
> $\forall k \in K$.

1.2 Let h denote the minimum allowable headway between trains.

Step 2. (Select route)

2.1 Check F^k, $\forall k \in K$. If F^k reaches the suggested frequency, $K = K \setminus \{k\}$. If $K = \varnothing$, terminate. Otherwise, let B denote a set of the routes which do not meet train frequency requirements. $B = K$.

2.2 If $B = \varnothing$, return to *Step 2.1*. Otherwise, select a route $k \in B$ with the longest travel time. $B = B \setminus \{k\}$. Construct a *train slot* along route k for the available track capacity A_T in G_T.

Step 3. (Construct train slot)

3.1 Choose feasible departure arc (α, u^q), where $u \in E_i$ and $\alpha = f$.

Based on the minimum allowable headway h, search the earliest available track time capacity between super-source node α and the origin terminal f in the arc set A_T in G_T. If a feasible train departure arc cannot be identified, $K = K \setminus \{k\}$ and return to *Step 2.2*. Otherwise, continue.

3.2 Choose feasible movement arc (u^q, v^r), where $u \in E_i, v \in E_i, u \neq v$.

Let q denote train i's earliest feasible departure time at terminal u. Starting from time q, search for the earliest available track time capacity from terminal u to terminal v. If a feasible movement arc cannot be identified, return to *Step 2.2*. If the arrival terminal v is a destination terminal g, go to *Step 3.4*. Continue, otherwise.

3.3 Choose feasible process/siding arc (v^r, u^q), where $v \in E_i, u \in E_i, v = u$.

Let r denote the train i's arrival time at terminal v and q denote the time when the train finishes its required activity at the terminal. Return to *Step 3.2*.

3.4 Choose feasible arrival arc (v^r, β), where $v \in E_i$ and $v = g$.

One *train slot* is obtained. Let $F^k = F^k + 1$. Return to *Step 2.2*.

This procedure terminates with a feasible train schedule, meeting suggested train frequencies.

4.4.2 Track capacity modification

The goal of the *track capacity modification technique* is to adjust the train arrival/departure times in an effort to reduce the *shipment delays* and attempt to generate additional train slots based on the residual track capacity. The timetable is then appropriately modified, where trains are assumed to be able to travel within a certain allowable range of speed. The resulting modified timetable is employed as input to the *train slot selection model* described in section 4.4.3.

A neighboring timetable of the current timetable is generated through reducing the *cluster delay* identified at a drop-off/pick-up terminal within a neighborhood due to a train's scheduled departure time (i.e. if the next scheduled departure is far from the shipment's arrival time). The largest *cluster delay* will be selected for reduction by shifting the train's departure time back to the mid-point of the range of activities that contribute to the *cluster delay*. The new departure time from a terminal must be later than the train's arrival time at that terminal. The departure and arrival times for each terminal visited by the train at points succeeding this terminal must be rescheduled by checking the track capacity. To allow the train to be rescheduled successfully (i.e. without conflicts with other train slots), the train is allowed to travel on the track by varying the speed within a reasonable range so that the disturbance caused by rescheduling the trains can be minimized. Additional shipments can be

accommodated at the pick-up/drop-off terminals. Such shipments would otherwise have to wait for the next available train.

After adjustment to the existing train slots is completed, additional train slots will be constructed by randomly creating train slots for every route. Each train slot will be constructed from its origin to its destination while keeping a minimum headway between trains that operate on the same track

For each train $i \in I^k$, let $E_i = \{f, (f+1)...,(g-1), g\}$ denote an ordered set of the terminals that train i will visit in sequence along the route in which f and g is defined as the route's origin and destination, respectively. The *track capacity modification technique* is described next.

Track capacity modification technique

Step 1. (Choose terminal)

Choose a terminal $l \in L$. Let M^l represent the set of trains that will load/unload shipments at terminal l. $L = L \setminus \{l\}$. If $L = \emptyset$, terminate.

Step 2. (Search neighborhood delay)

2.1 Choose one train $i \in M^l$ that departs earliest from terminal l. Remove train i from M^l. If $M^l = \emptyset$, return to *Step 1*.

2.2 Construct a neighborhood of train i's departure time at terminal l. Denote the train's departure time at terminal l as $\beta(l)$. Let $h_{i,j}$ denote the headway between two consecutive trains, i and j. The range of potential departure times is given by

$$\left[\beta(l) - \frac{1}{2} h_{i,i-1}, \beta(l) + \frac{1}{2} h_{i,i+1} \right],$$

where *i-1* is indicated as the train that departs before train *i* and *i+1* is indicated as the train that departs after train *i*.

Within the range of departure times, for shipments arriving approximately at the same time and waiting for the same train, *shipment delays* form a *cluster delay* for the terminal.

2.3 Search for the largest *cluster delay* indicated as *d* from the set of clusters defined as D_i for train *i* in the neighborhood constructed from *Step 2*. $D_i = D_i \setminus \{d\}$.

2.4 Let τ_i^d denote time duration of the *cluster delay*, $d \in D_i$ corresponding to the train *i*. If the *cluster delay* occurs before $\beta(l)$, go to *Step 3.1*. Otherwise, go to *Step 3.2*.

Step 3. (Modify train departure time at chosen terminal)

3.1 Force the train to depart earlier from terminal *l*. $\beta(l) = \beta(l) - \tau_i^d / 2$. Let $\mu = \beta(l) - \alpha(l)$. If $\mu \le 0$, $\beta(l) = \beta(l) + \mu + \varepsilon$, where ε is a small amount of time. Go to *Step 4*.

3.2 Force the train to depart later from terminal *l*. $\beta(l) = \beta(l) + \tau_i^d / 2 + \upsilon$, where υ is the time interval between the prior train departure time and the starting time of the cluster delay.

Step 4. (Adjust departure and arrival times to avoid conflicts)

Let $\alpha(n)$ and $\beta(n)$ indicate the departure and arrival times, respectively, for every terminal $n \in E_j$ visited in order by train $j \in I^k$, $k \in K$, $i \ne j$. Let trains *i* and *j* be two consecutive trains departing from or arriving at the same terminal. Based on the new departure time obtained from *Step 3*, find the new arrival time $\alpha(m)$ and departure

50

time $\beta(m)$ for every $m \in E_i$ that will be visited in order by train i after terminal l by identifying the new movement arcs and arrival arc along the space-time network, where no conflicts exist with train j:

For train i and j, let $|\Delta(\beta(m), \beta(n))|$, where $m = n$, denote the difference between the two trains' arrival times and let $|\Delta(\alpha(m), \alpha(n))|$, where $m = n$, denote the difference between the two trains' departure times.

(1) $|\Delta(\beta(m), \beta(n))| \geq$ minimum departure headway.

(2) $|\Delta(\alpha(m), \alpha(n))| \geq$ minimum arrival headway.

(3) $\alpha(m) > \alpha(n)$ and $\beta(m+1) < \beta(n+1)$, where $m+1=n+1$. Note that terminal $m+1$ is visited by train i immediately after terminal m and terminal $n+1$ is visited by train j after terminal n.

If any of the arrival/departure times cannot be adjusted, return to *Step 2.3*; otherwise, the *train slot* is successfully modified. Return to *Step 2.1*.

Step 5. (Select route for creating additional train slots)

5.1 Let H^k denote the frequency of the additionally constructed *train slots* on route k. $H^k = 0$, $\forall k \in K$. Let h denote the minimum allowable headway between trains. Go to *Step 5.3*.

5.2 Check every $k \in K$. If no residual capacity is available for additional *train slots* along route k, $K = K \setminus \{k\}$.

5.3 If $K = \emptyset$, terminate. Otherwise, let B denote the set of routes that do not meet train frequency requirements. $B = K$.

5.4 If $B = \varnothing$, return to *Step 5.2*. Otherwise, select a route $k \in B$ with the longest travel time. $B = B \setminus \{k\}$. Construct a *train slot* indicated as i along route k beginning from the origin, employing the available track capacity of A_T in G_T.

Step 6. (Construct a train slot along the space-time network)

6.1 Choose feasible departure arc (α, u^q)**,** where $u \in E_i, \alpha = f$.

Search the earliest available track time capacity between pseudo source node α and the origin terminal f in the arc set A_T. If a feasible train departure arc cannot be identified, return to *Step 5.4*.

6.2 Choose feasible movement arc (u^q, v^r), where $v \in E_i, u \in E_i, u \neq v$.

Let q denote train i's earliest feasible departure time at terminal u. Starting from time q, search the earliest available track time capacity from terminal u to terminal v. If a feasible movement arc cannot be identified, return to S*tep 5.4*. If the arrival terminal v is a destination terminal g, go to *Step 6.4*. Continue, otherwise.

6.3 Choose feasible process/siding arc (v^r, u^q), where $v \in E_i, u \in E_i, u = v$.

Let r denote train i's arrival time at terminal v and q denote the time when the train finishes its required activity at terminal v. Return to *Step 6.2*.

6.4 Choose feasible arrival arc (v^r, β), where $v \in E_i, v = g$.

Train i's slot is obtained. $H^k = H^k + 1$. Return to *Step 5.4*.

This procedure takes as input the train schedule produced by the *initial track capacity allocation technique* and produces a more efficient train schedule with reduced delays and improved utilization of residual track capacity.

4.4.3 The train slot selection model

The goal of the *train slot selection model* proposed herein is to construct a train timetable (i.e. select the set of train slots that are to be operated along a given set of routes) such that operational costs and *train delays* are minimized. This model, consisting of a binary multicommodity flow problem formulation, is described in subsection 4.4.3.1 and a solution technique based on concepts of column generation is proposed for its solution in subsection 4.4.3.2.

4.4.3.1 Mathematical formulation

Given an estimate of delay (see section 4.4.1.2 for addition detail), suggested routes (defined externally) and frequencies , the timetable will be constructed such that an additive function of the delays from scheduled arrival times at the destinations and total operational cost along the corridor is minimized. Operational costs considered herein include the service charges that arise from operation and swapping of locomotives, infrastructure charges and track access charges. The decision-maker's preference with respect to delay and cost minimization can be reflected by including appropriate weights on the delay and cost components of the objective function.

The train scheduling problem is formulated as an integral multi-commodity network flow problem that relies on a train slot representation of the track capacity of each route. Each decision variable in this model is a path-based binary variable representing whether or not the train slot is selected in the timetable. The train slot representation is constructed based on the space-time network described in section 4.3. Solution of the model can provide a train timetable for a given planning period for which demand is known.

Notation used in developing this formulation is given as follows.

<u>Notation</u>

K : set of routes

L : set of shipments loading/unloading terminals

I^k : set of candidate train slots operating on the route $k \in K$

c_i : operational cost for each train slot $i \in I^k$

η^k : suggested train frequency for each route $k \in K$

$\rho(\mu_i)$: penalty added to the cost of train slot $i \in I^k$ that is imposed when train delay μ_i

      exists

u_{lt} : maximum number of train slots which can transport shipments generated at

      unloading/unloading terminal $l \in L$ at time (i.e. day) $t \in T$

$$y_i = \begin{cases} 1, & \text{the train slot } i \in I^k \text{ is operated;} \\ 0, & \text{otherwise.} \end{cases}$$

$$\delta_{lt}(y_i) = \begin{cases} 1, & \text{the train slot } i \in I^k \text{ visit terminal } l \text{ at time } t; \\ 0, & \text{otherwise.} \end{cases}$$

Model formulation

      The formulation given in (8) through (11) is referred to herein as the path-based train slot generation formulation.

$$\text{Min } z(x) = \sum_{k \in K} \sum_{i \in I^k} c_i y_i + \sum_{k \in K} \sum_{i \in I^k} \rho(\mu_i) \tag{8}$$

subject to

$$\sum_{i \in I^k} y_i = \eta^k, \qquad \forall k \in K \tag{9}$$

$$\sum_{k \in K} \sum_{i \in I^k} \delta_{lt}(y_i) y_i \leq u_{lt}, \qquad \forall l \in L \quad \forall t \in T \tag{10}$$

$$y_i \in \{0,1\}, \qquad \forall i \in I^k \quad \forall k \in K \tag{11}$$

The objective given in equation (8) seeks to minimize the total delay incurred along the corridor and total operational costs required to transport the shipments within the corridor. Constraints (9) ensure that the total number of train slots employed along the corridor on each route must be operated to satisfy the suggested train frequency η^k on route k. Constraints (10) force the number of train slots that pass each drop-off/pick-up terminal l in a given time interval t to be no larger than the number of train slots necessary to transport the shipments at terminal l in time interval t. Binary integral requirements of the decision variables are given in constraints (11). Thus, a train slot-based binary multi-commodity network flow formulation is provided with simple constraint structure.

To ensure that a nonempty feasible region exists, the relationship between u_{lt} and η^k must be defined. u_{lt} is obtained by calculating the number of train slots required to carry the shipments that arrive at terminal l in time interval t. η^k is calculated from u_{lt}. Assume that the number of candidate train slots operated on route $k \in K$ is greater than the number of suggested trains, η^k.

Observation 1: If $u_{lt} \geq \eta^k$, where $l \in L$, $t \in T$, and l is on route k, the corresponding

constraint, $\sum_{k \in K} \sum_{i \in I^k} \delta_{lt}(y_i) y_i \leq u_{lt}$, is redundant.

Observation 2: If $\sum_t u_{lt} < \eta^k$, where $l \in L$ and l is on route k, the problem as

formulated is infeasible.

The lower bound of u_{lt} could be zero, since it is possible that no potential trains will pass terminal l at time t. For any $l \in L$ on the route k, $\sum_t u_{lt} \geq \eta^k$, ensuring that enough

candidate trains can be selected to meet the suggested train frequency. Thus, conditions described in observation 2 cannot arise and there exists at least one feasible solution.

In the train slot selection model, each binary decision variable, y_i, where $i \in I^k$, represents a potential feasible train slot for route $k \in K$. These train slots are generated to ensure that even if all train slots are operated, there will be no conflicts incurred within any track segment. This input will be generated based on a given demand at terminals and an initial timetable (described in the section 4.4.1).

4.4.3.2 Solution technique

The train slot generation formulation in (8)-(11) is a binary integer program with a block-angular structure. If constraints (9) were omitted, this problem formulation could be separated into a set of subproblems, one for each drop-off/pick-up terminal associated with each time interval. This structure can be exploited by a column generation based technique that can quickly generate an optimal or near-optimal solution. Column generation has been successfully applied to solve many large-scale optimization problems in, for example, vehicle routing (Desrochers et al., 1984), air crew scheduling (Lavoie et al., 1988), lot sizing and scheduling (Cattrysse et al., 1993).

Column generation is a price-directive method that, when applied in solving the multi-commodity flowproblem, decomposes the problem into single commodity network flows. Tolls (or prices) are placed on the bundle constraints that complicate finding a solution. This approach exploits the fact that constraints (9) are independent for each $k \in K$ and only constraints (10) are dependent among each $k \in K$.

The key idea in column generation is to never explicitly list all of the columns (i.e. decision variables) of the problem formulation, but rather to generate them only "as needed."

A column for the given formulation represents a train slot. Only a subset of the columns (train slots) are considered at each iteration. This smaller program is referred to as the restricted master problem. The restricted master problem is solved to optimality by the simplex method. Whether the solution is optimal for the original program or whether additional columns must be added to improve the solution is assessed at the end of each iteration. A subproblem for each commodity is used to generate a new column for the restricted master problem and solution of the subproblem is used to prove optimality of the current solution. The potential column (train slot) with the most negative cost in each subproblem will be added to the restricted master program.

The column generation procedure is given as follows (see Ahuja et al., 1993 and Hu T. C., 1963 for additional detail). Text in bold will be discussed in detail in the following subsections.

Step 1. Initialization. Choose a set of train slots as an initial basic feasible solution.

Step 2. Solve the **restricted master problem**. Solve the problem by using the simplex method to determine the value of the dual variables.

Step 3. Check if a new column can be generated. Use the dual variable values of the solved restricted master problem to update the cost coefficient of the **subproblem**. Get new columns with largest negative reduced costs based on subproblem solutions and add the new columns to the restricted master problem. Return to *Step 2* if a new column is found. Otherwise, terminate the procedure. The optimal solution is obtained.

Restricted master problem

The goal of the restricted master problem is to obtain the value of the dual variables so that the reduced cost for each train slot can be calculated for the subproblems. Since the

train slot selection formulation ((8)-(11)) has a block-angular structure, the formulation associated with a smaller set of variables in which the integrality constraints (11) are relaxed can be treated as the restricted master problem.

Subproblem

The goal of the subproblem is to find the column (train slot) with the minimum reduced cost to be added to the master problem. If the minimum reduced cost is nonnegative, then we can terminate the column generation procedure and the problem is solved to optimality. Let σ^k denote the dual variable corresponding to each route k in constraints (9) and ϖ_{lt} denote the dual variable corresponding to each terminal l at time (i.e. day) t in constraints (10). The reduced cost, λ_i^k, of the column corresponding to the restricted master problem is given by (12). Each route k has its own subproblem. The value of reduced cost of each column, λ_i^k, can be treated as the benefit (i.e. the reduction in the train operational and penalty costs) obtained by using a train slot i on route k.

$$\lambda_i^k = c_i + \rho(\mu_i) + \sum_{t\in T}\sum_{l\in L}\delta_{lt}(y_i)\varpi_{lt} - \sigma^k, \qquad \forall i \in I^k \ \forall k \in K. \tag{12}$$

Hence, equation (12) is considered for each train slot i operated on route k to check if the reduced cost of any column is negative. The column (train slot) with the most negative reduced cost will be added into the restricted master problem. Thus, the train operational and penalty costs incurred from delivery delay can further be reduced by selecting this train slot to transport the shipments.

The integrality constraints in the train slot generation formulation are relaxed while the model is solved by the column generation-based technique. Thus, the solution obtained by the column generation technique for the *train slot selection model* may contain non-

58

integer values. When this occurs, the variables that have non-integer values will be rounded to the closest integer value. If the resulting rounded solution is not feasible, the number of train slots on every route that must be removed from or added to the solution will be identified. When a feasible solution can be obtained by adding train slots to the timetable, train slots will be considered in order of cost. When train slots must be removed from the timetable to obtain feasibility, train slots with the highest cost will be eliminated first.

4.5 Elastic Demand

The solution approach when embedded in a simulation-based iterative framework (REORIENT Consortium, 2007), where demand for rail services is re-computed in accordance with the train schedule obtained by solving the freight train scheduling problem, results in an optimal or near optimal schedule given elastic demand.

4.5.1 Iterative simulation framework

The models and solution techniques described in section 4.3 address the scheduling problem, where demand is assumed to be fixed and known. However, the demand for the services depends on service characteristics, including the routes along which services are offered, frequency, expected arrival times at the destinations and other measures of service level. Likewise, the schedule is developed with the goal of providing a high level of service for the known demand, but at a low cost to the carrier. An iterative microscopic simulation-based framework is described that, employing techniques from section 4.4, produces a timetable for elastic demand, where the demand for the services depends on the services offered and services offered depend on the demand.

Interactions between the *train slot selection model* (including other supporting tools developed in section 4.4) and the simulation platform are presented in Fig. 4-3. The process begins by developing an initial timetable based on some estimate of demand. This initial

timetable can be produced through the procedures described in section 4.4 that seek an optimal timetable for the scheduling problem with inelastic demand. Alternatively, that process can be terminated after only the *initial track capacity allocation technique* described in section 4.4.1.2 is applied. This latter, simpler approach was employed in the experiments described in section 4.6.

Once an initial timetable has been developed, a mode choice model is applied to estimate the number of zone-to-zone shipments (defined as containers or carloads requiring transport from shipment origin to destination) that are attracted to the services offered based on the given timetable. The mode choice model employed within this framework was developed from the perspective of the shipper and is, therefore, sensitive to the characteristics of both the shipment and the usual transportation service level attributes. Additional detail can be found in (REORIENT Consortium, 2007). The simulation assigns the shipments to the train slots of the timetable, tracks their progress through the rail-based IM system, calculates shipment delays at terminals, classification yards and border crossing, and hence, computes the resulting *shipment delays*. These shipment delays are fed back to the *track capacity modification technique* and *train slot selection model* processes (of section 4.4) to produce a new timetable that better meets the new estimates of demand. The resulting updated timetable is fed back to the simulation model and the process is repeated until termination criteria are met. The simulation platform is described in detail in (Mahmassani et al., 2007; Zhang et al., 2007).
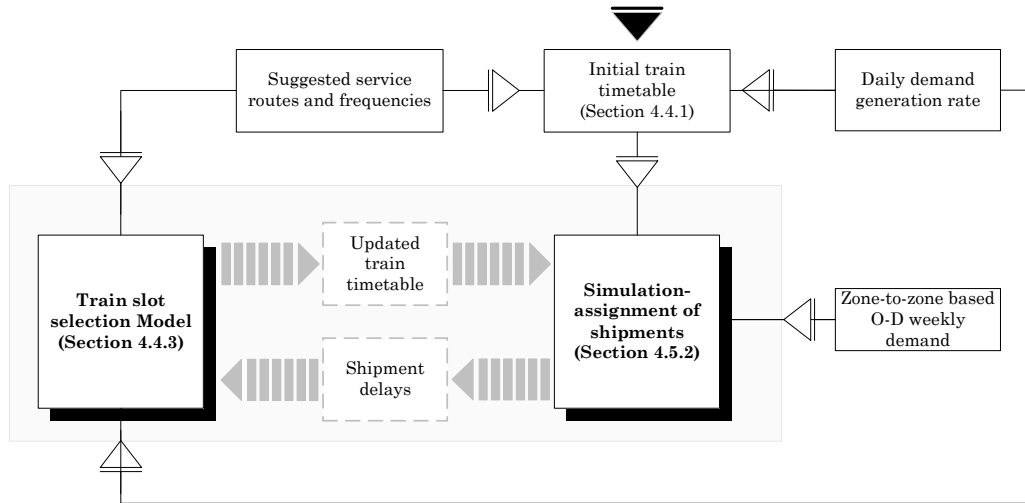
Figure 4-3 Iterative simulation-based scheduling approach for elastic demand

A state of equilibrium between the carrier and shipper is expected to be reached through the use of this iterative framework. The carrier improves its services by updating its timetable based on the response of the shippers to proposed services as estimated through the mode choice model. The train timetable is revised with consideration for the objectives of both the shipper and carrier. An equilibrium is met when changes to the timetable result only in a reduction in mode share for the rail-based IM services.

4.5.2 The simulation-assignment framework

The network modeling platform represents supply and demand in the network and captures their dynamic interaction through a simulation-assignment framework. The platform simulates freight at the disaggregate level of a shipment. The shipment will be transported along a sequence of links that are serviced by available modes (including existing rail services, proposed rail services, trucks, ships) with feasible IM transfers. A joint mode, path, and service choice is modeled for each shipment. A generalized cost function is used to determine the choice between truck-only paths and rail-based IM paths. Each path has associated attributes that are considered explicitly by the shipper: travel time, and transport

price. The commodity to be shipped and its value, time-sensitivity, and level of hazard, also are used to evaluate the utility of a particular path. The choice of a particular mode reflects the shipper's preference for a mode-path combination that is best suited for the commodity to be shipped.

The exact time at which a shipment will be set for loading on a train at a given terminal will depend on the time of arrival and required processing time (for classification, unloading and other operations at IM terminals and ports) at that terminal. Processing times cannot be known precisely in advance. Thus, in the simulation, such arrival and processing times are randomly assigned and *shipment delays at terminals* are then computed. These delays contribute to future IM shortest path computations and ultimate path assignment, and provide the necessary input for train slot adjustment in the *track capacity modification technique*. Additional details concerning the network modeling platform are described in Arcot et al. (2007), Mahmassani et al. (2007), and Zhang et al. (2007).

4.6 Real World Application

The iterative scheduling solution technique described in section 4.5 was employed over a European rail network (shown in Chapter 2) spanning 11 countries, connecting the Nordic Region to Romania and Greece. Interoperability legislation imposed by the European Commission (EC) of the European Union (EU) on all member states was designed to aid in transforming the European railways from currently existing nationally fragmented systems into an internationally integrated IM freight transport system. Numerous obstacles exist to the creation of an internationally integrated IM freight transport system in this region. The harmonization of these fragmented systems, however, is crucial to creating a competitive and profitable international system. Whether or not the new international system will be able to

provide competitive service as compared with other modes, e.g. transport by truck, depends on the efficiency of the system as measured by both carriers and shippers. The development of a timetable that is designed to serve this international freight business and that considers both decision-makers in the process will lead to service levels that are attractive to shippers, but sustainable for the carriers. That is, how the system is operated will affect the time required to complete deliveries and cost of operating the system. This, in turn, will affect demand for the system.

4.6.1 The REORIENT network

The network representation of the REORIENT network created for, and employed in, this work consists of 5,577 rail links, 5,753 rail nodes (i.e. terminals, classification yards, stations and border crossing points), 4,713 road links, 5,753 road nodes, 54 sea links and 21 port nodes. The rail link lengths range from 0.009 to 20 km. Approximately 20% of the links are single track and 80% are double track. The maximum speed on the tracks over the network is between 60 km/hr and 80 km/hr and depends on the track segment. The available terminals where shipments can be loaded or unloaded are primarily located in Sweden, Poland, Austria, Hungary, Romania, and Greece.

Zone-to-zone (origin-destination, OD) freight demand data is employed in this work. Approximately 3.2 million freight shipments traversed some portion of the REORIENT network in 2006 (TRANSTOOLS). These shipments are categorized into 22 commodity types. Each type can further manifest as either containerized or bulk units. Shipments are continuously generated from Monday to Thursday with 65% of the shipments generated split evenly between Monday and Thursday and the remaining 35% of shipments split between Tuesday and Wednesday. Thus, the node-to-node shipments are generated from a known,

fixed, and deterministic demand generation distribution model embedded in the simulation platform (REORIENT Consortium, 2007).

Four southbound service design options, developed in consultation with rail carriers from the region and founded on market-based research, have been proposed for the REORIENT Network. For the purposes of this analysis, these services are permitted to carry both bulk and unitized flows. The routes associated with these service designs are shown in Fig. 2-2.

T1: Halsberg-Trelleborg-Swinoujscie-Vienna/Bratislava-Budapest

T2: Trelleborg-Swinoujscie-Bratislava/Vienna

T3: Gdansk/Gdynia-Bratislava/Vienna-Budapest-Beograd-Thessalonica

T4: Bratislava-Budapest-Bucharest-Constantia

Fifteen loading or unloading terminals are specified for access to these routes, including Sofia, Arad, Bucuresti, Budapest, Thesssaloniki, Gdansk, Poznan, Wiencont, Swinoujscie, Constanta, Bratislava. Mutually beneficial multi-carrier train timetables were developed with the proposed CDM strategies for the operation of these four routes. Operations along these routes will also affect the temporal and spatial patterns of flows traversing other portions of the REORIENT network.

4.6.2 Results of experiments

A five-day planning horizon is considered (i.e. Monday through Friday) in this simulation analysis. The resulting periodic timetable scheduled along the four service routes is assumed to be used repetitively, i.e. repeating every Monday. The input required for the simulation-based analysis includes: the REORIENT Network topology, the attributes of the network (rail link length, number of tracks, terminal and classification yard locations, travel

speeds), zone-to-zone (origin-destination) freight demand data, service routes and a train timetable for operating the service routes.

The initial track capacity allocation and track capacity modification techniques and the train slot selection model were coded in Microsoft Visual C++ 6.0 and CPLEX 10.1 was employed. These codes were run on an Intel Pentium 4 with 3.6 GHz CPU and 2 GB RAM using the Windows XP operating system. The simulation platform was run on a Dell Precision PWS690 Intel(R) Xeon(R) with two 3 GHz CPU5160s and 16 GB RAM. Three iterations were conducted (including the initial run of the initial track capacity allocation technique). The timetable produced by the initial track capacity allocation technique in the first iteration produced a timetable with 103 trains (16 on T1, 24 on T2, 25 on T3, 38 on T4), with average train delay of 22.9 hours. 378 cluster delays were identified by the train slot modification technique and an additional 138 trains were generated from the resulting residual track capacity for inclusion in the timetable. The train slot selection model selected 171 of the 241 potential trains (27 on T1, 39 on T2, 40 on T3, 65 on T4) with a final objective value of 782,167 and average train delay of 14.1 hours. In the final iteration, 236 cluster delays were identified by the train slot modification technique and an additional 101 trains were generated from the resulting residual track capacity for inclusion in the timetable. The final timetable was composed of 197 trains (31 trains on T1, 47 trains on T2, 49 trains on T3, 70 trains on T4) selected from 272 trains through solution of the train slot selection model. The resulting objective function value was 758,167 and an average train delay of 10.2 hours was obtained. Note that the resulting train timetable was able to accommodate a greater number of trains (197 as compared to 171) with reduced average delays (10.2 hours as

compared to 14.1 hours) and reduced total objective function value (758,167 as compared to 782, 167).

The resulting 197 trains are scheduled to carry over 370,000 tons in a given week along the new services. Introduction of these new services led to a 6.5% increase (representing an increase of nearly 700,000 tons) in rail-based IM flows over the entire REORIENT network as compared with the 2006 base (where no new services are introduced). Thus, with improvements that aid in overcoming administrative, technological, and business barriers to interoperability and optimally or near optimally allocating track capacity can lead to significant increases in IM market share for the rail-based IM services.

4.7 Conclusions

The *train slot selection model* based on concepts of multicommodity network flows, along with supporting tools (i.e. the *initial track capacity allocation* and *track capacity modification techniques*), is developed for determining freight train timetables for scheduling international rail services along a network of interconnected routes. A column generation-based approach is proposed for its solution. The train slot selection model seeks to minimize operational costs and delays in delivery from scheduled arrival times. To address the elasticity of demand to service levels, the solution approach is embedded in a simulation-based iterative framework, where demand for rail services is re-estimated in light of the train schedule that results from solution of this freight train scheduling problem.

The solution technique was employed over a pan-European rail-base IM network. Existing rail service schedules are fragmented, with little coordination across international boundaries. Results of the simulation experiments demonstrate the potential for new international rail freight service in the REORIENT network and the potential role that

optimal track capacity allocation can play. Establishment of an integrated, international, rail-based IM freight transport system may contribute to opening new markets and increased trade along the supply network, thereby fostering economic growth in the region.

The optimal or near-optimal allocation of track capacity in a rail-based IM freight transport network, as can be completed through proposed modeling and solution techniques, will permit efficient and cost-effective rail-based IM transport, resulting in a competitive alternative to truck transportation.

Chapter 5   Collaborative Decision-Making in Train Scheduling

5.1 Introduction

This paper proposes collaborative decision-making (CDM) strategies for the collaborative operation of international rail-based intermodal (IM) services by multiple carriers. The benefits of the proposed techniques are assessed using a discrete-time carrier collaboration simulation-assignment framework on a real-world European IM network spanning 11 countries from the Baltic (Scandinavia) to the Mediterranean, Greece through Bulgaria, Czech Republic, Hungary, Poland, Romania and Slovakia, termed the REORIENT network (depicted in Fig. 5-1). Existing rail-based IM services are fragmented and are typically operated by publicly owned rail-companies. In fact, Network Statements (see for example the Network Statement for Finland (2006) from REORIENT countries indicate that at least one carrier exists in every country with the exclusive business of national rail transport. Often, the rail infrastructure is state-owned. Track access rights must be obtained for carriers of foreign countries to operate their trains internationally. Despite that European Commission (EC) directives have both legally and functionally separated rail operations from infrastructure ownership and management, the government-owned national railways still maintain a symbiotic relationship with infrastructure providers. New entrants to rail business face considerable hurdles in terms of access to infrastructure and operations at border crossings. Moreover, passenger traffic has precedence over freight traffic. Thus, train timetables are created with priority for national carriers, leaving only residual track capacity for international freight traffic.

Figure 5-1 The REORIENT network

The EC's Interoperability Directives envision an environment in which new sufficiently-capitalized entrants could enter and meet market needs through various types of specialized freight services. To enable this requires the ability to request and obtain slots (a slot, referred to as a train slot, is defined herein as the use of track capacity along a specific stretch of track for a given short period of time) in a timely manner. While the process is progressively becoming more transparent, rules for allocating slots remain riddled with inefficiency. In a low traffic environment, slots may not be scarce resources, and some inefficiency in its allocation may be tolerated. However, there are indications that certain portions of the rail network under consideration are already exhibiting high levels of

69

utilization, and slots will eventually come to be viewed as the valuable resources they are. Under the objectives of the European Commission which motivated this work, it is envisioned that considerable increases in rail freight traffic could be expected for new services coupled with various technological, administrative and operational improvements (REORIENT Consortium, 2007). In such an environment, flexible means for utilizing slots become essential to attaining the desired service levels and associated efficiencies necessary to contain the cost of providing the service. Such flexible means fall under the general umbrella of CDM schemes, which constitute a class of approaches for the management of shared or public resources by a collection of private and public entities or agents with individual goals.

The available slots for operating international trains given national timetables can be patched together to form international train timetables and routes. These available slots (or bundles of slots) are sold for operation by various carriers. This mechanism of allocating slots can lead to inefficiencies that can be mitigated through cooperative agreements between carriers. Three strategies for cooperation, i.e. CDM schemes, designed to overcome these inefficiencies associated with operating across the countries of the REORIENT network, are proposed in this work: (1) train slot cooperation, (2) train space leasing and (3) train slot swapping techniques.

The three proposed CDM strategies take as input the carrier train timetables and the predetermined routes along which the trains will operate. Through the CDM schemes involving slot leasing, swapping and other mechanisms, these timetables can be improved, mutually benefiting all carriers. Thus, if different carriers, possibly from different countries, could cooperate with each other (through the sharing of information and resources, e.g. slots

or locomotive power), barriers to entry or to reliable service that may exist in such fragmented IM networks as the REORIENT network, could be overcome.

Industry structure in Europe continues to evolve, with various possible business models emerging in different parts and segments of the market. The result will be a mix of multinational carriers operating services across borders, as well as evolved national undertakings with integrated services, and other possible combinations. In all of these cases, the problem of slot allocation and management will play a critical role in the efficient and competitive use of the infrastructure.

CDM strategies proposed in this work will be assessed through a discrete-time carrier collaboration simulation model (REORIENT Consortium, 2007) that replicates services, carrier operations and shipper response to the revised (more efficient) timetables. The platform makes it possible to model variability in such aspects as delays at the classification yards; time required for IM transfer at terminals, ports and border crossings; and required travel times. The increase in rail-based IM market share that results from the introduction of more efficient CDM-based timetables is estimated in the simulation platform. This can be compared with the market share anticipated from non-collaboratively derived timetables.

5.2 Background Review on CDM

CDM involves teamwork through communication, cooperation and coordination among each of the agents in the team (Berzins and Dhavala, 1998). While earlier forms of CDM were envisioned and performed through debate and negotiation among a group of people, modern incarnations rely extensively on sophisticated collaboration support systems that allow most activities and interaction to occur virtually through well-defined frameworks and protocols. Conflicts of interest are inevitable and support for achieving consensus and

71

compromise is required. For problems in which agents compete, but where there is an opportunity to cooperate, an improved solution for each agent might be achieved by incorporating CDM. CDM has been applied in many works addressing, for example, air traffic flow management, supply-chain systems, submarine command and control, engineering design projects, and homeland security problems (Ball et al., 2000; Carlson, 2000; Chang et al., 2001; Sherali, et al., 2003; Groothedde et al., 2005).

Among these works, the works in air traffic flow management are the most relevant, especially the aircraft arrival and departure slot arrangement, which, like track capacity allocation in rail-based IM freight transport in the REORIENT network, is a capacity allocation problem (Ball et al., 2000; Vossen and Ball, 2005; Vossen and Ball, 2006). The goal of the aircraft arrival/departure slot arrangement is to minimize delays incurred at congested airports. Through a procedure built by CDM, arrival or departure slots to an appropriate aircraft are assigned to minimize the total delay of airlines; thus, arranging slots more efficiently. Airlines can benefit from cooperating with each other despite that they are inherently competitive.

5.3 The REORIENT network

A five-day planning horizon is considered (i.e. Monday through Friday) in this simulation analysis. The resulting periodic schedule is assumed to be used repetitively, i.e. repeating every Monday. The input required for the simulation-based analysis includes: the REORIENT network topology, the attributes of the network (rail link length, number of tracks, terminal and classification yard locations, travel speeds), zone-to-zone (origin-destination) freight demand data, service routes and a train timetable for operating the service routes.

The network representation of the REORIENT network created for, and employed in, this work consists of 5,577 rail links, 5,753 rail nodes (i.e. terminals, classification yards, stations and border crossing points), 4,713 road links, 5,753 road nodes, 54 sea links and 21 port nodes. The rail link lengths range from 0.009 to 20 km. Approximately 20% of the links are single track and 80% are double track. The maximum speed on the tracks over the network is between 60 km/hr and 80 km/hr and depends on the track segment. The available terminals where shipments can be loaded or unloaded are primarily located in Sweden, Poland, Austria, Hungary, Romania, and Greece.

Zone-to-zone (origin-destination, OD) freight demand data is employed in this work. Approximately 3.2 million freight shipments traversed some portion of the REORIENT network in 2006 (TRANSTOOLS). These shipments are categorized into 22 commodity types. Each type can further manifest as either containerized or bulk units. Shipments are continuously generated from Monday to Thursday with 65% of the shipments generated split evenly between Monday and Thursday and the remaining 35% of shipments split between Tuesday and Wednesday. Thus, the node-to-node shipments are generated from a known, fixed, and deterministic demand generation distribution model embedded in the simulation platform (REORIENT Consortium, 2007).

Four southbound service design options, developed in consultation with rail carriers from the region and founded on market-based research, have been proposed for the REORIENT network. For the purposes of this analysis, these services are permitted to carry both bulk and unitized flows. The routes associated with these service designs are shown in Fig. 2-2.

T1: Halsberg-Trelleborg-Swinoujscie-Vienna/Bratislava-Budapest

T2: Trelleborg-Swinoujscie-Bratislava/Vienna

T3: Gdansk/Gdynia-Bratislava/Vienna-Budapest-Beograd-Thessalonica

T4: Bratislava-Budapest-Bucharest-Constantia

Fifteen loading or unloading terminals are specified for access to these routes, including Sofia, Arad, Bucuresti, Budapest, Thesssaloniki, Gdansk, Poznan, Wiencont, Swinoujscie, Constanta, Bratislava. Mutually beneficial multi-carrier train timetables were developed with the proposed CDM strategies for the operation of these four routes. Operations along these routes will also affect the temporal and spatial patterns of flows traversing other portions of the REORIENT network.

5.4 Collaborative Strategies

The three CDM strategies proposed in this work, the train slot cooperation, train space leasing, and train slot swapping techniques, rely on various mechanisms for collaboration among carriers. The means of collaboration considered in this work include: joint operation of train slots, exchange of train slots between carriers, and leasing of train capacity. These three CDM strategies are described next.

5.4.1 Train slot cooperation

In the train slot cooperation approach, two or more carriers can join forces to jointly operate a train slot. Carriers operate over separate portions of the train slot's route (e.g. nearly all carriers operating within the REORIENT network operate only within a specific country), such that operation along the entire route is carried out through the cooperation of multiple carriers. Thus, through collaboration with other carriers, carriers can transport shipments with origins or destinations that are not covered by the carrier's own service routes. This strategy is similar to that of code share concept in the airline industry where two or

more airlines operate several portions of an entire route. Thus, passenger can fly from his (or her) origin to destination through the cooperation of the multiple airlines.

This method of joint operation of a train line is particularly relevant in the REORIENT network, where track access rights may not be granted to foreign carriers in one or more of the countries on a particular route. Even if track access rights could be obtained, operation across borders is often cumbersome and costly, requiring alternative or specialized equipment (as differences in, for example, power or track gauge often exist), training, and knowledge (e.g. of local language). In such instances, the shipment will be transferred from one carrier's train to another's at the border of two countries. Certain operations may be required at borders, where one of the border countries does not provide track access rights to rail carriers from the other border country or where partnership agreements for joint operation have been enacted. However, the time required for such operations may be reduced if two carriers, each of which is permitted or better suited to operate within its own country, were to collaborate on the shipment through information sharing.

Fig. 5-2 illustrates such operations at a border. Carriers A and B operate in bordering countries. Assume that neither carrier is given track access rights to operate in the other's country. Carriers A and B co-transport a shipment from origin X to destination Y. The shipment is transported by Carrier A from origin X to terminal Z, which is located at the border between the two countries. The shipment is then unloaded from Carrier A's train and reloaded to Carrier B's train. Alternatively, the shipment can be transported directly from origin to destination if the carriers are willing to share their rolling stock. It is also possible that they might choose to simply switch locomotives such that the locomotive running the

train is owned by the carrier that is operating the train. This requires appropriately gauged railcars. Transport by Carrier B of the shipment continues until destination Y is reached.

Suppose a carrier is able to obtain track access rights in all countries en route, but the carrier has only enough cargo to fill a train a portion of the time. The latter implementation of the train slot cooperation method would allow two or more carriers to jointly use the train slot over time.
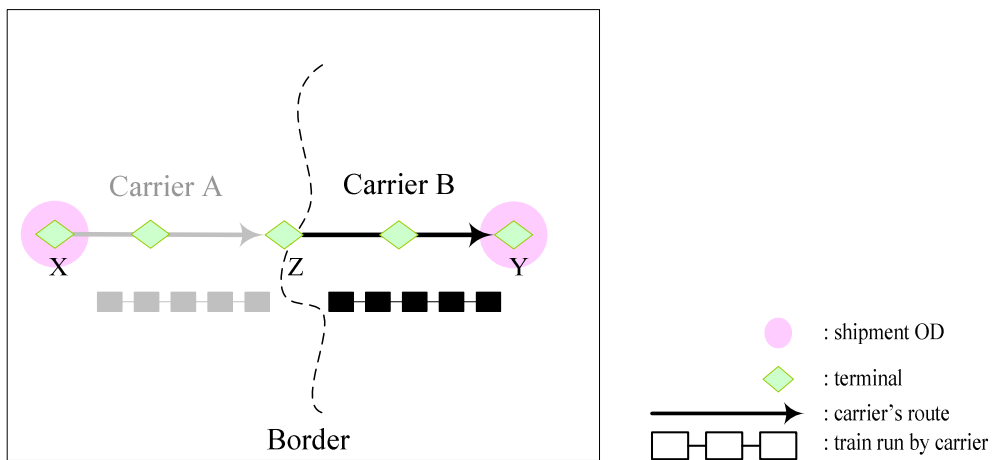
Figure 5-2 Train slot cooperation approach

5.4.2 Train slot swapping

The train slot swapping approach allows two carriers that each owns a train slot to exchange capacity rights for the slots. This can facilitate cooperation when one carrier has excess capacity in a slot and the other has newly arising need for transport along the other carrier's route. Alternatively, when two carriers have excess capacity in their train slots, each carrier might be able to improve its level of service by swapping train slots for given trains or given days of the week. Such swaps can also help the carriers to maintain delivery time windows promised to the shippers.

In Fig. 5-3, the train slots with the same O-D pair shown in gray and black are owned by Carriers A and B, respectively. The arrival time at destination Y is 11:00 p.m. on Carrier

A's slot and is 2:00 p.m. on Carrier B's slot. Suppose a delivery must be made by Carrier A between 1:00 p.m. and 6:00 p.m. Carrier A, however, will not make the deadline if it uses its own slot. Thus, Carrier A may exchange its own slot with Carrier B for the black slot that is not currently in use, thereby avoiding some penalties imposed by the shipper for late arrival. Carrier B may then choose to use the newly received train slot or may even choose to swap or lease it.
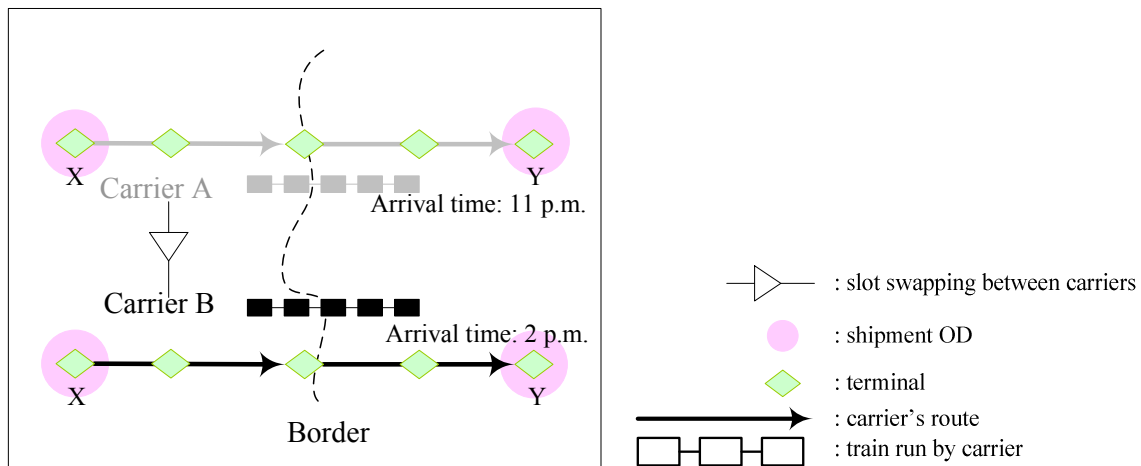
Figure 5-3 Train slot swapping approach

5.4.3 Train space leasing

Presume for a moment that slots are sold in bundles of time and must be purchased for every day of the week if purchased for a single day. It may be the case that a single carrier that owns a particular train slot cannot fill an entire train every single day of the week. The train space leasing approach proposed herein allows the carrier to lease a portion of the train capacity to other carriers. It is assumed that no carrier is willing to sell all of a train's capacity. That is, it is assumed that it would be more lucrative to swap train slots than to operate a train carrying entirely other carriers' shipments. A fixed percentage of the train's capacity will, therefore, be reserved for the train slot's owner. More than one carrier can lease a train's excess capacity. Through such an approach, the carrier who owns the slot can increase its revenue by opening the residual train capacity to other carriers. Fig. 5-4

77

illustrates this cooperation method. In this figure., a train slot that is operated from origin X to destination Y is owned by Carrier A. Suppose a container must be delivered from origin X to destination Y by Carrier B. If the residual capacity (a single train car shown in black) of a train operating in this train slot can be leased to Carrier B, both carriers can benefit. That is, Carrier A gains additional revenue by charging Carrier B and Carrier B gains by renting space on Carrier A's train without having to operate a train.
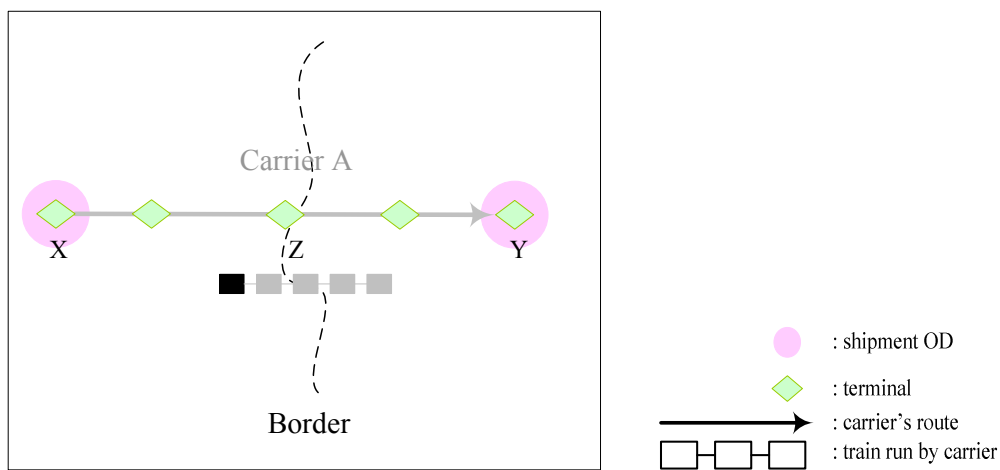
Figure 5-4 Train spacing leasing approach

5.5 Simulation-Based Framework

Analysis of the complex interactions over space and time associated with the movement of freight between OD pairs over IM freight networks with rail services involving the cooperation of multiple carriers involves many difficult subproblems. As a result, it is very difficult to describe the problem using a quantitative optimization-based model. Therefore, a carrier collaboration simulation-assignment framework (REORIENT Consortium, 2007) was developed to analyze and evaluate the proposed carrier CDM strategies that result in various IM rail freight services contemplated in the REORIENT network. The carrier collaboration simulation-assignment framework is shown in Fig. 5-5. The simulation platform is employed to evaluate services (i.e. timetables) that are generated

78

by optimization-based scheduling algorithms (described in Chapter 4) exploiting the chosen CDM strategy.

This carrier collaboration simulation-assignment platform extends an existing network modeling platform developed to analyze and evaluate proposed operational improvements and various IM rail freight services contemplated in the REORIENT network. Specific details of the simulation environment and other core network modeling and analysis capabilities developed to evaluate the effectiveness of service scenarios and operational strategies in the REORIENT network are given in (Arcot et al., 2007; Mahmassani et al., 2007). This modeling approach integrates a mode choice modeling process within a network flow assignment framework. For a given specification of services and operational strategies, this platform and its extension that explicitly recognizes multiple-carrier operations provide detailed information on flows by mode and service between the various origins and destinations in the study area. An overview of the carrier collaboration simulation-assignment platform extension is given in Fig. 5-5, followed by more detailed description of its main components.
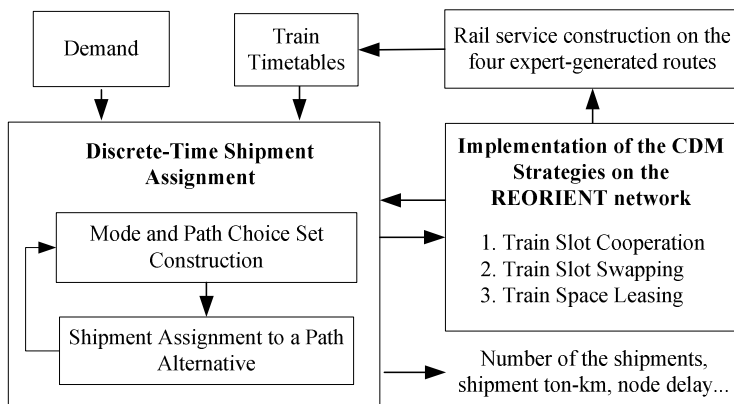
Figure 5-5 Carrier CDM simulation-based analysis

5.5.1 Implementation of the CDM strategies on the REORIENT network

The three proposed CDM strategies described in Section *Collaborative Strategies* were employed in creating mutually beneficial train timetables for the four expert-generated routes (Fig. 5-6(a)). Each implementation results in a suggested timetable. In the train slot cooperation implementation (Fig. 5-6(b)), the access-rights to the expert-generated routes (i.e. ability to operate along the routes) are assumed to belong to one carrier. Thus, in this implementation, four carriers (one associated with each of the four service routes) can collaborate with one another. That is, a shipment is permitted to be transported by any of the four carriers.

In the train slot swapping implementation (Fig. 5-6(c)), there are two carriers (Carrier A and B) operating train slots on the service routes. The train schedules for Carriers A and B were created by alternating slot assignments to carriers over time, resulting in an equitable distribution of train slots. In this scenario, to allow shorter delivery times, a shipment originally transported by Carrier A (or B) can be transferred to another train slot owned by Carrier B (or A) at any of the intermediate terminals.

In the train space leasing implementation (Fig. 5-6(d)), as in the implementation of the train slot swapping strategy, there are two carriers (Carrier A and B) operating train slots on the service routes. Unlike in the former implementation, where shipments carried by either carrier can switch carriers, in this implementation, such swapping is restricted. Carrier A can transport its shipments in a slot owned by Carrier B, but the reverse is not permitted. This replicates the renting of space by a carrier on another carrier's trains.

(a) Four expert-generated routes

(b) Four carriers in train slot cooperation

(c) Two carriers in train slot swapping

(d) Two carriers in train space leasing
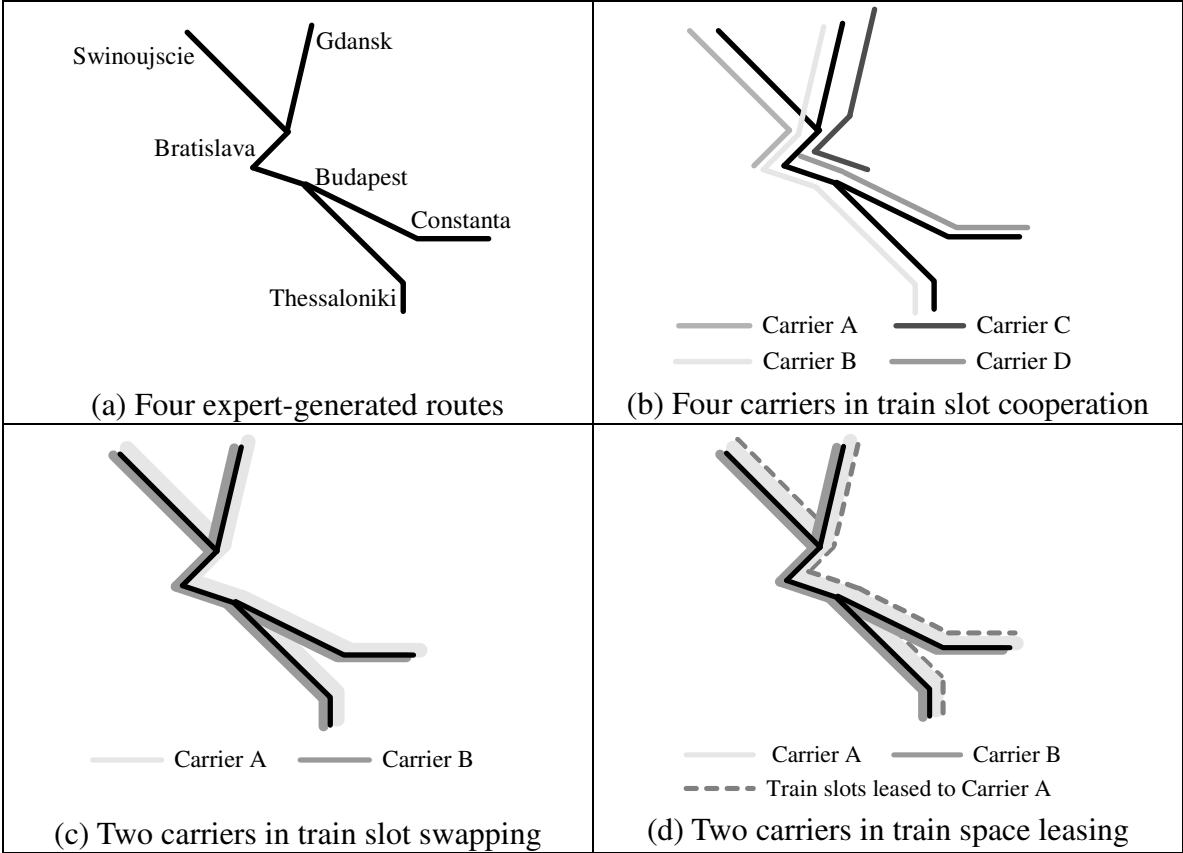
Figure 5-6 Collaborative decision-making strategies on the expert-generated routes

5.5.2 Rail service construction on the four expert-generated routes

Once a strategy is adopted, train timetables are created. Given the suggested routes, frequencies and the residual network capacity (i.e. remaining capacity after passenger and national traffic are assigned), train timetables are constructed for each carrier using a model that employs a binary multicommodity network flow program in generating a timetable for each carrier. Model formulation and proposed solution approach designed for its solution are given in Chapter 4. The model seeks to minimize an additive function of the delays from scheduled arrival times at the destinations and total operational cost along the corridor. Operational costs considered include the service charges that arise from swapping of locomotives, infrastructure charges and track access charges. The decision-maker's

preference with respect to delay and cost minimization can be reflected by including appropriate weights on the delay and cost components of the objective function. In addition to constraints that ensure that enough train frequency exists to ship the demand along the routes between origins and destinations, the model contains other constraints related to the track capacity usage that must be imposed while constructing the train timetable. Such constraints include train siding, train overtaking, and track capacity usage constraints.

5.5.3 Shipment assignment

A shipment is defined as the smallest unit of cargo (i.e. a container or carload) that will be transported from shipment origin to destination. The shipment will be transported along a sequence of arcs that are serviced by available modes with feasible IM transfers (referred to herein as a path alternative). Each path alternative is operated by a carrier. Link costs and travel times are assumed to be additive, as are node (i.e. terminal or intersection) costs and transfer delays. When faced with a joint mode and route choice set, a shipper will choose a path that minimizes the shipper's generalized cost of transporting a shipment from shipment origin at the time that the shipper takes responsibility for the shipment to its destination.

A dynamic freight assignment problem, addressed within the carrier collaboration simulation-assignment framework in an IM network, where carriers collaborate with one another in the transport of shipments is solved by determining the number of shipments for each alternative and the resulting temporal-spatial loading of shipments and conveyances. The framework features three main components: (1) freight traffic simulation, (2) a shipper behavioral model, and (3) path processing along with shipment assignments as permitted by acceptable CDM strategies. The freight traffic simulator depicts freight flow propagation in

the IM network. This facilitates the evaluation of network performance for the given set of modal and route decisions made by individual shippers. The shipper behavioral component models a shipper's mode and route selection decision in a stochastic utility maximization framework with multiple evaluation criteria. The third component is intended to generate realistic route choice sets based on the chosen CDM strategy and perform stochastic network loading required to solve the shipment assignment problem. Different CDM strategies will lead to the generation of different realistic route choice sets within the network. Very large service transfer penalties are imposed on the terminal nodes to prevent shipments from transferring to train slots operated by carriers that do not collaborate. For additional detail on the first two components of this assignment framework, see Arcot et al., (2007) and Mahmassani et al., (2007).

5.5.4 Evaluation criteria for the CDM strategy

Several evaluation criteria are proposed to assess the performance of the overall system under different service design options in the CDM scenarios. From the system's perspective, the objective is to attract more shipments to use the services and to transport these shipments in a more efficient way. That is, under the implementation of a CDM strategy, it is expected that more of the shipments will choose the proposed services than had chosen these services over truck under non-CDM operations due to improvements in distance or time required to reach the final destination. The performance is evaluated based on the number of the shipments attracted by the freight transport system, shipment ton and ton-km.

5.6 Preliminary Findings from Experimental Results

The train timetables for the proposed service routes generated by the optimization model described in Chapter 4 employing each of the selected collaborative strategies were

evaluated with the aid of the carrier collaboration simulation-assignment platform. Flows along the services in terms of tons and ton-km were generated through the assignment mechanism of the simulation framework. Changes in flow can be employed to assess changes in market share that result from the introduction of improved services that follow from the implementation of collaborative strategies for operating the rail system. Such comparisons can be made for the proposed services by considering results obtained from running the simulation model. Results of the runs are shown in Fig. 5-7, along with accompanying Table 5-1. Specifically, in the Fig. 5-7., the improvements due to the introduction of the three CDM strategies described in Section *Collaborative Strategies* are assessed by subtracting the amount of flow in tons or ton-km attracted to the services in which no collaboration among carriers is permitted from the amount of flow attracted to the services where a given CDM strategy is adopted. Related numerical results are given in Table 1, where this difference is shown for each of the four proposed service routes by adopted CDM strategy. Additionally, this difference is divided by the flows produced where no collaboration is permitted and is shown as a percentage, indicating the percent increase in flows resulting from the introduction of each specific CDM strategy.

Table 5-1 Improvement on CDM strategies compared with noncollaboration

Route	Train Slot Cooperation		Train Slot Swapping		Train Space Leasing	
	tons (%)	ton-km (%)	tons (%)	ton-km (%)	tons (%)	ton-Km (%)
T1	785 (0.64)	99,605 (0.36)	124,241 (4.46)	593,107 (2.18)	2,302 (1.88)	616,168 (2.27)
T2	412 (0.24)	1,424,630 (3.77)	172,114 (25.7)	14,040,913 (55.78)	14,921 (9.49)	7,944,896 (25.41)
T3	13,108 (4.32)	6,804,861 (5.91)	316,887 (12.09)	33,384,719 (37.67)	27,552 (9.52)	22,698,017 (22.86)
T4	9,470 (2.53)	6,048,579 (4.55)	384,146 (43.09)	46,336,764 (50.05)	85,644 (28.69)	30,132,865 (29.69)
Total	23,775 (2.44)	14,377,675 (4.59)	997,388 (23.59)	94,355,502 (40.41)	130,419 (15.04)	61,391,946 (23.03)
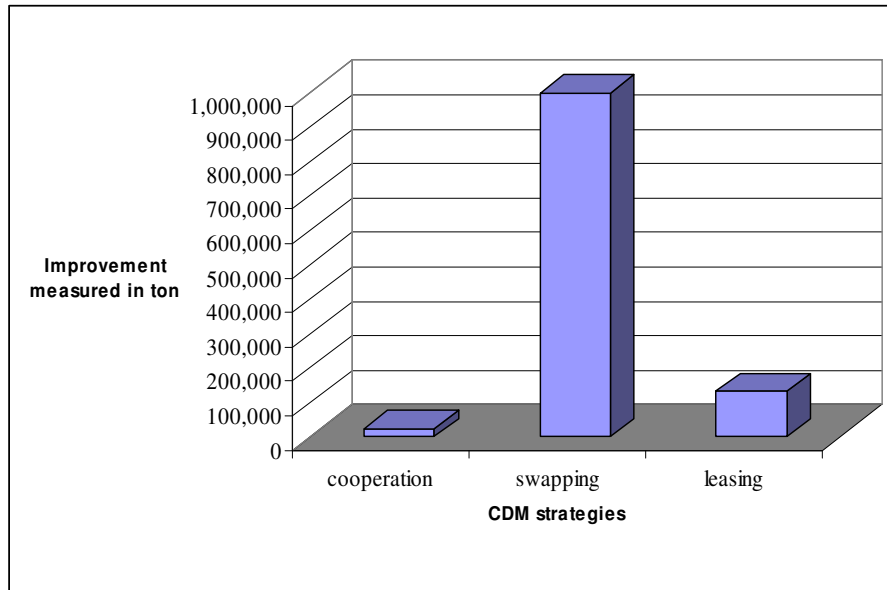
Figure 5-7a Improvement in ton by CDM strategies compared with noncollaboration
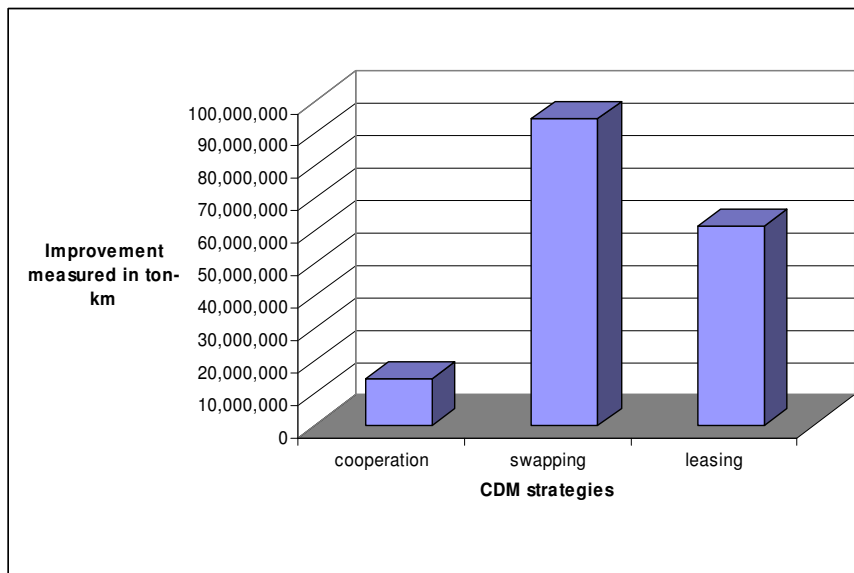
Figure 5-7b Ton-km produced by scenarios with/without train slot cooperation

Findings

Train Slot Cooperation

The experimental results show that the total improvement due to the introduction of carrier collaboration in the form of train slot cooperation among four carriers, as measured in

tons or ton-km transported by newly proposed rail-based IM services, are on the order of two and five percent, respectively. That is, increases of 25,000 tons and 15,000,000 ton-km were predicted along the newly proposed services as a consequence of permitting train slot cooperation between various carriers. This increase was noted primarily for the T3 and T4 services. Little is indicated to change on T1 and T2 services. This can be explained by the significant overlap in T1 and T2 services, permitting shippers to choose the best of the two routes for their purposes and existing slack in their current timetables. With greater usage of T1 and T2 services, greater benefit could be gained from collaboration.

Train Slot Swapping

Significant gains (on the order of 24% and 40% in terms of tons or ton-km, respectively) are predicted where the carriers jointly operate train slots on the service routes, i.e. where train slot swapping is permitted. This strategy appears to outperform other proposed CDM strategies, resulting in the greatest increase in market share for the IM rail freight services. This superior performance may be due to certain characteristics of the proposed services and OD demand within the region. For example, most shipments travel relatively short distances on the IM network (on the order of three zone lengths). With short travel distances, the probability of transferring between services is likely to be small. It is expected that if average travel distances were to increase, the relative performance of the train slot cooperation strategy would improve. While the train slot cooperation strategy is tested assuming the operation of four carriers along four service routes and the train slot swapping strategy is tested assuming the operation of only two carriers, such conclusions can be drawn, because collaboration among four, instead of two, carriers can lead to greater opportunities for collaboration and, therefore, is advantageous.

Note that the improvements due to the train slot swapping strategy are found primarily along the T2, T3 and T4 routes. This appears to be the result of the fact that most shipments are carried by T2 instead of T1; thus, better connections will exist for transferring to T3 or T4 from T2. In addition, the majority of shipments employ routes in the Czech Republic, Austria, Slovakia and Hungary, where several services are offered. Transfers to other service routes and border crossings are required in these regions even for short travel distances. Since service schedules (timetables) offered with train slot swapping have greater frequencies than those offered with train slot cooperation, train slot swapping outperforms train slot cooperation.

Train Space Leasing

Considerable increase (on the order of 15% in tons and 23% in ton-km) in flows along the proposed services is predicted where the train space leasing strategy is applied. While large, this increase is significantly smaller than the increase predicted for the train slot swapping strategy. This may be due to the fact that only one carrier is permitted to lease some subset of train slots from other carriers. If additional swapping options were permitted (e.g. greater percentage of a carrier's train slots could be swapped or multiple carriers were permitted to swap their train slots), improvement in the performance of this strategy would be expected.

5.7 Conclusions

Three CDM strategies are proposed (train slot cooperation, train space leasing and train slot swapping) for operating a multi-carrier rail-based IM freight transport system. The strategies were assessed through a carrier collaboration simulation-assignment framework to

manage collaboratively competing demands for the use of the infrastructure. Experiments were run to assess the potential impact of employing such strategies within proposed services along the REORIENT network, a real-world international, rail-based IM freight transport network. Results of these experiments indicate that the proposed strategies are expected to result in significant improvements in terms of shipments that are attracted to the proposed services. The best performing CDM strategy (the train slot swapping strategy) led to a more than 40% increase in terms of ton-km attracted to the services.

Like other CDM strategies, the proposed strategies result in a win-win situation for all parties. In addition to attracting more demand, cost savings in terms of rolling stock and labor and reduced shipment delays are achieved.

More sophisticated collaborative mechanisms can be proposed and assessed. For example, three or more carriers might jointly operate separate portions of a route, where they might swap train slots. There may be a limit on the number of swaps that is permitted between any pair of carriers. Train capacity can be leased to more than one carrier. Additionally, these experiments included only those scenarios where collaboration is permitted among all carriers on any route. However, it may be the case that only a portion of the carriers may enter into collaboration agreements along a given route. Assessment of the potential of these and other more advanced CDM strategies would require further investigation.

Chapter 6   Track Capacity Allocation among Multiple Carriers Competing in Spot Markets

6.1 Introduction

Shippers seeking transport for one-off loads (e.g. containers) or other cargo arriving on the freight transport market at irregular points in time, such as might originate from companies with highly irregular production and shipment schedules, will often place the shipments on the spot market. Such shipments will be carried by either truck or rail and it is typical that the shipments will be time-sensitive. This chapter focuses on the handling of these one-off loads by rail. To accommodate these shipments upon their arrival on the spot market, and the rail carriers that will transport them, the carriers' trains are scheduled based on residual track capacity. In rail networks, where excess track capacity is abundant, inefficiencies in constructing slots for new train lines as required can be tolerated. Where such excess track capacity is scarce, as may be the case in highly congested systems, efficiency and flexibility in residual capacity utilization is critical if desired service levels are to be achieved. The development of a framework for issuing track access rights to multiple carriers seeking to run additional trains on an as-needed basis to accommodate shipments that arise on the spot market is needed. To serve this sector of the freight transport business, such a framework must be transparent and must lead to an efficient and, in some parts of the world, socially acceptable allocation of track capacity.

This work was motivated by the desire to create an internationally integrated IM freight transport system in a particular region of Europe, where at least one carrier exists in every country with the exclusive business of national rail transport, as described in Chapter 2. In most countries in this region, ownership of track access rights is retained by national

government agencies. Infrastructure ownership and train operations are vertically separated in this region of the world. See Affuso (2003) for related discussion pertaining to such vertical separation in Italy. Carriers, therefore, must apply to the infrastructure managers for track access rights to operate trains in a given rail corridor. Typical objectives of the infrastructure manager in allocating track capacity are to create economic efficiency and greater competition and to meet social obligations of provided services. Rail operations differ in the United States (U.S.), where infrastructure management and train operations are vertically integrated. In the U.S., the objective used in track capacity allocation is to maximize profit. An additional difference between Europe and the U.S. in the methods used for track capacity allocation is that in Europe this process is administrative-based; whereas, in the U.S., optimization techniques are employed. To further illustrate the process employed in Europe, one can consider national network statements that are supplied by member states. In the Swedish network statement (2006), for example, it is stated that requested track access rights for trains are prioritized and conflicts are resolved based on rankings. Resolution of unresolved conflicts is attained through discussion. An automated approach that can respond quickly to carrier requests for track access rights as they arise and that exploits optimization-based techniques can lead to more efficient use of available capacity and greater transparency. Moreover, such an approach can permit quick construction of a train slot, i.e. a combination of segments of track-time pairs that form a route, in which a train can be operated and quick resolution of conflicting requests to allow rail carriers to compete in the freight spot market.

Fig. 6-1 illustrates the process of allocating residual track capacity to trains carrying shipments arising on the spot market. A multi-period time-scale, where each period represents one week, is used. As is shown in the figure, carriers compete for shipments that

arrive unexpectedly within a period. Once allotted to carriers, the carriers assign their shipments to trains based on shipment origins and destinations and preferred arrival and departure times. Each carrier applies to the authority controlling the track access rights for track capacity along desired routes to accommodate the newly composed trains. Since more than one carrier may request to use the same portion of track at overlapping times, there may be competition between carriers for track capacity. The authority issues track capacity to the competing carriers based on a given set of objectives.
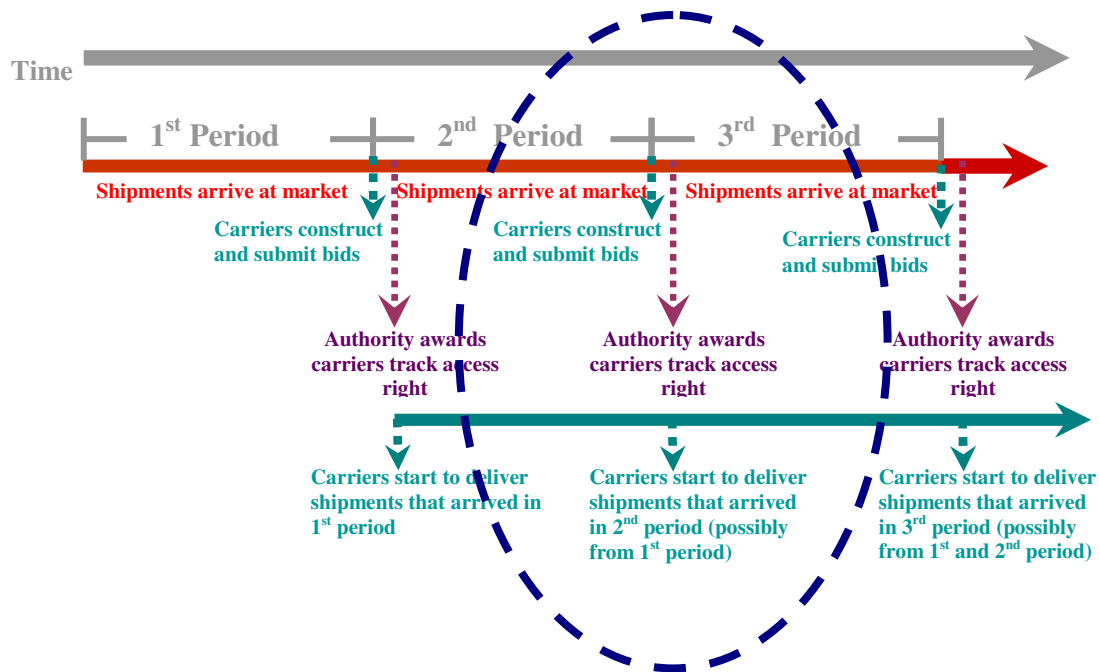
Figure 6-1 Issuance of residual track capacity to trains carrying shipments arising on the spot market

In this chapter, the problem of allocating residual track capacity among multiple carriers to facilitate the delivery by train of shipments arising on the spot market is addressed. A combinatorial auction (CA)-based track capacity allocation framework is proposed that allows carriers to fully express their preferences, enables transparency in the assignment of

track access rights to carriers, permits more efficient utilization of the track capacity, and seeks to maximize the number of shipments that can be delivered.

The framework consists of two main components as illustrated in Fig. 6-2: bid construction and bid allocation (i.e. the selection of bids to which track access rights will be awarded). To begin the auction process, carriers (i.e. the bidders) construct their bids. Carriers make a package bid that contains at least one train slot. After receiving bids from carriers, an authority that retains ownership of the track access rights or represents such an owner (i.e. the auctioneer) awards the track access rights to selected carriers through the bid allocation process. The objective of the bid allocation process is to issue track access rights to multiple carriers such that select social benefits are maximized, i.e. carrier competition and number of shipments delivered are maximized. One might choose an alternative objective of maximizing profit, where track access rights to a given rail network are provided by a private infrastructure owner operating its own fleet, as in the U.S.
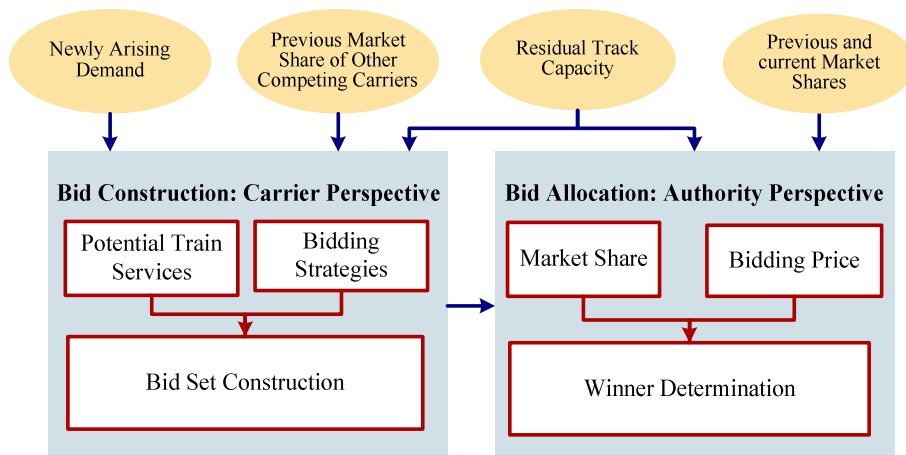
Figure 6-2 Combinatorial auction-based track capacity allocation framework

Two bid set construction approaches are proposed herein in which carriers form their bids: all-or-nothing and train slot-based. In both approaches, the carrier is assumed to have knowledge of its shipments and residual track capacity that can be purchased. The all-or-nothing bid set construction approach will lead to an XOR bid, consisting of atomic bids connected by XOR bidding language. Each atomic bid consists of requests for train slots that together can accommodate trains that can meet shipment requirements. If any single atomic bid is obtained by solving a winner determination problem (WDP), all proposed trains can be operated and if no bid is obtained, no trains will be operated.

Unlike the all-or-nothing bid set construction approach, where each atomic bid contains requests for sufficient train slots to accommodate all trains that the carrier wishes to operate, in the train slot-based bid set construction approach, sets of atomic bids are constructed such that if one bid from each set were obtained, a set of train slots would be issued to accommodate all trains that the carrier wishes to operate. Thus, multiple bids can be obtained (at most one from each set) and if no bid is obtained from a given set, only a fraction of the carrier's desired trains can be operated. The train slot-based bid set construction approach will lead to an OR-of-XOR bid.

Regardless of the approach used in constructing bids, all bids are evaluated through a WDP. A multi-carrier residual track capacity allocation problem is formulated for this purpose. The WDP seeks to maximize social benefits. A multidimensional knapsack problem formulation is proposed. Solution of the WDP provides the optimal winning combination of segments of track-time pairs for the given set of bids.

Two existing auction-based approaches have been proposed in the literature for track capacity allocation. Both approaches use single-unit auction methods, preventing the carriers

from expressing their preferences for desired combinations of train slots. These works are reviewed in section 6.2. Assumptions made in developing the combinatorial-based track capacity allocation framework proposed in this chapter are provided in section 6.3. In section 6.4, participants, including an auctioneer, seller and bidders are identified and the bidding languages employed within the proposed framework are discussed. The all-or-nothing and train slot-based bid set construction approaches are presented in sections 6.5 and 6.6, respectively. The auctioneer will award the carriers track access rights through solving the WDP that is formulated in section 6.7 Findings from experiments conducted to compare the proposed bid set construction approaches under varying levels of available track capacity, demand level, and carrier market share are presented in section 6.8. In section 6.9, conclusions and future considerations are discussed.

6.2 Existing Auction Approaches for Track Capacity Allocation

Recently, a number of works (Affuso, 2003; Gibson, 2003; Newbery, 2003) have debated the feasibility of distributing track access rights to carriers through the use of auctions and two auction-based approaches have been proposed for this purpose (Brewer and Plott, 1996; Nilsson, 1999).

Brewer and Plott (1996) proposed a game-theoretic approach to address the freight train scheduling problem. A binary conflict ascending price mechanism, a decentralized mechanism based on a first-price auction, was proposed. Through numerical experiments, it was shown that this mechanism can be used for capacity allocation. In their experiments, ten study participants represented ten bidders, each of whom bid on train paths along preferred routes. If two or more bidders sought to operate a train along all or part of the same path at overlapping times, the train slot with highest price was awarded so that the train could be

operated on that path at the desired time. In the experiments, nine feasible paths were available on a single track line connecting three stations. All of the paths had the same origin and destination with different departure and arrival times.

Nilsson (1999) developed a second price iterative auction-based procedure to address the train scheduling problem. Since in a second price auction it is beneficial for bidders to truthfully report their valuations, a better allocation could be generated compared with using a first-price auction (as proposed in Brewer and Plott, 1996). Nisson considered an experiment involving ten bidding rounds with three train carriers and nine paths. Before the auction began, path arrival and departure times were adjusted to resolve arrival and departure time conflicts. Train carriers were permitted to bid on their top two preferred paths of all available paths.

While the two proposed track capacity allocation approaches of Brewer and Plott (1996) and Nilsson (1999) provide a basic foundation for designing game-theoretic mechanisms for addressing the track capacity allocation problem, the applications were overly simplistic. These problems consider only one train line containing up to three stations and a single track employed in only one direction. Few conflicts exist in such simple networks. Moreover, a single item (e.g single path or single origin-destination (O-D) pair) bidding process as implemented in these works ignores path and arc interdependencies, as well as complementary attributes of the bidding items. Findings from these works do not extend to more complex rail networks and, therefore, these approaches are limited in their ability to address real-world scheduling problems. Combinatorial auctions (CAs) may be more promising.

CAs are auctions in which bidders can place bids on combinations of items called "packages." CAs have been employed in a variety of industries, including truckload transportation, bus routing, industrial procurement, airport arrival and departure slots arrangement, and allocation of radio spectrum for wireless communications services (Hogan, 1992; Brunekreeft, 2001; Jain, 2001; Melody, 2001; Klemperer, 2002; Milgrom, 2004; Plott and Salmon, 2004; Bourbeau et al., 2005; Elmaghraby, 2005; Strandenes et al., 2005; Triki et al., 2005; Cramton et al, 2006). The advantage of CAs is that the bidders can more fully express their preferences, leading to improved economic efficiency (allocating the items to those who value them the most) and greater auction revenues. CAs are of particular relevance when items are complementary (Xia, et al., 2004; Xia, et al., 2005), i.e. when a set of items has greater utility than the sum of the utilities of the individual items.

In the problem of train scheduling, complementarities exist among arcs, because access rights of track segments exhibit strong interdependencies. The utility of serving a set of track segments, which together form a train slot, by one carrier may be greater than or equal to the utility of serving them by separate carriers. The utility of serving only a portion of a train slot may be zero, because utility can only be gained if the track segments allocated to a carrier can be combined to create a path between a desired O-D pair. Combined operations by one carrier may be more cost-effective than the separate operation by multiple carriers. In addition, by applying CAs in track capacity allocation, carriers (bidders) can fully express their preference for combinations of track capacities. In this work, a CA-based track capacity allocation framework is proposed for addressing the track capacity allocation problem. Unlike in single-unit auctions, where bids can be simply constructed and the winner is determined based solely on the value of the highest bid, in CAs, techniques for

constructing bids are required and a WDP based on both bidding prices and availability of resources must be formulated. Methods for constructing bids and a binary, linear multidimensional knapsack problem with side-constraints formulation of the WDP are described in following sections.

6.3 Problem Assumptions

Assumptions made in developing the CA-based track capacity allocation framework and relevant solution techniques are provided in this section.

1. *An auction involving track capacity allocation will only be held after customer commitments for regular use of services is scheduled.*

   It is known that the one-off loads arise at irregular points in time and, thus, will likely be shipped only once. In addition, any profit gained from accommodating these one-off loads is less significant than that gained from handling regular commitments. Thus, arrangement of train slots for regular commitments has higher priority over similar arrangement for the one-off loads.

2. *Details of one-off loads are known by all carriers.*

   Before the auction begins, all carriers involved in the auction are given information on the one-off loads' preferred delivery times and O-D pairs. Each carrier is not knowledgeable of competing carriers' customers. However, information pertaining to shipments made by all carriers in prior time periods is open to the public.

3. *Carriers will operate all awarded train slots.*

   To avoid negative competition among carriers, it is assumed that carriers will operate all train slots purchased through an auction.

4. *Carriers are assumed to have private values and behave rationally.*

Carriers will bid their true valuation for their desired train slots.

5. Each bidding package must be a directed, elementary path.

6.4 Bidding Languages in the Proposed CA-based Track Capacity Allocation Framework

Let A and T denote a set of arcs within a rail network and a set of consecutive time intervals, respectively. Track segment-time pairs (i.e. train slot), denoted a^t, where $a \in A$ and $t \in T$, are auctioned. When bidding, carriers must specify their desired routes, as well as departure and arrival times along these routes. The carriers bid on a subset of the network arcs for particular time periods that together form train slots.

Definition 1 Q is defined as a set of train slots specified by the carrier. Let $i = \{a^1, a^2, ..., a^z\}$ denote a set of z track segment-time pairs forming train slot $i \in Q$.

Definition 2 To deliver a set of shipments, two train slots i and j, where $i, j \in Q$ and $i \neq j$, are substitutable if and only if they can be employed to transport the same shipments.

Definition 3 Two train slots i and j, where $i, j \in Q$ and $i \neq j$, are complementary if and only if they can only be employed to transport different shipments.

In this work, package bids containing at least one train slot are submitted by carriers. Bidding language introduced in Nisan (2000) is employed herein to represent package bids in CAs. Specifically, in this work, atomic bids, XOR bids, OR bids, and OR-of-XOR bids are used by carriers to express their preferences for train slots. These concepts are described next in the context of auctioning of train slots.

Atomic Bids

Suppose a carrier has k atomic bids, $\{q_1, p_1\}, \{q_2, p_2\}, \{q_j, p_j\}..., $ or $\{q_k, p_k\}$. An atomic bid q_j is a set of one or more train slots and a price p_j that the carrier is willing to pay for this set of train slots. All train slots in a single atomic bid must have identical O-D pairs.

XOR Bids

XOR bidding language can represent situations in which each carrier can submit an arbitrary number, k, of atomic bids. Assuming for simplicity that each atomic bid contains only one train slot, a typical XOR bid will be expressed as follows:

$$\{q_1, p_1\}XOR\{q_2, p_2\}...XOR\{q_k, p_k\}, \text{ for } q_1, q_2, ..., q_k \text{ in } Q.$$

This expression implies that the carrier is willing to obtain at most one of the proposed atomic bids, $\{q_1, p_1\}, \{q_2, p_2\}, ..., $ or $\{q_k, p_k\}$. XOR bids can represent all valuations, including complementary or substitutable train slots.

OR Bids

Similar to XOR bidding language, each carrier can submit an arbitrary number of atomic bids. A typical OR bid, assuming for simplicity that each atomic bid includes only a single train slot, can be expressed as follows.

$$\{q_1, p_1\}OR\{q_2, p_2\}...OR\{q_k, p_k\}, \text{ for } q_1, q_2, ..., q_k \text{ in } Q.$$

The OR bidding language implies that the carrier is willing to obtain any number of atomic bids; substitutable (not complementary) train slots can be represented.

OR-of-XOR Bids

OR and XOR bidding languages can be combined to succinctly represent many desirable valuations. In the OR-of-XOR bidding language, a bidder can submit an arbitrary number, m, of XOR bids connected by ORs. Each XOR bid can consist of any number of atomic bids. However, the atomic bids among m XOR bids cannot be the same. In such an OR-of-XOR

bid, the carrier is willing to obtain any number of atomic bids; however, at most one bid from each OR clause will be assigned. This reflects the fact that the atomic bids of each OR clause are designed to handle the same shipments.

6.5 Bid Set Construction Approaches

Two bid set construction approaches are proposed in this section in which carriers form their bids: all-or-nothing and train slot-based. The all-or-nothing bid set construction approach will lead to an XOR bid, consisting of atomic bids connected by XOR bidding language. If any single atomic bid is obtained by solving the WDP, all proposed trains can be operated and if no bid is obtained, no trains will be operated. The train slot-based bid set construction approach leads to an OR-of-XOR bid. Thus, multiple bids can be obtained (at most one from each set) and if no bid is obtained from a given set, only a fraction of the carrier's desired trains can be operated.

6.5.1 All-or-nothing bid set construction approach

The all-or-nothing bid set construction approach that relies on the XOR bidding language is proposed for use in constructing a bid set given the carrier's objectives of transporting a given set of one-off loads arising on the spot market that must be delivered within a predetermined time window. It is assumed that the carrier is knowledgeable of the residual track capacity and can estimate the minimum number of train slots required to accommodate all shipments. Each atomic bid, consisting of a minimal number of train slots (thus, minimizing operating costs) that simultaneously meet customer demands, is constructed such that if any is obtained, the carrier will be able to transport all of the shipments within delivery time windows.

The all-or-nothing bid set construction approach seeks a collection of competitive atomic bids. At most $2^{|Q|} - 1$ unique atomic bids can be generated through this approach. However, many of these bids will be insufficient or inefficient for the number of trains required to accommodate all shipments. To reduce the effort required to determine the set of atomic bids with the minimal number of train slots necessary to accommodate the shipments, an estimation technique is employed to calculate the minimum number of required train slots for the given shipments. An exhaustive search approach can then be applied to screen possible atomic bids with the minimal or near-minimal number of train slots. The all-or-nothing bid set construction approach is an iterative approach as described next. Text in bold will be discussed in detail in following subsections.

All-or-nothing bid set construction approach

Step 1. **Minimum required train slots estimation.** Compute the minimum number, η, of train slots according to the number of shipments.

Step 2. Enumeration of potential atomic bids. Enumerate all possible train slot combinations, each consisting of m train slots. Let θ represent a given train slot combination.

Step 3. **Feasible atomic bid identification.** Examine each train slot combination, θ, for feasibility. If θ is feasible for use as an atomic bid, calculate the price of the atomic bid. The price will be a function of the total travel time required to complete the trip from origin to destination.

Step 4. Termination check. If at least one identified combination, θ, is feasible, terminate the procedure. Otherwise, let $\eta = \eta + 1$ and return to *Step 2*.

This procedure terminates with a set of atomic bids that will be submitted to the WDP. Any of the atomic bids, if obtained, can deliver all of the shipments.

Minimum required train slots estimation

The problem of calculating the minimum number of required train slots, i.e. the minimum number of trains that must be operated, as input for the process of generating the atomic bids for a given set of shipments (*Step1*), is described in this subsection. This problem is formulated as a linear, binary integer program and a quick estimation method is presented for its solution.

Notation

N : set of shipments

Q : set of train slots

u : maximum number of shipments that can be transported in any given train slot

$$x_q = \begin{cases} 1, & \text{if train slot } q \in Q \text{ is employed;} \\ 0, & \text{otherwise.} \end{cases}$$

$$\delta_{nq} = \begin{cases} 1, \text{if shipment } n \in N \text{ can be delivered within train slot } q \in Q \text{ in a given time window;} \\ 0, \text{otherwise.} \end{cases}$$

Mathematical formulation

The problem of estimating the minimum number of required train slots is formulated in (1) through (4) as a set packing problem with side constraints.

$$\text{Min } z(x) = \sum_{q \in Q} x_q \tag{1}$$

subject to

$$\sum_{q \in Q} \delta_{nq} x_q = 1, \qquad \forall n \in N \tag{2}$$

$$\sum_{n \in N} \delta_{nq} x_q \leq u, \qquad \forall q \in Q \tag{3}$$

$$x_q \in B, \qquad \forall q \in Q \tag{4}$$

Binary decision variable, x_q, represents whether or not the train slot, q, associated with a certain period of time can be used to transport a subset of the shipments. The objective (equation (1)) seeks to minimize the number of trains required to transport the given set of shipments between their O-D pairs. Constraints (2) ensure that each shipment, $n \in N$, is assigned to one and only one train slot, $q \in Q$. Constraints (3) are train capacity constraints associated with the trains that will be operated in the train slots $q \in Q$. It is assumed that train capacity limitations are identical on all trains. Binary integrality requirements are given in Constraints (4).

The outcome of the formulation ((1)-(4)) is a set of train slots that together can be employed to transport all shipments such that the number of train slots required to accommodate the shipments is minimized. The set packing problem is known to be NP-hard, and thus, one can expect the number of decision variables and constraints in the formulation to be quite large for a real-world problem instance. Solution to optimality would be formidable. Instead of obtaining the number of required train slots by solving the model to optimality, the number of required train slots, η, can be quickly estimated through the use of a simple approximation method shown in (5). In this method, shipments are permitted to be delivered by trains whose O-D pairs might not be the same as that of the shipment. Thus, equation (5) produces a lower bound on the minimum required train slots.

$$\eta = \lceil \text{total number of the shipments} / u \rceil \tag{5}$$

Feasible atomic bid identification

Whether not a given set of train slots is feasible, i.e. whether or not shipment assignment constraints (2) and train capacity constraints (3) are met, must be determined. An

exhaustive search procedure is applied to examine feasibility by enumerating all the possible combinations (i.e. which shipment is transported by which train). Once a feasible solution is found, an atomic bid containing the associated set of train slots is created.

6.5.2 Train slot-based bid set construction approach

In this section, a train slot-based bid set construction approach that employs OR-of-XOR bidding language is proposed for constructing a bid set that enables the carriers to fully express their preferences. The construction procedure permits quick generation of a manageable number of competitive bid combinations. Competitive and manageable bid sets will require fewer decision variables (i.e. atomic bids) in the WDP as compared with bid sets constructed through the use of XOR bidding language; thus, decreasing the computational burden required for its solution.

The proposed train slot-based bid set construction approach is comprised of carrier and revised carrier models described in section 6.5.2.1 and 6.5.2.2, respectively. The carrier model is a binary integral program that seeks an optimal train timetable based on a given set of shipments and assigns the shipments to the train timetable. The train timetable is created such that operational costs and delays in delivery from scheduled arrival times (i.e. delay) are minimized. For each set of shipments with the same O-D pair and desired arrival time at the destination, an atomic bid is constructed from the train timetable. That is, the requested train slot(s) in the atomic bid are equivalent to the train slot(s) allocated in the optimal train timetable for the given O-D pair and given arrival time. The carrier model produces the most desirable combination of atomic bids (one atomic bid for each OR clause) such that together the train slots associated with the bids are sufficient to accommodate all shipments. If these bids are awarded, the total operational cost of transporting the shipments and total incurred

shipment delay will be minimized. Once employed within an atomic bid, a train slot is removed from future consideration.

It is possible that more than one carrier is interested in the same portion of the track capacity. To increase the probability of obtaining train slots that can be used to transport the shipments, a number of backup bids are needed. Therefore, the bid set is augmented by iteratively solving the revised carrier model. That is, the revised carrier model is employed to produce the remaining set of atomic bids for each OR clause of the OR-of-XOR bid. The revised carrier model is identical to the carrier model with additional shipment delivery restriction constraints. These constraints ensure that shipments carried by different trains in the timetable produced by the carrier model will also be carried on different trains in the resulting solution of the revised carrier model. This ensures that the atomic bids in the same OR clause can be used to deliver an identical set of shipments. The revised carrier model is employed repetitively until a predefined bid size is reached. The procedure terminates when the number of atomic bids in each OR clause exceeds the number of trains operated by competing carriers in the prior time period. It is assumed that the competing carriers will bid on at least as many train slots as were awarded in prior time periods. This train slot-based bid set construction approach is illustrated in Fig. 6-3.
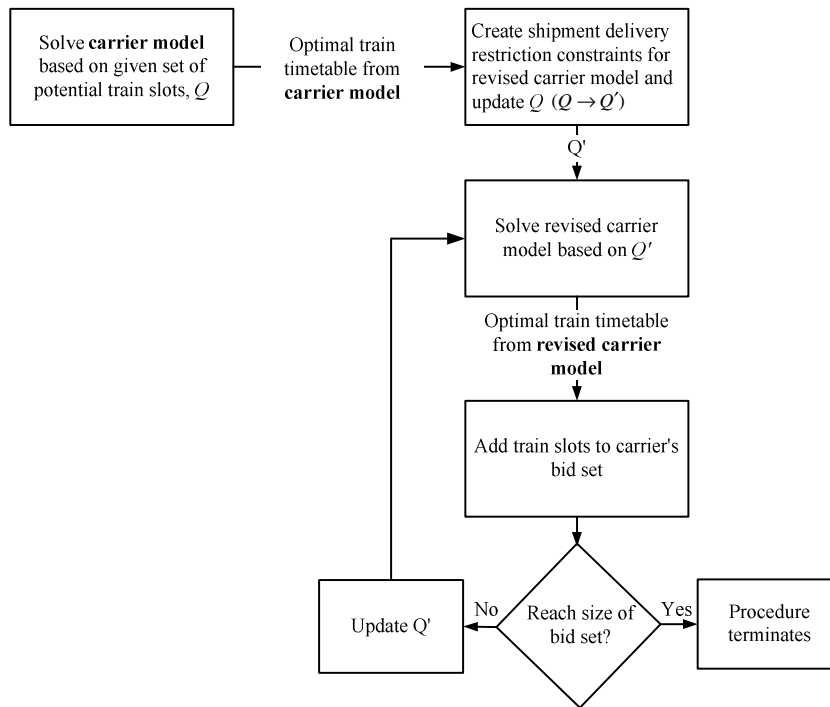
Figure 6-3 Train slot-based bid set construction approach

Fig. 6-4 illustrates the structure of the bid set. The first bid in each OR clause is constructed from the solution to the carrier model. A solid rectangle is placed around these bids in the figure. An OR bid could be constructed entirely from these atomic bids. If all of the atomic bids in such an OR bid were obtained, all proposed trains could be simultaneously operated and all desired shipments would be delivered. A rectangle with dashed lines is used to indicate each set of bids produced from the optimal solution of a run of the revised carrier model. The atomic bids created from each solution of the revised carrier model can form a single OR bid. The resulting OR-of-XOR bid is composed, therefore, of OR clauses from which at most one bid will be chosen. The combination of bids from each OR clause (one from each such clause) form a set of train slots that will support the transport of all desired shipments.

$$\text{Number of XOR bids}^{(1)} \left\{ \begin{array}{l} \text{OR} \\ \text{OR} \\ \vdots \\ \text{OR} \end{array} \right.$$

$\{q_1,p_1\}$	XOR	$\{q_2,p_2\}$	...	XOR	$\{q_k,p_k\}$
$\{q_1,p_1\}$	XOR	$\{q_2,p_2\}$	...	XOR	$\{q_k,p_k\}$
$\{q_1,p_1\}$	XOR	$\{q_2,p_2\}$	...	XOR	$\{q_k,p_k\}$

Size of each XOR bids$^{(2)}$
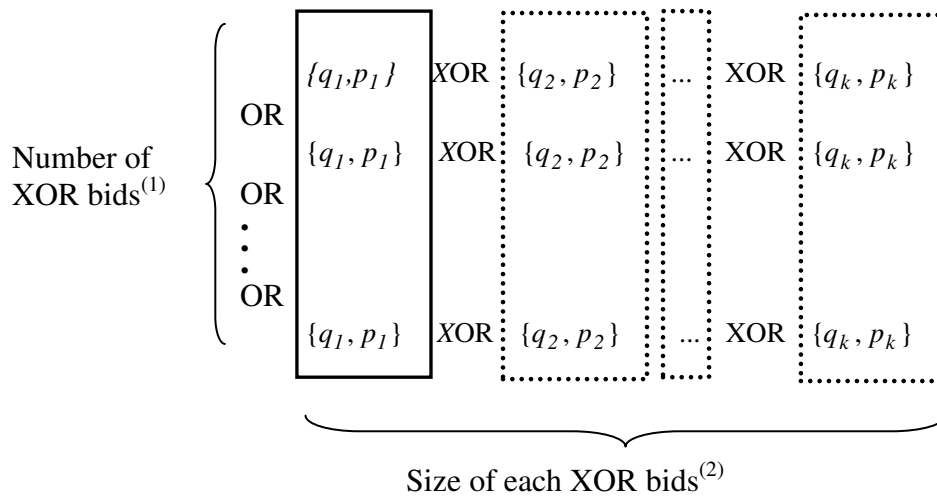
⬜ : Atomic bids produced by carrier model

⬜ : Atomic bids augmented by revised carrier model

(1): Depends on solution of the carrier model
(2): Depends on prior market share of competing carriers for each O-D pair

Figure 6-4 An OR-of-XOR bid based on the train slot-based bid set construction approach

It is important to note that while the bids in each OR clause are associated with a unique O-D pair, train slots that are awarded across the O-D pairs employ competing resources, i.e. track capacity within a single network. That is, there are interdependencies in track capacity requests among bids associated with the varying O-D pairs. Thus, if the OR-of-XOR bid were broken into separate single-unit auctions of XOR bids (one for each unique O-D pair), the order in which the auctions are conducted would affect the outcome.

In the all-or-nothing bid set construction approach described in Section 6.5.1, the number of atomic bids in a bid set is directly proportional to the number of possible train slots that can be constructed given the residual track capacity. In the bid set construction approach, however, the size of the bid set is not directly influenced by this number of potential train slots. Note that given the number of atomic bids connected by XOR in an OR

clause, r_i, and the number of OR clauses in a bid set, m, the size of the bid set can be computed by $\sum_{i=1}^{m} r_i$. This number, therefore, is a function of the number of train slots required by the optimal solution to the carrier model and knowledge of the train slots operated by competing carriers in the prior time period.

If instead of employing the revised carrier model to select desirable backup bids for each OR clause of the OR-of-XOR bid one wished to consider all feasible combinations of train slots, a more complicated bidding language would be required to express the carrier's preference. Moreover, bid selection via solution of the WDP will require exponentially increasing effort with increasing problem size. Let Q define a set of potential train slots specified by the carrier based on residual track capacity within the network and ξ_i denote the set of the potential train slots that can be used in creating each OR clause i. Given m OR clauses, assume that $\{\xi_i \cap \xi_{i+1}...\cap \xi_m\} = \varnothing$ and $\{\xi_i \cup \xi_{i+1}...\cup \xi_m\} = Q$. If each atomic bid is assumed to contain only one train slot, the OR clause i can be selected from $C_{r_i}^{\xi_i}$ combinations. For a given number of OR clauses in the bid set, m, the total number of combinations in the OR-of-XOR bid set is $\sum_{i=1}^{m} C_{r_i}^{\xi_i}$. Solution of the resulting WDP would require enormous computational effort. The approach proposed herein based on the carrier and revised carriers models, by contrast, can generate competitive bids within a manageable bid size, i.e. $\sum_{i=1}^{m} r_i$ as compared with $\sum_{i=1}^{m} C_{r_i}^{\xi_i}$.

6.5.2.1 Carrier model

The carrier model seeks a subset of potential train slots in which to operate trains that can transport a desired set of one-off loads (i.e. shipments) with given O-D pairs and desired arrival times. Solution of the model results in a train timetable. The resulting timetable is used to construct bids for inclusion in an OR bid. Notation used in developing this formulation is given next.

Notation

A: set of arcs

T: set of time intervals

N: set of shipments

Q: set of train slots

W_q: set of shipments delivered using train slot $q \in Q$

M_q: set of shipments whose origins and destinations are both one of the intermediate

  terminals of train slot $q \in Q$, $M_q \subseteq N$

f_{nq}: cost of train slot $q \in Q$ used to carry shipment $n \in N$

e_{nq}: delay of shipment $n \in N$ transported in train slot $q \in Q$

u: maximum number of shipments transported in a single train slot

$x_{nq} = \begin{cases} 1, & \text{if train slot } q \text{ is employed to deliver shipment } n \in N; \\ 0, & \text{otherwise.} \end{cases}$

$\delta_{nq} = \begin{cases} 1, & \text{if O-D pair of shipment } n \in N \text{ is the same as that of train slot } q \in Q; \\ 0, & \text{otherwise.} \end{cases}$

$\lambda_{qa^t} = \begin{cases} 1, & \text{if train slot } q \text{ employs arc } a \in A \text{ in a time interval } t \in T; \\ 0, & \text{otherwise.} \end{cases}$

Mathematical formulation

The carrier model is given in (6) through (11).

$$\text{Min } z(x) = \sum_{q \in Q} \sum_{n \in N} f_{nq} x_{nq} + e_{nq} x_{nq} \tag{6}$$

subject to

$$\sum_{q \in Q} \delta_{nq} x_{nq} = 1, \qquad \forall n \in N \tag{7}$$

$$\sum_{n \in N} \delta_{nq} x_{nq} \leq u, \qquad \forall q \in Q \tag{8}$$

$$\sum_{\substack{n \in N \\ n \neq m}} \delta_{nq} x_{nq} \geq x_{mq}, \qquad \forall q \in Q, \; \forall m \in M_q \tag{9}$$

$$\sum_{q \in Q} \sum_{n \in N} \frac{\lambda_{qa^t} x_{nq}}{\sum_{n \in W_q} x_{nq}} \leq 1, \qquad \forall a \in A, \; \forall t \in T \tag{10}$$

$$x_{nq} \in B, \qquad \forall n \in N, \; \forall q \in Q \tag{11}$$

The objective given in equation (6) seeks to minimize the total incurred shipment delay and total operational cost of transporting the desired shipments within a network given a set of potential train slots. Constraints (7) ensure that each shipment is assigned to only one train slot. Train capacity constraints for each train slot and shipment bundle constraints are given in Constraints (8) and (9), respectively. The bundle constraints ensure that if a train slot is employed, it must carry at least one shipment with identical origin and destination to that of the train slot. Constraints (10) ensure that track $a \in A$ at time $t \in T$ is allocated to at most one atomic bid. Binary integrality requirements are given in Constraints (11).

The outcome of the carrier model is a set of train slots that can be employed to transport a set of desired shipments such that operational costs and delays in delivery are

minimized. The outcome is treated as a single OR bid, where each constituent atomic bid contains at least one train slot. The resulting OR bid is the most desirable bid combination for the carrier.

Ideally, the carrier will obtain all of the atomic bids in this single OR bid. However, since there may be more than one carrier interested in the same track capacity, it is not guaranteed that each of the atomic bids in such a bid set will be obtained. Therefore, backup candidate atomic bids are needed for each atomic bid in the single OR bid and an OR-of-XOR bidding language is applied to express the carrier's preference for given track capacity. A set of backup candidate atomic bids for each OR clause of the OR-of-XOR bid can be created through a revised carrier model proposed in the next subsection.

6.5.2.2 Revised carrier model

The revised carrier model is identical to the carrier model with additional shipment delivery restriction constraints as given in constraints (12).

$$x_{iq} + x_{jq} < 2, \qquad\qquad \forall i, j \in S_q,\ i \neq j,\ \forall q \in Q. \qquad\qquad (12)$$

The shipment delivery restriction constraints are derived based on the optimal solution to the carrier model. That is, train slots employed in the OR bid constructed from the carrier model will be eliminated from consideration in solving the revised carrier model. Additionally, these shipment delivery restriction constraints ensure that all shipments transported in the same train slot in the resulting train timetable of the carrier model will be transported together in the solution to the revised carrier model. This guarantees that all atomic bids in each OR clause will be interchangeable.

The revised carrier model formulation is given in (13). Like the carrier model, the outcome of the revised carrier model is a set of train slots that are employed to transport a set of desired shipments such that operational costs and delays in delivery are minimized.

Solution of the revised carrier model can produce a set of backup candidate atomic bids for each OR clause of the OR-of-XOR bid.

$$z'(x) = \min\{\sum_{q \in Q'}\sum_{n \in N} f_{nq}x_{nq} + e_{nq}x_{nq} : (7), (8), (9), (10), (11), (12)\}, \tag{13}$$

where $Q' \subset Q$ is defined as a set of train slots that have not been selected in an atomic bid through the use of the carrier model or revised carrier model..

To avoid the situation where the same train slot is included in XOR bids of more than one OR clause, train slots selected for inclusion in an atomic bid are removed from future consideration. If more than one atomic bid that is obtained contains the same train slot, it will result in significant inefficiencies. Thus, the OR-of-XOR bid set is constructed such that atomic bids in different OR clauses are disjoint.

6.6 Winner Determination Problem

The WDP seeks the optimal atomic bids proposed by the carriers to which train slots will be awarded such that the social benefits of the winning bids are maximized, i.e. carrier competition and the number of shipments that will be transported by trains operated in the awarded train slots are maximized. The WDP is formulated as a multidimensional knapsack problem, solution of which provides a train timetable for a given set of one-off loads to be transported in a given planning period.

The proposed formulation relies on a train slot representation of limited residual track capacity. It is assumed that all proposed bid sets have either an XOR or an OR-of-XOR bid set structure, as developed from the all-or-nothing and train slot-based bid set construction approaches, respectively. Carriers that obtain track access rights for the purpose of

transporting one-off loads will only retain these rights during the considered planning period.

Notation used in developing this formulation is given next.

<u>Notation</u>

A : set of arcs

T : set of time intervals

K : set of carriers

F : set of carriers employing all-or-nothing bid set construction approach

E : set of carriers employing train slot-based bid set construction approach

S^b : set of atomic bids (i.e. train slots) submitted by carrier $b \in K$

$c(s)$: competition factor of atomic bid $s \in S^b$ by carrier b

$v(s)$: bidding price of bid $s \in S^b$

$w(s)$: number of shipments delivered within the train slots in bid $s \in S^b$

ϖ_b : maximum number of atomic bids submitted by carrier $b \in K$ that can be obtained

G^b : set of OR clauses in OR-of-XOR bid set submitted by carrier $b \in E$

$$y_s = \begin{cases} 1, & \text{if bid } s \text{ is awarded by solving WDP;} \\ 0, & \text{otherwise.} \end{cases}$$

$$\delta_{a^t}(s) = \begin{cases} 1, & \text{if train slot in bid } s \text{ employs track } a \text{ at time } t \in T; \\ 0, & \text{otherwise.} \end{cases}$$

$$\lambda_g(s) = \begin{cases} 1, & \text{if bid } s \text{ belongs to OR clause } g \in G^b; \\ 0, & \text{otherwise.} \end{cases}$$

<u>Mathematical formulation</u>

Formulation (WDP) is presented next in equations (14) through (18).

$$\text{Max } WDP = \sum_{b \in K} \sum_{s \in S^b} \frac{1}{c(s)} \frac{1}{v(s)} w(s) y_s \tag{14}$$

subject to

$$\sum_{b \in K} \sum_{i \in I^k} \delta_{a^t}(s) y_s \le 1, \qquad \forall a \in A, \ \forall t \in T \tag{15}$$

$$\sum_{s \in S^b} y_s \le \varpi_b, \qquad \forall b \in K \tag{16}$$

$$\sum_{s \in S^b} \lambda_g(s) y_s \le 1, \qquad \forall g \in G^b, \forall b \in E \tag{17}$$

$$y_s \in B, \qquad \forall s \in S^b, \forall b \in K \tag{18}$$

The objective given in equation (14) seeks to maximize the social benefits, including carrier competition and number of shipments transported within the network. The objective considers the competition factor, $c(s)$, in awarding track access rights. $c(s)$ is computed for each carrier from a weighted average of the carrier's market share for the spot market in the prior time period and desired market share as determined from the number of requested train slots. The greater the competition factor, the less weight that is placed on awarding track access rights to that carrier. Bid price $v(s)$ of carrier b for bid s, $s \in S^b$ and $b \in K$, is assumed to be the train operational cost determined by carrier b. The number of shipments that can be delivered times the inverse of the bid price (i.e. $w(s) \times \frac{1}{v(s)}$) gives the average cost of delivering each shipment in atomic bid s. Thus, track access rights requested in atomic bids with low products of average cost and inverse competition factors are more likely to be awarded.

Constraints (15) ensure that track $a \in A$ at time $t \in T$ can be allocated to at most one atomic bid. Constraints (16) ensure that the number of atomic bids proposed by carrier $b \in K$ from which track access rights are awarded is no greater than a given limit imposed by the structure of the bid set. For a carrier $b \in E$ that proposes an OR-of-XOR bid set, constraints (17) ensure that track access rights from at most one atomic bid of each OR clause, $g \in G^b$, of the OR-of-XOR bid set can be awarded to that carrier. Binary integrality requirements are enforced in constraints (18). When considered in conjunction with constraints (17), one may observe that constraints (16) are unnecessary for bids with the OR-of-XOR bid set structure. Thus, constraints (16) can be rewritten as in (16'), where only atomic bids contained in the XOR bids are considered:

$$\sum_{s \in S^b} y_s \leq \varpi_b, \qquad \forall b \in F \qquad (16')$$

6.7 Results of Numerical Experiments

Results of computational experiments designed to illustrate the feasibility of applying CAs for track capacity allocation to carriers seeking to transport one-off loads arriving on the spot market in a given time period are provided and analyzed. The all-or-nothing and train slot-based bid set construction approaches proposed in Section 6.5 were employed in the experiments and the competitiveness and size of resulting bid sets are compared under varying levels of residual track capacity, prior period market share and number of shipments arriving on the spot market. The competitiveness of the bid sets is determined from solution of the WDP through which train slots are awarded to the carriers.

The bid set construction approaches and WDP were coded in Microsoft Visual C++ 6.0 and CPLEX 10.1 was employed. These codes were run on an Intel Pentium 4 with 3.6

GHz CPU and 2 GB RAM using the Windows XP operating system. Track capacity of a 6-vertex, 8-edge network (Fig. 6-5) with double tracks along each edge is assumed to be allocated among two carriers over a one-week period. Three sets of experiments were run. In the first set, four levels of residual track capacity (20%, 30%, 40%, and 50%) were considered for a given setting of prior period market share (50%) and number (1,000) of shipments. In the second set, either 1,000 or 2,000 shipments were assumed to arrive on the spot market that are of interest to the carriers; market share and residual track capacity were fixed at 50% each. Finally, 10 levels of prior period market share were studied for a fixed number (1,000) of shipments and residual track capacity (50%). For a given combination, shipments (i.e. shipment O-D pairs and desired departure and arrival times) were randomly generated. Five runs, associated with five sets of randomly generated shipments, were made for each combination, results of which are reported individually. One carrier is assumed to employ the all-or-nothing bid set construction approach while the second carrier is assumed to employ the train-slot based bid set construction approach. Residual track capacity is computed based on a train timetable generated for the forward market. This timetable was developed using the initial track capacity allocation technique introduced in Chapter 4 for a randomly designed demand set.
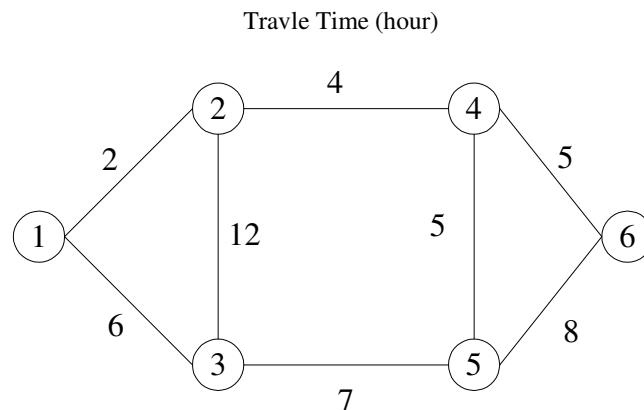
Figure 6-5 Illustrative network

116

Results of the experiments in terms of size of the bid sets that are constructed and average overall shipment delay are provided in Tables 6-1 through 6-3. The number of shipments that can be accommodated based on the awarded track capacity is provided. Note that if residual capacity is too restricted to support the desired train slots requested to accommodate all desired shipments, only a fraction of the shipments will be transported.

Table 6-1 Size of bid sets and resulting average delay given varying levels of residual track capacity

| Track capacity | Carrier A(market share=0.5) | | | Carrier B(market share=0.5) | | |
| | Train slot-based approach | | | All-or-nothing approach | | |
	Number of atomic bids	*Number of shipments delivered	Average shipment delay(hr)	Number of atomic bids	*Number of shipments delivered	Average shipment delay(hr)
50%	161	1,000	0.4	5,570	1,000	19.6
	125	1,000	0.8	6,741	1,000	16.1
	98	1,000	0.0	5,967	1,000	18.3
	176	1,000	0.0	3,161	1,000	15.6
	135	1,000	1.5	2407	1,000	17.0
40%	122	1,000	0.0	2,354	1,000	23.1
	154	1,000	2.7	5,305	1,000	14.0
	156	1,000	0.3	6,224	1,000	17.0
	138	1,000	1.3	3,651	1,000	16.9
	158	1,000	1.7	2,142	1,000	12.0
30%	148	1,000	1.0	1,754	1,000	19.4
	169	1,000	2.3	1,230	1,000	12.1
	135	1,000	4.0	1,591	0	--
	156	1,000	3.7	1,374	1,000	20.4
	167	1,000	0.6	1,424	1,000	16.7
20%	139	827	3.3	958	1,000	25.5
	156	1,000	2.5	205	0	--
	142	1,000	6.1	760	0	--
	127	1,000	5.1	756	0	--
	59	1,000	3.8	1,856	0	--

Table 6-2 Size of bid sets and resulting average delay given varying numbers of shipments that arrive on market

| Number of shipments | Carrier A(market share=0.5) | | | Carrier B(market share=0.5) | | |
| | Train slot-based approach | | | All-all-nothing approach | | |
	Number of atomic bids	Number of shipments delivered	Average shipment delay(hr)	Number of atomic bids	Number of shipments delivered	Average shipment delay (hr)

	161	1,000	0.4	5,570	1,000	19.6
	125	1,000	0.8	6,741	1,000	16.1
1,000	98	1,000	0.0	5,967	1,000	18.3
	176	1,000	0.0	3,161	1,000	15.6
	135	1,000	1.5	2,407	1,000	17.0
	128	2,000	2.9	3,354	0	--
	157	2,000	5.3	5,305	0	--
2,000	169	2,000	6.0	6,224	0	--
	163	2,000	9.4	5,651	0	--
	175	2,000	13.3	2,142	0	--

Table 6-3 Size of bid sets and resulting average delay given varying carriers' market shares

Carriers' previous market shares (Carrier A, Carrier B)	Carrier A			Carrier B		
	Train slot-based approach			All-all-nothing approach		
	Number of atomic bids	Number of shipments delivered	Average shipment delay(hr)	Number of atomic bids	Number of shipments delivered	Average shipment delay(hr)
(0.1, 0.9)	201	1,000	0	1,766	1,000	23.0
	196	1,000	0	1,312	0	--
	184	1,000	0.3	1,396	1,000	23.4
	186	1,000	2.5	4,399	1,000	14.6
	198	1,000	2.3	1,158	0	--
(0.2, 0.8)	195	1,000	0	1,083	1,000	18.9
	178	1,000	1.3	1,504	1,000	18.8
	159	1,000	3.2	3,202	1,000	17.1
	161	1,000	1.7	1,371	1,000	22.4
	175	1,000	1.8	3,161	1,000	18.7
(0.3, 0.7)	185	1,000	3.6	3,466	1,000	17.9
	169	1,000	0.7	2,017	1,000	20.5
	163	1,000	3.2	1,903	1,000	23.3
	178	1,000	1.7	2,202	0	--
	157	1,000	2.0	1,630	0	--
(0.4, 0.6)	142	1,000	0.2	1,656	0	--
	156	1,000	2.3	1,517	1,000	15.4
	143	1,000	1.3	2,399	1,000	18.2
	154	1,000	2.0	1,378	1,000	23.8
	158	1,000	6.4	1,056	0	--
(0.5, 0.5)	148	1,000	1.0	1,754	1,000	19.4
	169	1,000	2.3	1,230	1,000	12.1
	135	1,000	4.0	1,591	0	--
	156	1,000	3.7	1,374	1,000	20.4
	167	1,000	0.6	1,424	1,000	16.7
(0.6, 0.4)	102	853	1.8	3,451	1,000	23.3
	94	925	5.7	3,702	1,000	12.4
	98	1,000	0.4	1,054	1,000	17.9
	110	1,000	0.7	1,071	1,000	18.4
	117	1,000	4.7	1,770	1,000	16.4
(0.7, 0.3)	92	872	2.0	1,732	1,000	16.4
	88	1,000	5.5	5,784	1,000	17.8
	84	1,000	1.6	1,106	1,000	21.7

	92	886	0.4	2,173	1,000	15.4
	91	1,000	2.0	1,279	1,000	9.8
(0.8, 0.2)	72	1,000	9.7	5,652	1,000	19.3
	86	1,000	5.1	2,135	1,000	17.7
	82	828	--	1,308	1,000	14.1
	76	867	--	1,474	1,000	18.0
	69	1,000	1.6	2,352	1,000	15.7
(0.9, 0.1)	31	905	4.4	2,410	1,000	10.6
	56	817	10.8	1,684	1,000	17.4
	54	652	4.8	1,876	1,000	17.9
	66	876	0.7	2,042	1,000	18.2
	52	821	3.9	1,234	1,000	12.8

Results of the runs displayed in Tables 6-1 through 6-3 indicate that employing the train slot-based bid set construction approach, as compared with the all-or-nothing bid set construction approach, results in fewer atomic bids required to create viable bid sets, a greater number of shipments that can be accommodated by awarded train slots, and lower average shipment delays. In addition, the number of atomic bids in each bid set constructed though the use of the train slot-based bid set construction approach is fixed and, thus, does not fluctuate between experiments. Bid sets developed using the all-or-nothing bid set construction approach vary in size and may grow exponentially with increase in residual track capacity. That is, the number of atomic bids in the OR-of-XOR bid set developed by the train slot-based bid set construction approach depends primarily on the outcome of the carrier model and competing carrier market share of the spot market in the prior time period. The number of atomic bids in the XOR bid set developed by the all-or-nothing bid set construction approach considers all possible combinations of train slots from the available residual track capacity and is often quite large even for small problem instances.

As observed in Table 6-1, when only limited residual track capacity exists, the average shipment delay incurred by shipments transported by trains along the awarded train slots was relatively large, irrespective of whether or not the train slots were awarded to the

carrier using the all-or-nothing bid set construction approach. However, in general, the average shipment delays were lowest for those shipments carried by trains employing train slots that were awarded to the carrier employing the train slot-based bid set construction approach.

Table 6-2 shows that when 4,000 one-off loads arrive on the spot market, assuming that each carrier will seek track access rights to support trains to carry half of these loads, and provided that residual track capacity is limited, which is the case in the experiments, the carrier that employs the all-or-nothing bid set construction approach to create an XOR bid set obtained as little as zero train slots, while the carrier employing the train slot-based bid set construction approach obtained at least a substantial portion of its requested train slots. This is because the latter carrier is willing to accept a portion of the requested train slots, while the former carrier, due to the construction of the bid set, prefers to have no request awarded if all requests cannot be awarded. Additionally, the average shipment delay incurred was high for all shipments transported via trains using the awarded train slots, regardless of carrier or bid set construction technique used when enough track capacity was awarded to each carrier to operate trains to carry at least 1,000 shipments. This implies that residual track capacity is scarce.

Table 6-3 shows that the likelihood that a carrier will be awarded sufficient track capacity to support the operation of trains to carry all desired shipments decreases with increasing prior period market share, regardless of the bid set construction approach employed. This is because the WDP seeks to maximize social benefits. Thus, if a carrier has a high prior period market share as compared with its competitors, this carrier is less likely to

receive its requested train slots and average shipment delay is likely to increase slightly (as noted in the experiments).

6.8 Conclusions and Future Considerations

A CA-based track capacity allocation framework and necessary bid set construction approaches are proposed for use in developing a train timetable for delivering one-off loads (e.g. containers) or other cargo arriving on the freight transport spot market at irregular points in time. Application of this framework can aid in creating an efficient and transparent process for allocating track capacity within an IM network. Such a process, which seeks to maximize social benefits, is mutually beneficial to the authority that controls track access rights and carriers who wish to operate trains within the network.

Results of experiments employing the proposed CA-based framework and bid set construction approaches show that the train slot-based bid set construction approach performs best in terms of number of atomic bids in the bid sets, number of shipments that are accommodated and average shipment delay. The experiments were conducted on a network with numerous O-D pairs and corresponding potential paths, thus, capturing many of the complexities that would arise in a real-world application. Application of the proposed framework to large-scale networks, e.g. the REORIENT network described in Chapter 2, would preclude the use of the all-or-nothing bid set construction approach, because this approach requires the enumeration of all potential atomic bids. The train slot-based bid set construction technique, on the contrary, generates a restricted set of most desirable atomic bids through the use of proposed carrier and revised carrier models. This procedure, thereby, permits the quick generation of a manageable number of competitive bid combinations and, thus, solution of large problem instances.

121

In spot markets, it may be the case that underutilized trains will be operated. This situation may be improved by encouraging carriers to form alliances, where carriers can purchase train space from another carrier that is operating a less-than-full train between a desired O-D pair for a desired arrival time. In Chapter 5, collaborative decision-making based approaches were proposed for collaborative freight transport in forward markets. In Chapter 7, CA-based techniques that support carriers in co-transporting one-off loads arriving on spot markets are developed. Techniques that support carrier collaboration can lead to even greater utilization of train and track capacities.

Chapter 7   Collaboration among Multiple Carriers in the Transport of Goods Arriving on Spot Markets

7.1 Introduction

In this chapter, the optimal allocation of residual track capacity among multiple competing carriers that desire to transport one-off loads arriving on the freight transport spot market is addressed. In Chapter 6, a combinatorial auction (CA)-based framework is proposed for optimally allocating residual track capacity among multiple competing carriers. This auction-based framework, when implemented, can result in an efficient and socially acceptable allocation of track capacity among the carriers. It was assumed, however, that carriers will operate trains to complete these shipments even if there is an insufficient number of shipments to fill a train for a given O-D pair. The operation of such underutilized trains is likely to be unrealistic and would lead to inefficient utilization of the track capacity and rolling stock. A more efficient and profitable utilization of these limited resources could result from collaboration among carriers. Unlike in prior works, it is assumed that carriers can co-transport shipments, i.e. carriers can lease space on competitors' trains. Alliance-based freight transport frameworks are proposed for this purpose. These frameworks may make it feasible to operate profitable trains that might otherwise have been less than full. Moreover, carriers will be able to serve a greater share of the spot market. Such a flexible framework to utilizing limited residual track capacity is essential for providing desired service levels at an acceptable cost.

Three options for accommodating shipments arising on the spot market are available for each carrier: (1) carry the shipment on an existing train with excess capacity operated by the carrier (to serve the forward market), (2) lease space from competing carriers with excess

capacity on one of their operated trains, (3) form new train lines to be operated in desired train slots for which track access rights are requested. The new train is allowed to carry competing carriers' shipments if the new train's capacity is not filled by the carrier's own shipments.

A greedy-based train space leasing approach (i.e. train space leasing approach) for accommodating shipments arising on the spot market on trains operated for the forward market (i.e. option (2)) is proposed. This technique seeks to maximize the number of one-off loads that can be accommodated by the collaborating carriers given available capacity on existing trains and characteristics of the shipments. A CA-based train slot creation approach is developed to award track access rights to operate newly formed trains (i.e. options (3)) to transport those shipments that cannot be accommodated on existing trains. In this mechanism, carriers must construct package bids for train slots. An authority that retains ownership of track access rights or that represents such an owner (i.e. the auctioneer) awards the requested track access rights to select bidders through a bid allocation process. Techniques for bid set construction (i.e. for construction of OR-package bids) and bid allocation are proposed. A Winner Determination Problem (WDP) with the objective of maximizing the average per shipment revenue over all carriers is formulated that provides the optimal allocation of track access rights to bidders for the given bid sets.

The existing auction-based approaches proposed for track capacity allocation described in the literature are reviewed in section 7.2. Assumptions made in developing the alliance-based freight transport framework are provided in section 7.3. In section 7.4, a CDM-based technique for train space leasing among existing trains is presented. In section 7.5, the CA-based train slot creation that supports cooperation among competing carriers is

presented. Results of computational experiments are presented in section 7.6. Conclusions are discussed in section 7.7.

7.2 Literature Review on Collaborative Modeling

In this chapter, alliance-based freight transport frameworks for track capacity allocation are developed that take concepts from both CDM and CA. In addition to techniques proposed in Chapter 6, Brewer and Plott (1996) and Nilsson (1999) appear to be the only works in the literature to develop auction-based approaches for track capacity allocation. As discussed in Chapter 6, their works considered only a single route; thus, obviating the need to consider interdependencies and time conflicts between trains and the routes upon which they operate. While their auction techniques provide an initial foundation for designing auction-based approaches to facilitating the allocation of track capacity, neither work captures the complexities that would arise in a real-world application of such an approach. Moreover, collaboration among carriers was not considered. In Chapter 5, the potential benefits of collaborative operation of rail-based intermodal (IM) freight services among multiple carriers were recognized. Thus, CDM strategies, including (1) train slot cooperation, (2) train space leasing and (3) train slot swapping, were proposed for this purpose. In train slot cooperation, two or more carriers operating over separate portions of the train slot's route can join forces to jointly operate a train slot. In train space leasing, a carrier is allowed to lease a portion of its excess train capacity to other carriers. Finally, in train slot swapping, two carriers are permitted to exchange track access rights for different train slots. A discrete-time carrier collaboration simulation model that replicates services, carrier operations and shipper response to proposed timetables was developed to assess the potential increase in rail-based IM market share resulting from the introduction of such

collaborative operations, results of which were shown for a large European ferry and rail network (i.e. REORIENT network). In addition to attracting more demand as was shown in Chapter 5, cost savings in terms of rolling stock and labor and reduced shipment delays can be achieved. The authors know of no additional works in the literature that address track capacity allocation with tools of auction theory or CDM.

As discussed in Chapter 5, CDM has been applied extensively in addressing, for example, air traffic flow management, supply-chain systems, submarine command and control, engineering design projects, and homeland security problems (Ball et al, 2000; Carlson, 2000; Chang et al., 2001; Sherali et al., 2003; Groothedde et al., 2005; Vossen and Ball, 2005; Cramton et al., 2006; Vossen and Ball, 2006). Among these, the works in air traffic flow management are the most relevant, especially the aircraft arrival and departure slot arrangement, which, like track capacity allocation in rail-based IM freight transport in a network, is a capacity allocation problem (Ball et al., 2000; Vossen and Ball, 2005; Cramton et al., 2006; Vossen and Ball, 2006). The goal of the aircraft arrival/departure slot arrangement is to minimize delays incurred at congested airports. Through a procedure based on CDM, arrival or departure slots to an appropriate aircraft are assigned to minimize total delay incurred by the airlines; thus, arranging slots more efficiently. Airlines can benefit from cooperation despite that they are inherently competitive.

Methodologies that combine CDM concepts with CAs have been developed for numerous applications, including freight transport procurement in the trucking industry (Song and Regan, 2004; Krajewska and Kopfer, 2006), aircraft arrival/departure slot allocation (Vossen and Ball, 2005; Vossen and Ball, 2006; Cramton et al., 2006; Ball et al., 2000), military airlift planning services (Godfrey et al., 2004), and other logistics activities

(Choi and Han, 2007; Hunsberger, 2001; Hunberger and Grosz, 2000; Ito et al., 2000). Potential cost savings and efficiencies that can be derived from such a combined approach are demonstrated on small example problems.

Song and Regan (2004) develop mechanisms to permit post-market collaborations in which carriers can coordinate their operations and derive greater economic efficiency. Song and Regan, as well as Krajewska and Kopfer (2006), developed CA-based collaboration mechanisms that permit the coordination of operations among independent freight carriers. Both works conclude that a system-wide optimal solution can be achieved and that incorporation of CDM strategies within an auction-based framework, where information and resources are shared, is promising. An example presented in Krajewska and Kopfer's (2006) illustrates that, from a user (carrier) perspective, each carrier that participates in an alliance generates no losses as a consequence of the collaboration and does so with increased profit.

Combinatorial auctions involving collaborative procurement (i.e. collaboration of multiple independent buyers who band together as a single bargaining unit to obtain goods or services at reduced wholesale prices) have been considered (Chen et al., 2002; Chen et al., 2006; Matsuo, 2005; Choi and Han, 2007). While the procurement procedures designed in these works differ in terms of auction rules and collaboration methods, results of numerical experiments from all works show that such collaborative auction-based approaches benefit both buyers and sellers.

Commercial airline carriers collaboratively schedule flights to serve the military by auctioning flights for military missions and allowing the carriers to swap flights when it is mutually beneficial after the results of the auction are known (Godfrey et al., 2004). Such

collaboration through swapping creates flexibility that can be beneficial for both the carriers and the military.

Horty and Pollack (2001) address the initial-commitment problem, i.e. where utility-maximizing agents must decide whether or not to commit to a collaborative activity. They propose a methodology based on concepts of CDM to evaluate such collaborative opportunities given pre-existing commitments. Their technique led to solutions that required some agents to cancel their prior commitments. Thus, Hunsberger and Grosz (2000) proposed a CA-based collaboration mechanism that seeks to minimize the effects on existing commitments. Hunsberger recognized that this approach required an enormous number of constraints in the WDP. Thus, he proposed an approach that generates an efficient set of bids for new activities while protecting an agent's pre-existing commitments (Hunsberger, 2002).

While applied widely, methodologies based on concepts of CDM and CAs developed for these applications cannot be readily applied to address the track capacity allocation problem addressed herein. In this paper, a collaborative CA-based approach to allocating track capacity is proposed. Results of numerical experiments show that such a technique that allows carriers to form alliances can lead to increased profits with little risk in terms of loss for the carriers. A greater number of shipments can be transported through the network as compared with allocation to individual carriers that are not permitted to work together; thus, leading to increased track capacity utilization efficiency.

7.3 Problem Assumptions

Assumptions made in developing the alliance-based freight transport framework are provided in this section.

1. Residual track capacity within the network is known.

Residual track capacity based on periodic schedules of trains formed to accommodate shipments in forward markets is known. All remaining track capacity can be used to form new train lines.

2. *Carriers' areas of operation are known.*

Carriers' areas of operation in the forward market can be geographically identical, overlapping or adjacent.

3. *Carriers willing to participate in an alliance.*

Every carrier seeking to place bids in the auction is willing to join an alliance of all carriers, and are, therefore, able to transport shipments arriving on the spot market, and is willing to offer its excess train capacity for lease.

4. *When a carrier transports a shipment belonging to another carrier, the carrier (i.e. the subcontractor) must pay no less than the transportation cost incurred.*

The carrier's costs consist of a fixed cost for operating a train that is shared by all shipments to be transported on the same train and a variable cost that depends on the distance traveled.

7.4 Train Space Leasing

When one-off loads arrive on the spot market and a carrier is selected to transport such shipments, the carrier will first consider whether or not it can accommodate the shipments on its already scheduled trains. When this is not possible, and an alliance exists, the carrier can decide between offering new train service (assuming it can obtain necessary track access rights to operate the service) and leasing space on existing trains of competing carriers in the same alliance. In this section, a train space leasing model is proposed that seeks optimal decisions for leasing train space for one-off loads to be carried on trains of

competing carriers in the alliance. The formulation can be viewed as a binary multiple knapsack problem. The objective is to maximize the number of shipments that can be accommodated on existing trains belonging to the carriers of the alliance such that all shipments are scheduled to arrive at the destination by their preferred arrival times (i.e. on time). A train space leasing methodology is proposed for its solution. It is assumed that space can be leased from any train with available capacity provided by any carrier participating in the alliance. The benefits of train space leasing across carriers in an alliance in terms of the number of shipments arriving on the spot market that can be accommodated are evaluated.

Solution produced through the use of the train space leasing approach will result in an assignment of subcontracts that efficiently uses existing capacity of scheduled trains, presumably with lower cost than operating additional train lines, while maintaining on-time performance. Such an approach, therefore, results in a win-win situation for all parties.

In subsection 7.4.1, an example is provided to illustrate the potential benefits of train space leasing. Formulation of the train space leasing problem is given in subsection 7.4.2. The train space leasing methodology for its solution is described in subsection 7.4.3.

7.4.1 Illustrating the benefits of train space leasing within an alliance

When a carrier receives requests to ship one-off loads, if no alliance exists, the carrier will choose between transporting the shipments on existing trains, forming and operating new trains on which the shipment will be placed, or rejecting the jobs. The formation and operation of new trains can be expensive, requiring the use of limited resources, such as locomotives and track capacity. Additionally, ideal train slots may not be available, resulting in shipment delays. Transportation on existing trains may not always be feasible and, even if feasible, may not guarantee on-time arrival. Thus, if an ideal situation does not exist for

accommodating the shipments on existing trains, customer satisfaction may be in jeopardy. By leasing capacity on trains of competing carriers, on-time performance can be improved and costs of transporting the shipment can be maintained at a reasonable level. Moreover, a more efficient utilization of existing train capacity can be achieved, resulting in greater economic efficiency as considered from the society as a whole.

PDT	PAT	Carrier	Shipments
Mon	Tue	A	1
Tue	Thur	A	2
Wed	Fri	A	3
Tue	Wed	B	4

Trains	Carrier	Train Type	SDT	SAT
1	A	Existing	Tue	Wed
2	A	New	Wed	Fri
3	B	Existing	Mon	Tue
4	B	Existing	Tue	Thur
5	B	Existing	Wed	Fri

PDT: Preferred departuretime   → : Carrier collaboration   SDT: Scheduled departure time

PAT: Preferred arrival time   ⇢ : Non-carrier collaboration   SAT: Scheduled arrival time
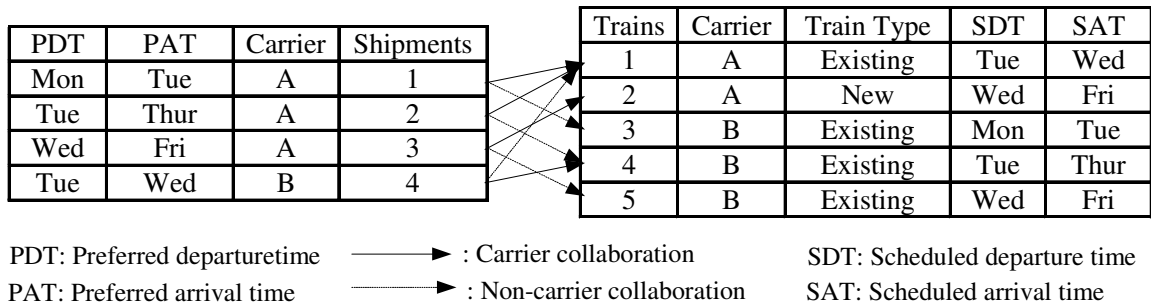
Figure 7-1 Benefits of carrier collaboration

An example that illustrates the benefits of train space leasing is shown in Fig. 7-1. Four shipments are assumed to arrive on the spot market, three belonging to Carrier A and one belonging to Carrier B. While Carrier B can accommodate its shipment, Carrier A can only accommodate two (shipments 1 and 2) of its three shipments. To transport shipment 3 by its desired arrival time, Carrier A would need to form a new train. Assuming that shipment 3 is not large enough to fill an entire train, such train formation and operation will not likely be profitable. If Carriers A and B were to form an alliance, Carrier A could take advantage of available capacity on Carrier B's trains; thus, obviating the need for additional train formation.

7.4.2 Train space leasing model

Formulation of the train space leasing problem as a binary multiple knapsack program on a bipartite graph is presented. Only those shipments that cannot be accommodated by a carrier's existing trains are considered. A train can carry such shipments with any O-D pair for which the origin and destination are contained in the train's itinerary. Binary decision variables represent whether or not a shipment is to be carried on an existing train of a competing carrier in the alliance in the optimal solution. The model exploits a bipartite representation of shipments and train itineraries in terms of track usage. In Fig. 7-2, such a representation is depicted with four shipments and two train itineraries. Edges are placed between shipments and itineraries in the bipartite graph if the train associated with the itinerary has sufficient capacity and can guarantee on-time arrival for the shipment. Solution of the train space leasing model provides an assignment of shipments to existing trains for a given planning period.
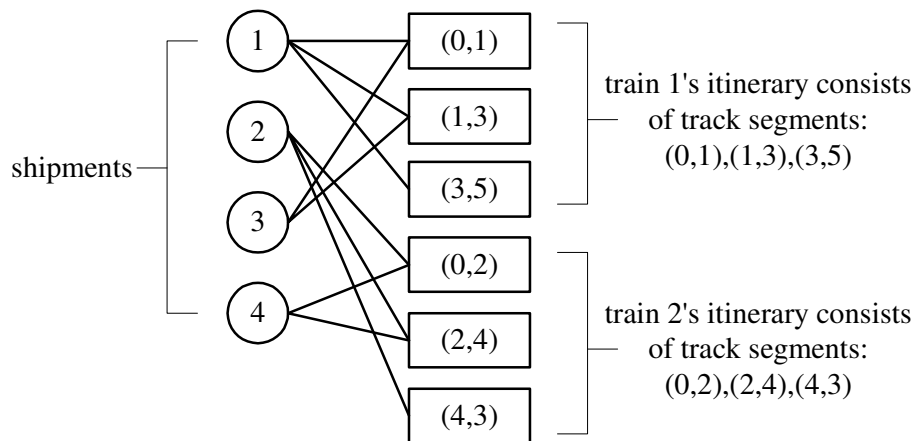

Figure 7-2 Bipartite graph representation for train space leasing model

Let J and j^A denote a set of existing trains and a set of track segments employed by train $j \in J$, respectively. S denotes the set of all shipments and R denotes track segments employed by all trains in J. Bipartite graph $G = \{V^S, V^R, E\}$, where $V^S \curlyvee V^R$ and E denote a set of vertices and edges, respectively. V^S and V^R denote a set of vertices corresponding to the shipments and track segments, respectively. $(v^s, v^r) \in E$, $v^s \in V^S$ and $v^r \in V^R$ is an edge connectivity vertices given that shipment can be transported by a train employing track segment. Additional notation required for articulating the train space leasing model are given next.

Notation

u_r: residual capacity of the train employing track segment $r \in j^A$, $j \in J$

$\varpi(s, j)$: number of track segments employed by train $j \in J$ to transport shipment $s \in V^S$

$$x_{sr} = \begin{cases} 1, \text{if shipment } s \in S \text{ is selected for transport along track segment } r \in \underset{j \in J}{\curlyvee} j^A; \\ 0, \text{otherwise.} \end{cases}$$

$$\delta_r(s) = \begin{cases} 1, \text{if shipment } s \in S \text{ can be transported along track segment } r \in \underset{j \in J}{\curlyvee} j^A; \\ 0, \text{otherwise.} \end{cases}$$

Formulation

The train space leasing model is presented in (1) through (4).

$$\text{Max} \quad z = \sum_{s \in S}(\sum_{j \in J} \sum_{r \in j^A} \frac{x_{sr}}{\varpi(s, j)}) \tag{1}$$

subject to

$$\sum_{j \in J} \frac{\sum_{r \in j^A} \delta_r(s)x_{sr}}{\varpi(s, j)} \leq 1, \qquad s \in S \tag{2}$$

$$\sum_{s \in S} \delta_r(s) x_{sr} \leq u_r, \qquad\qquad r \in j^A, j \in J \qquad\qquad (3)$$

$$x_{sr} \in B, \qquad\qquad s \in S, r \in j^A, j \in J \qquad\qquad (4)$$

The objective given in equation (1) seeks to maximize the total number of shipments that can be delivered through train space leasing, thus, maximizing utilization of existing train capacities. Constraints (2) ensure that each shipment $s \in S$ is transported by at most one train $j \in J$. Constraints (3) force the sum of shipments $s \in S$ that will be accommodated by a train $j \in J$ to be at most u_r, $r \in j^A$, $j \in J$. Binary integrality requirements are enforced through constraints (4).

The structure of the bipartite graph is amenable to simultaneous modeling of multiple O-D pairs. Each train ($j \in J$) itinerary consists of multiple track segments and includes a subset of vertices $v^r \in V^R$ representing these track segments. Let V_{od}^R be a subset of V^R for a given O-D pair. An edge exists in G between vertex v^s that corresponds to a shipment with a given O-D pair and each vertex $v^r \in V_{od}^R$. All shipments with a given O-D pair will be connected to the same vertices of the train itineraries. If shipment vertices are connected to different subsets of vertices representing track segments in train itineraries, then these shipments must have different O-D pairs.

The number of decision variables and constraints of the train space leasing formulation given in (1) through (4) depends on the number of track segments employed, number of trains, and number of shipments. One can expect the number of decision variables and constraints to be quite large for a real-world problem instance, because the number of decision variables will increase exponentially with the number of track segments, trains and shipments. In addition, the problem as formulated, is a binary multiple knapsack problem,

which is well-known to be NP-hard (Martello and Toth, 1990). Thus, a formidable computing task would be expected if solution to optimality is desired.

7.4.3 Train space leasing approach

A train space leasing approach is proposed for solution of the train space leasing model. The technique iterates over each O-D pair and terminates with a feasible solution. The technique begins by identifying residual capacity on existing trains for a given O-D pair. The trains with residual capacity are ranked in nonincreasing order based on their scheduled arrival times.

The technique consists of two main phases: train selection and shipment selection. In each iteration of the train selection phase, the train with the latest arrival time at the destination is selected (if not previously selected). In the shipment selection phase, the shipment with the latest arrival time at the destination (if not already assigned) is considered. If the shipment can be accommodated by the identified train from the train selection phase, the shipment is assigned to the train. Identification of a shipment in the shipment selection phase continues until either no remaining capacity exists on the selected train or all shipments have been considered and the procedure returns to the train selection phase. The technique iterates until all trains have been evaluated. This train space leasing approach is described in more detail next.

Train space leasing approach

Step 1. Sort the trains and shipments. Let J define a set of existing train slots that are ordered by decreasing scheduled time of arrival at destination. Let *S* define shipments that are ordered by decreasing preferred arrival time at destination.

Step 2. Select train. If $J = \varnothing$, procedure terminates. Otherwise, select the train slot $j \in J$

that arrives at the destination latest among the trains in the J and $J = J \setminus \{j\}$. Let

SDT_j and SAT_j represent the train j's scheduled departure time at origin and arrival

time at destination, respectively. Let j^R denote a set of track segments employed by

train $j \in J$. Let u_j^r denote the excess capacity of the track segment donated as

$r \in j^R$ employed by train j.

Step 3. Select shipment. Select the shipment $s \in S$ that has the latest arrival time at the

destination and has not been selected in *Step 3*. If all shipments have been selected or

$S = \varnothing$, go back to *Step 2*. Let PDT_S and PAT_S represent the preferred departure

and arrival times of the shipment s, respectively.

Step 4. Upload the shipment to the selected train. If $PDT_S \leq SDT_j$, $PAT_S \geq SAT_j$ and

$u_j \neq \varnothing$, the shipment s is scheduled to be delivered by the selected train j. If the

shipment s can be delivered by train j, $S = S \setminus \{s\}$ and remove the capacity of track

segments in j^R that is employed by shipment s. Return to *Step 3*.

Proposition 1. The train space leasing approach results in the optimal solution to the train

space leasing problem ((1)-(4)) for each independent O-D pair.

Proof. Let J and S be a set of trains and shipments, respectively, ordered by decreasing

scheduled time of arrival at the destination, according to the train space leasing approach.

Trains in J and shipments in S all have different arrival times. Suppose that each train in J

can only accommodate one shipment.

Let Y be a solution generated through the use of the train space leasing approach. Through the train space leasing approach, l^{th} shipment is delivered via the first available train (see *Step 2 and 3*). Therefore, at most $|J|$ shipments can be accommodated in solution Y obtained by the train space leasing approach.

Suppose in the solution Y, the lth shipment delivered by the kth train in J. Suppose that this is not the case, i.e. the lth shipment is assigned to other available train, the mth train in J, where $m > k$.

When the $(l+r)^{th}$ shipment where $r \geq 1$ tries to choose a train from a set of eligible trains, $J^{(l+1)^{th}}$, it is certain that $J^{(l+1)^{th}}$ will not contain the mth train. However, for the kth train, we need to discuss two situations: (1) k^{th} train is included in $J^{(l+r)^{th}}$. (2) k^{th} train is not included in $J^{(l+r)^{th}}$.

Situation (1): k^{th} train is included in $J^{(l+r)^{th}}$

> If k^{th} train is included in $J^{(l+r)^{th}}$, since k^{th} train can be chosen to accommodate the $(l+1)^{th}$ shipment, the number of shipments that could be accommodated is at most $|J|$.

Situation (2): k^{th} train is not included in $J^{(l+r)^{th}}$

> If the kth train is not included in $J^{(l+r)^{th}}$, it will not be included in any set of eligible trains $J^{n^{th}}$, where $n > l + r$. The $(l+r)^{th}$ shipment has latest arrival time among the shipments that have not accommodated by trains. In this case, the maximum number of shipments that can be accommodated is at most $|J|$ -1, since the kth train will not be able to accommodate any additional shipment.

From situation (1) and (2), if the shipment is not transported by the first available train, we conclude that the number of shipments that can be accommodated in solution Y is less or equal to $|J|$. Since the at most $|J|$ shipments can be obtained though the use of the train space leasing approach, therefore, we can conclude that the problem can be solved to optimality through the use of the train space leasing approach.□

7.5 Train slot creation

If insufficient capacity exists on the existing trains of carriers in the alliance to accommodate the desired shipments, new trains will need to be formed. Trains are formed to accommodate shipments with varying, but synergistic O-D pairs. Typically, however, each carrier will have an insufficient number of such synergistic shipments to form profitable trains. If carriers cooperate in the formation of trains, profitability may be achieved. In this section, the problem of creating trains for accommodating one-off loads arriving on the spot market through the cooperation of multiple carriers is considered. Carriers may operate in overlapping operational areas. It is assumed that a single carrier will operate each additional train and that carriers in an alliance can lease space on these trains at a pre-determined price. A carrier will not choose to operate a train unless at least some portion of the train slot is in its operational area.

A CA-based train slot creation approach that seeks to maximize the total profit obtained by members of the alliance (employing a proxy based on the number of high-valued shipments that can be accommodated) is proposed. A set of candidate atomic bids that allow for cooperation is constructed through a bid set construction technique presented in subsection 7.5.1. Resulting candidate bids are used as input to a winner determination problem (formulated in section 7.5.2) that seeks the optimal allocation of track capacity to

requested train slots. This CA-based approach results in a cooperative timetable for multiple carriers that facilitates the co-transport of shipments that cannot be accommodated by existing trains already operated by the carriers.

7.5.1 Bid set construction

In the proposed bid set construction technique, it is assumed that all carriers in the alliance will inform the alliance members of any shipments for which it seeks transport by an alternate carrier. A carrier first constructs a train slot in its operational area on which a train will be operated. The carrier considers its own shipments and the shipments offered for subcontract by other carriers in the alliance when forming the train slots. When considering shipments of other carriers for inclusion in a train, only those shipments that can be accommodated with direct service by the shipment's preferred arrival time can be included. The bids are, thus, constructed in such a way that cooperation is permitted.

The proposed bid set construction technique is comprised of two main components: atomic bid construction and atomic bid selection, described next.

Atomic bid construction

Atomic bids are generated by constructing a set of candidate train slots with their associated shipments from the residual track capacity. For simplicity, it is assumed that these train slots employ only the operational routes used in serving the forward market. A subset of the candidate train slots are selected based on how well they might serve the carrier's desired additional shipments (i.e. the one-off loads). If residual train capacity exists, the carrier will seek to transport shipments from other carriers in the alliance. Note that if more than one additional train will be operated between a given O-D pair, shipments will be first included on the train of the train slot with the shortest distance between the given O-D pair. Once the

139

capacity of such a train is depleted, remaining shipments will be placed on trains employing the longer distance route.

Each atomic bid consists of, not only a train slot request and its associated shipments that will be delivered through that train slot, but also a bid price. Typically, in an auction, the bid prices are equivalent to the amount that the carrier is willing to pay for the bidding item. In this context, however, carriers operate in an alliance and the goal of the auctioneer is to award track access rights such that social benefits are maximized. Thus, the bid prices are statements of revenue, as opposed to an amount that the carrier is willing to pay. It is assumed that a carrier will not bid on a train slot that is not profitable. This atomic bid construction procedure is described next.

Atomic bid construction procedure

Step 1. (Initialization)

   1.1 Let E^R denote a set of routes operated by the carrier that are included in the existing train schedules. Let F denote a set of routes along which the train slots will be constructed. $F = E^R$.

Step 2. (Select route)

   2.1 If $F = \varnothing$, returns to *Step 1.1*. Otherwise, randomly choose a route $f \in F$ and remove it from further consideration:

Step 3. (Construct train slot)

   3.1 Construct a train slot, denoted g, along selected route f from the route's origin to its destination. Based on the residual track capacity, assign the earliest available track capacity for each track segment of the train slot's route. Remove this track-time segment from the residual track capacity.

3.2 If no train slot can be formed along route f, $E^R = E^R \setminus \{f\}$. If $E^R = \varnothing$, procedure

terminates. Otherwise, return to *Step 2*.

Step 4. (Collect shipments)

4.1 Collect a set of shipments that belong to carrier and competing carriers indicated as

M^b and M^c, respectively, where shipments are eligible for assignment to train slot g.

Order the shipments in M^b and M^c in decreasing order of revenue.

Step 5. (Assign shipments)

5.1 If $M^b \neq \varnothing$, go to next step. Otherwise, choose M^c .

5.2 If $M^b = \varnothing$ or $M^c = \varnothing$, go to *Step 6*. Otherwise, choose shipment m from the

shipment set (M^c or M^b).

5.2 Assign shipment m to train slot g and remove m from the shipment set, if sufficient

capacity exists. Repeat *Step 5.1*. If train capacity is filled, go to next step.

Step 6. (Create an atomic bid)

6.1 The shipments assigned to train slot g and the train slot comprise an atomic bid.

6.2 Calculate the cost and revenue of delivering the shipments through the use of train

slot g and set the price for the associated atomic bid. Return to *Step 2.1*.

This procedure terminates with a set of atomic bids, each consisting of a train slot and

a set of shipments that will be delivered by the train operated along that train slot. These

atomic bids may serve overlapping sets of shipments. Thus, the train slots are substitutable.

The XOR bidding language described in Chapter 6 can be used by the carrier to explicitly

express preferences among atomic bids, where atomic bids are additive, substitutable, or

complementary. Note, however, that such XOR bidding language may require $2^{|J|} - 1$ atomic

bids, where J is the set of train slots represented by the atomic bids. Thus, the number of the

bids increases exponentially with increasing number of train slots. As the XOR bids are fed to the WDP for selection, the problem of solving the WDP may become insurmountable. Thus, an atomic bid selection technique is employed that selects an efficient subset of these atomic bids. This procedure is described next.

Atomic bid selection

Some atomic bids that are constructed through the use of the atomic bid construction procedure may have substitutabilities. Thus, a XOR bidding language that may increase the computational efforts of the WDP is required to express the carrier's preference. To reduce the number of atomic bids constructed by the atomic bid construction procedure, an atomic bid selection model is proposed whose solution results in a subset of competitive, nonsubstitutable atomic bids. Thus, if all atomic bids in the reduced set are awarded, trains will be run on all associated train slots and all desired shipments can be accommodated. That is, the smaller set of atomic bids can be expressed as an OR bid.

Before proceeding to the presentation of the mathematical formulation of the atomic bid selection model, notation employed in the formulation is given.

Notation

J^b : set of carrier b's atomic bids generated from the atomic bid construction procedure

S : set of shipments

r_{j^b} : revenue of bid j^b

$$y_{j^b} = \begin{cases} 1, & \text{if bid } j^b \text{ is selected;} \\ 0, & \text{otherwise.} \end{cases}$$

$$\rho_j^s = \begin{cases} 1, & \text{if shipment } s \in S \text{ is included in bid } j^b; \\ 0, & \text{otherwise.} \end{cases}$$

<u>Mathematical formulation</u>

The atomic bid selection formulation is given in (5) through (7).

$$\text{Max} \quad z = \sum_{j^b \in J^b} r_{j^b} y_{j^b} \tag{5}$$

subject to

$$\sum_{j \in J^b} \rho_{j^b}^s y_{j^b} = 1, \qquad \forall s \in S \tag{6}$$

$$y_{j^b} \in B, \qquad \forall j^b \in J^b \tag{7}$$

The objective given in equation (5) seeks to maximize total revenue of running trains along provided routes within the network such that all shipments are accommodated. Constraints (6) ensure that each shipment $s \in S$ is accommodated by one train slot $j^b \in J^b$. Binary integrality requirements for every edge are given in constraints (7). Note that it is assumed that the carrier seeks track access rights for train slots to accommodate shipments for which sufficient residual track capacity exists. The resulting bid set will be used as input for the winner determination problem, which determines the optimal bid allocation, described next.

7.5.2 Winner determination problem

A WDP formulation is presented in this section whose objective is to determine the optimal bid allocation. Solution of the WDP results in a cooperative train timetable for operating additional trains to transport one-off loads arriving on the spot market. That is, residual track capacity given the train timetable for trains accommodating shipments in the forward market is allocated among the carriers based on the carrier bids. The WDP is formulated as a binary multiple knapsack problem. The objective is to maximize the number of high priority or high-valued shipments that are accommodated. Before proceeding to the

presentation of the mathematical formulation of the WDP, notation employed in the formulation is given.

<u>Notation</u>

C : set of carriers

B : set of bidders

T : set of time intervals

A^t : set of tracks at time interval $t \in T$

S^c : set of shipments owned by carrier $c \in C$

J^b : set of bids constructed by bidder $b \in B$

p_{j^b} bidding price associated with bid $j^b \in J^b, b \in B$

$\omega(j^b)$: number of shipments contained in bid $j^b \in J^b, b \in B$

a^t : track $a \in A$ at time interval $t \in T$

$$x_{j^b} = \begin{cases} 1, & \text{if bid } j^b \in J^b, b \in B, \text{ is selected;} \\ 0, & \text{otherwise.} \end{cases}$$

$$\theta_{j^b}^{s^c} = \begin{cases} 1, & \text{if shipment } s^c \text{ will be transported within a train slot contained in bid } j^b; \\ 0, & \text{otherwise.} \end{cases}$$

$$\delta_{j^b}^{a^t} = \begin{cases} 1, & \text{if train slot contained in bid } j^b \text{ employs track } a^t; \\ 0, & \text{otherwise.} \end{cases}$$

<u>Formulation</u>

The WDP formulation is given in (8) through (11).

$$\text{Max} \quad z = \sum_{b \in B} \sum_{j \in J^b} \frac{p_{j^b}}{\omega(j^b)} x_{j^b} \tag{8}$$

subject to

$$\sum_{b \in B} \sum_{j \in J^b} \theta_{j^b}^{s^c} x_{j^b} \leq 1, \qquad \forall s^c \in S^c, \ c \in C \tag{9}$$

$$\sum_{b \in B} \sum_{j \in J^b} \delta_{j^b}^{a^t} x_{j^b} \leq 1, \qquad \forall a^t \in A^t, \ t \in T \tag{10}$$

$$x_{j^b} \in B, \qquad \forall j^b \in J^B, \ b \in B \tag{11}$$

The objective given in equation (8) seeks to maximize the average per shipment revenue over all carriers. Constraints (9) ensure that each shipment s^c can only be included in at most one bid x_{j^b}. Constraints (10) ensure that each track segment for a given time period can be assigned to at most one train slot (i.e. bid) j^b. Binary integrality requirements are enforced through constraints (11).

Solution of the WDP results in a cooperative timetable of trains to be operated by multiple carriers. It is possible that not all shipments will be accommodated in the final schedule. It may be the case that two competing carriers with overlapping operational areas produce identical bids (including identical bidding prices), where the bids are constructed under the assumption that shipments belonging to the competitor will be shipped by a train employing the awarded train slot. The proposed WDP formulation would treat these bids as equivalent bids and if one should be awarded in the optimal solution, an alternative optimal solution exists in which the identical bid of the competing carrier is awarded. This can be avoided by including a factor (e.g. prior market share of each carrier or equity in awarding train slots among the carrier) in the objective function.

Observation. Consider an identical auction approach, where no collaboration is permitted. It is always beneficial for the carriers to participate in the alliance. That is, for the given set of

shipments arriving on the spot market in the given time interval, assuming revenues and costs to be identical in both collaborative and noncollaborative approaches, no carrier will lose profit by virtue of the collaboration.

Discussion. Cooperation in an alliance, where train space leasing is permitted and bid allocation is based on revenue as in equation 8, will lead to lower total operational costs incurred by all members of the alliance and no loss in profit for any member.

Consider a network with only one route. Let J' be a set of train slots along the route awarded through the use of the WDP. C and X^c denote a set of carriers and set of carrier shipments $c \in C$ to be transported along the route, respectively. H_{j^c} is the total operational cost of a train slot $j^c \in J'$ operated by carrier $c \in C$. The operational cost is assumed to consist of a fixed cost, F_{j^c}, (i.e. the cost for operating a train) and a variable cost, V_{j^c}, which depends on requirements of the shipments to be carried on the train. The variable cost for each shipment $x \in X^c$ is denoted as $v(x)$, i.e. $\sum_{c \in C} \sum_{x \in X^c} v(x) = V_{j^c}$, and is assumed to be constant regardless of which carrier handles its final transport. Let $f(x)$ denote a single shipment's ($x \in X^c$) share of the fixed cost. Finally, r_x denotes the revenue obtained from delivering shipment $x \in X^c$.

Suppose shipments $\underset{c \in C}{\mathrm{Y}}\, X^c$ can be transported by a single train $j^c \in J'$. If only one train is operated, the operational cost for transporting these shipments can be expressed as in (12):

$$H_{j^c} = F_{j^c} + V_{j^c} = F_{j^c} + \sum_{c \in C} \sum_{x \in X^c} v(x). \tag{12}$$

146

If, however, each carrier with one or more shipments in X^c were to operate a train to accommodate its own shipments, the operational costs would be expressed as in (13):

$$\sum_{c\in C}\sum_{j^c\in J'^c} H_{j^c} = \sum_{c\in C}\sum_{j^c\in J'^c} F_{j^c} + \sum_{c\in C}\sum_{j^c\in J'^c} V_{j^c} = \sum_{c\in C}\sum_{j^c\in J'^c} F_{j^c} + \sum_{c\in C}\sum_{x\in X^c} v(x).\qquad(13)$$

Assuming that slot j^c will be operated under both scenarios, the following inequality (14) holds:

$$\sum_{c\in C}\sum_{j^c\in J'^c} H_{j^c} - H_{j^c} = [\sum_{c\in C}\sum_{j^c\in J'^{cc}} F_{g^c} + \sum_{c\in C}\sum_{x\in X^c} v(x)] - [F_{j^c} + \sum_{c\in C}\sum_{x\in X^c} v(x)] \geq 0.\qquad(14)$$

That is, the total cost to the alliance of accommodating a set of shipments employing train space leasing is no more than the total cost incurred when no such collaboration is permitted. This result can be immediately extended to more than one route.

A train slot is included in an atomic bid if and only if the inequality (15) holds:

$$\sum_{c\in C}\sum_{x\in X^c} R_{x^c} \geq H_{j^c}.\qquad(15)$$

If $|C| = 1$, the revenue obtained by this single carrier will be $\sum_{x\in X^c} R_x$ if the carrier operates a train in train slot j^c. If $|C| > 1$, the carrier, c, operating a train in train slot j^c will transport one or more shipments originally contracted to a competing carrier. Suppose $|C| = 2$ and that carrier c operates a train in train slot j^c. Carrier $r \in C, r \neq c$, will send its shipments on carrier c's train, assuming that there is enough capacity and that the total revenue $\sum_{x\in X^r} r_x$ obtained by carrier r exceeds the price that carrier c charges for the transport. Moreover, carrier c will charge a positive fee that will cover, at a minimum, the additional operating

costs, $\sum_{x \in X^r} f(x) + \sum_{x \in X^r} v(x)$, incurred for transporting carrier r's shipments. Thus, the revenue obtained by carrier c will be the sum of the revenue obtained for transporting its own shipments and the fees generated by transporting carrier r's shipments.

In the worst case, carrier c will receive zero revenue, i.e. carrier c is no worse off as a result of the collaboration. Therefore, carrier collaboration will have lead to at least no loss in profit for any members of the alliance. □

Cooperative scheduling of trains to accommodate one-off loads arriving on the spot market permits efficient, cost-effective use of residual track capacity, resulting in a competitive alternative to truck transport.

7.6 Results of Numerical Experiments

Results of computational experiments designed to illustrate the feasibility of applying the proposed cooperative CA for track capacity allocation to carriers seeking to transport one-off loads arriving on the spot market in a given time period are provided and analyzed. The carriers are assumed to operate within an alliance, where each carrier can choose to lease space on trains operated by competing carriers in the alliance. Results from the train space leasing approach for assigning one-off loads to existing trains of carriers in the alliance and the alliance-based train slot cooperation approach for allocating track capacity for operating new trains and assigning shipments to these trains were examined in the experiments. Results of both techniques were evaluated in terms of the number of shipments delivered. The number of trains scheduled and average train capacity utilization of newly scheduled trains were also considered in assessing the latter technique. To assess the potential gains that

might be achieved through cooperation, these measures were considered in both cooperative and noncooperative environments.

The bid set construction approaches and WDP were coded in Microsoft Visual C++ 6.0 and CPLEX 10.1 was employed. These codes were run on an Intel Pentium 4 with 3.6 GHz CPU and 2 GB RAM using the Windows XP operating system. Track capacity of the 6-vertex, 8-edge network (Fig. 7-3) employed in Chapter 6 with double tracks along each edge is assumed to be allocated among two carriers over a one-week period. Either 1,000 or 2,000 one-off loads are assumed to be of interest to both carriers. Residual track capacity is fixed at 30%. Five runs, associated with five sets of randomly generated shipments, were made for each combination, results of which are reported individually. Residual track capacity is computed based on a train timetable generated for the forward market. This timetable was developed using the initial track capacity allocation technique introduced in Chapter 4 for a randomly designed demand set.

Travle Time (hour)
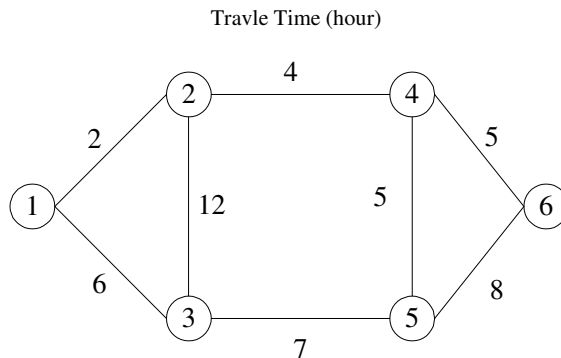

Figure 7-3 Illustrative network

For each set of shipments generated, three approaches were employed to allocate the shipments to trains: (1) train space leasing, (2) CA-based train slot creation, and (3) a non-cooperative CA approach. The third approach is identical to the CA-based train slot creation approach; however, carrier which operates the train cannot accommodate shipments that

belong to other competing carriers. Results of the train space leasing approach can be compared to a "do-nothing" approach, where no shipment is placed on a train operated by a competing carrier. Recall from Section 7.4 that only those shipments that cannot be accommodated on a carrier's existing trains will be considered by the train space leasing approach.

Results of the experiments are provided in Tables 7-1 and 7-2. For each set of shipments, results are provided for two cases: non-cooperative CA and carrier collaboration. The former is generated by the non-cooperative CA approach (approach (3)), while the latter is produced from results of runs of two procedures (approaches (1) and (2)). That is, it is assumed that shipments are first accommodated on existing trains and only once this existing capacity is depleted are new trains introduced. Thus, the number of shipments transported under carrier collaboration is the sum of the number transported under train space leasing and the number transported under train slot creation. Similarly, the number of trains scheduled and average train capacity utilization are computed from results of both approaches. Details of the number of shipments to be handled only through train space leasing on existing trains and the additional shipments accommodated through the train slot creation approach (i.e. CA-based approach) are provided in Table 7-2. Note that if residual capacity is too restricted to support the desired train slots requested to accommodate all desired shipments, only a fraction of the shipments will be transported.

Table 7-1 Experimental results with and without carrier collaboration

Number of shipments that arrive on market	1,000 shipments	2,000 shipments

	Non-cooperative CA	Carrier collaboration	Non-cooperative CA	Carrier collaboration
Number of shipments accommodated	589	917	1,195	1,940
	615	943	1,074	1,886
	593	921	897	1,908
	637	990	1,205	1,806
	654	898	1,163	1,828
Number of trains scheduled	23	19	39	41
	25	20	38	39
	24	22	36	43
	27	23	32	42
	26	20	35	40
Average train capacity utilization	51%	84%	61%	82%
	49%	78%	57%	82%
	49%	74%	56%	80%
	47%	78%	67%	76%
	50%	76%	66%	80%

Table 7-2 Number of shipments accommodated under carrier collaboration

1,000 shipments		2,000 shipments	
Train space leasing	Train slot cooperation	Train space leasing	Train slot cooperation
121	796	267	1,673
159	784	295	1,591
104	817	196	1,712
98	892	219	1,587
137	761	223	1,605

The results in Table 7-1 indicate that, given either 1,000 or 2,000 one-off loads arriving on the market, through collaboration (i.e. the joint results of train space leasing and alliance-based train slot cooperation) a greater number of shipments (on average greater by no lower than 50%) and a higher train capacity utilization (with an increase on average of no less than 30%) are noted as compared with the non-cooperative CA approach.

One may also observe that, through the use of the non-cooperative CA approach, the average train capacity utilization given 2,000 shipments arriving on the market is higher (on the order of 24%) than where 1,000 shipments arrive on the market. However, through

collaboration, the average train capacity utilization given 2,000 shipments arriving on the market is only slightly higher (on the order of 2%) than where 1,000 shipments arrive on the market. This may be due to the fact that when 2,000 shipments, as compared with 1,000 shipments arrive on the market, more shipments were accommodated by the existing trains through train space leasing. Note that the average train capacity utilization does not include utilization rates for existing trains. Additionally, a greater number of new train lines were scheduled to operate.

The results in Table 7-2 indicate that under both shipment scenarios (i.e. 1,000 or 2,000 shipments arriving on the market), of the shipments that are to be transported, approximately 12% on average were accommodated through leasing space on existing trains and 88% on average were handled by new trains on which carriers in the alliance can lease space.

7.7 Conclusions

Collaborative techniques for developing a train timetable from residual track capacity for delivering one-off loads arriving on a spot market to be shipped by multiple carriers that form an alliance are proposed. A train space leasing approach is described that assigns these shipments to existing trains with excess capacity belonging to the competing carriers. Such existing trains are operated to accommodate freight in a forward market. For a given O-D pair, employing the train space leasing approach can lead to maximizing the number of shipments that can be accommodated by existing trains. A train slot creation technique, a CA-based approach, is proposed for allocating track capacity among the carriers that seek to operate additional trains to ship the one-off loads that cannot be accommodated through train

152

space leasing. The proposed bid set construction technique can lead to an OR bid, consisting of competitive atomic bids, and thus, does not increase the computational burden of the WDP.

The proposed techniques are compared with a non-collaborative CA approach for adding trains to accommodate these shipments, where it is assumed that each carrier will transport its own shipments without assistance from competing carriers. Results of numerical experiments show that application of the proposed collaborative techniques can aid in creating an efficient allocation of track capacity within an IM network. Such a process, which seeks to maximize social benefits, is mutually beneficial to the authority that controls track access rights, as well as the carriers who wish to operate trains within the network. The results also indicate that employing the proposed collaborative framework, as compared with a comparative non-cooperative CA, results in a greater number of shipments that can be accommodated by awarded train slots with fewer trains; thus, leading to higher train capacity utilization and more efficient use of limited public and private resources.

The experiments were conducted on a network with numerous O-D pairs and corresponding potential paths, thus, capturing many of the complexities that would arise in a real-world application. The methodologies developed in this work can be applied to a large-scale network, e.g. the REORIENT network described in Chapter 2. Instead of solving the train space leasing model, which is well-known to be NP-hard to optimality, the proposed train space leasing methodology can efficiently assign shipments to existing trains given multiple O-D pairs to generate a near optimal solution. The proposed bid construction approach not only enables the carrier to submit a competitive bid set, but also decreases the complexity of the WDP. Although the carrier's preference can be fully expressed based on the complete set of atomic bids through the use of an XOR bidding language, the complexity

of the WDP will quickly increase with increasing network size. A balance between carrier expressiveness and ease of handling the WDP is sought. The bid set construction technique, which produces a manageably sized bid set that generally captures the carrier's preferences is proposed for this purpose.

It is assumed that the transfer of cargo between trains at intermediate terminals is not permitted, i.e. that trains will be operated with shuttle service with intermediate drop-off and pick-up locations, requiring no train classification at such intermediate locations. Such service will be particularly relevant in Europe, where there are strong advocates for the creation of shuttle trains offering international service that can handle high priced goods competitively as compared with truck. The proposed techniques can be extended to address the train space leasing and train slot creation problems considered herein with transfers.

Chapter 8   Conclusions and Extensions

8.1 Conclusions

This dissertation addresses freight train scheduling on forward and spot markets. The problems addressed in this dissertation were not only motivated by a particular European region, where it is hoped that the European railways will be transformed from nationally fragmented rail-based IM freight systems into an internationally integrated network, but also by U.S. rail-based freight transport, where train operators retaining infrastructure ownership seek to optimize on-time performance.

This work contributes to this field in several respects. A train slot selection model was proposed to address the multiple decision-maker freight train scheduling problem and a column generation-based solution approach was proposed for its solution. The solution approach when embedded in a simulation-based iterative framework, where demand for rail services is re-computed in accordance with the train schedule obtained by solving the freight train scheduling problem, results in an optimal or near optimal schedule given elastic demand.

A combinatorial auction (CA)-based track capacity allocation framework that can lead to efficient, transparent and socially acceptable allocation of track capacity among the carriers, as well as the necessary bid set construction and allocation methodologies, were developed for allocating residual track capacity among multiple carriers to facilitate the delivery of shipments on an as needed basis.

The potential benefits of collaborative operation of freight services among multiple carriers in both forward and spot markets were recognized. Collaborative decision-making (CDM) based strategies, including trains slot cooperation, train slot swapping, and train space leasing proposed for forward markets, as well as train space leasing and CA-based

train slot creation approaches for spot markets were proposed. It is anticipated that results of such techniques will lead to more efficient and profitable utilization of limited track capacity and rolling stock.

The introduction of train timetables over rail networks through the use of the scheduling techniques will permit efficient and cost-effective rail-based freight transport, reducing the amount of freight shipped by trucks. This, in turn, can positively impact the environment through reductions in pollution due to freight traffic. Rail services that can respond to demand from spot markets will be better positioned to compete with truck transport.

CDM strategies proposed for use in forward and spot markets can aid in integrating carriers' operations and, thus, reducing shipment delays. In addition, cost savings in terms of rolling stock and labor that can be achieved through such strategies enables carriers to provide competitive prices, shifting the demand from truck transport.

8.2 Extensions

In Chapter 5, three CDM strategies were proposed that involves collaborative operation of international rail-based services by multiple carriers seeking to operate within an alliance. To implement these strategies in actual rail operations, operational information of competing carriers must be shared among members of the alliance. An authority jointly selected by members of the alliance would work on behalf of the alliance. The shared information is provided to facilitate the assignment of resources to carriers with transportation needs.

Information required to implement the first of the three proposed strategies, train slot cooperation, may include existing itineraries of trains operated on the network and

knowledge of the desired shipments that cannot be transported on a member's scheduled trains. Information that is needed to implement the train slot swapping strategy, the second proposed CDM strategy, includes train slots that carriers are willing to swap for train slots of competing carriers, desired train slots belonging to competing carriers, and knowledge of competing carriers' new transportation needs. Information that is required to employ the train space leasing strategy, the third collaborative strategy, includes excess train capacity of trains operated by carriers in the alliance and the details of shipments for which carriers seek for transport. Rules for allocating available resources, such as available train capacity and train slots that can be swapped, to the carriers of the alliance for each carrier must be constructed and agreed upon.

Likely objectives of the alliance in creating such rules for implementing the CDM strategies are to maximize total revenue (through efficient use of track capacity) and ensure equality in allocating or re-allocating resources and in revenue distribution. Mechanisms that might be employed to create a collaborative environment in which the incentives for competing carriers to operate given determes of sharing proprietary information about their business are discussed next.

1. Number of train slots traded in by a carrier must equal to those assigned from the authority. For all carriers in the alliance, the value to a carrier of the train slots traded to other carriers must equal, or nearly equal the value of train slots received from other carriers for a given time period.

2. A carrier "A" that leases space on a competing carrier "B"'s train will give priority to carrier B when that carrier seeks to lease space on one of carrier "A"'

157

trains. Alternatively, carrier A can pay carrier B to lease the space with no further obligation.

3. Two carriers will only agree to the joint operation of a train slot if it is beneficial to both. Such benefit can be derived through payment received from the shippers directly or by one carrier to another.

4. Any train slot purchased jointly by two or more carriers will be shared by the carriers in proportion to the fee that the carrier pays.

Such rules for implementing the proposed CDM strategies will promote fair and efficient resource sharing among multiple competing carriers, where no carrier will be worse off as a result of the collaboration. The revenue resulting from delivering shipments must be equitably distributed among the carriers that operate the trains or own the shipment delivery contracts. One approach that could support a fair distribution of revenue among the carriers would be to ensure that the carrier operating the train on which a competing carrier's shipment is transported is compensated for more than the marginal cost of including the shipment on the train.

More sophisticated collaborative mechanisms can be proposed based on the proposed CDM strategies in Chapter 5. For example, three or more carriers may jointly operate separate portions of a route, where they may choose to swap train slots. There may be a limit on the number of swaps permitted between any pair of carriers. Train capacity can be leased to more than one carrier. Additionally, only those scenarios where collaboration is permitted among all carriers on any route were considered herein. However, it may be the case that only a portion of the carriers may enter into collaboration agreements along a given route.

Assessment of the potential of these and other more advanced CDM strategies would require further investigation.

Disruptions can also be mitigated by applying strategies that are developed based on concepts of CDM. For example, if disruptions occur along a route, the train slot swapping strategy can be employed to gain the assistance of a competing carrier operating an alternate train route for a relative O-D pair. This train slot swapping strategy can be viewed as a re-routing strategy.

In Chapter 5, the proposed three CDM strategies were evaluated through simulation. Analytical optimization models that involve multiple objectives (i.e. fairness and efficiency) may be employed to implement these strategies that can be viewed as a multi-agent (i.e. multi-carrier) decision-making process.

A collaborative CA-based approach is proposed in Chapter 7 that results in a cooperative timetable for multiple carriers. Such a timetable facilitates the co-transport of shipments (i.e. a train is allowed to carry competing carriers' shipments if the train's capacity is not filled by the carrier's own shipments) that cannot be accommodated by existing trains already operated by the carriers. To implement this approach, details of each carrier's shipments (i.e. O-D pair, route, desired arrival time) must be shared with all carriers in the alliance. Excess train capacity can be allocated among carriers in proportion to the number of shipments that will be transported by trains employing this track capacity. Conflicts between requests can be resolved by allocating the desired track capacity to the carrier that will obtain the greatest revenue from the allocation.

Chapters 6 and 7 of this dissertation focus on bid set construction and formulation of the winner determination problem for track capacity allocation through the use of CAs.

Because it appears that no prior published works have applied CAs for track capacity allocation, the use of combinatorial auctions for this purpose offers many research opportunities. Bidding prices were assumed to be true valuation costs in this dissertation. A bidding approach that can incorporate the stochastic nature of travel times and resulting operational costs of the transportation services in creating bids and setting appropriate bidding prices is needed. In reality, carriers may exhibit different bidding behaviors in participating in combinatorial auctions. Some carriers are conservative while some may be aggressive at bidding on train slots. It could be useful to embed this bidding behavior in the bid set construction approaches to reflect these real-world concerns.

The incorporation of various bidding behaviors in the bid set construction process may require more complicated bidding language to express bids. However, the simpler the bidding language, the easier it is to solve WDP. Bidding language that can be employed in efficiently creating competitive bid sets for carriers with varying bidding behaviors that are both simple and expressive is desired.

Bid allocation among carriers can be a formidable task if the WDP intends to allocate the track capacity within an actual network (i.e REORIENT network) that involves a large number of time intervals and track segments. In addition, bidding languages proposed in this dissertation used in constructing bid sets may also increase the complexity of the WDP, because the number of required constraints may increase with number of O-D pairs. It is possible that the WDP will be intractable. An efficient heuristic that is specially designed for solving the WDP may be needed.

References

1. Affuso, L., 2003. Auctions of rail capacity. Utilities Policy, Vol. 11, pp. 43-46.

2. Aggarwal, A. K., Oblak, M., Vemuganti, R. R., 1995. A heuristic solution procedure for multicommodity integer flows. Computers & Operations Research, Vol. 22, No.10, pp.1075-1087.

3. Ahuja, R., Magnanti, K., L., Orlin, J. B., 1993. Network flows: theory, algorithms, and applications. Prentice Hall, Englewood Cliffs, NJ.

4. Arcot, V. C., Caprara, A., D'Ambrosio, C., Dong, J., Kozuki, A., Kuo, A., Lu, C.-C., Mahmassani, H., Malaguti, E., Martello, S., Miller-Hooks, E., Nair, R., Toth, P., Zhang, K., 2007. Modelling network freight transport for demonstration of seamless international rail freight services, Technical Report, University of Maryland.

5. Arnold, P., Peeters, D., Thomas, I., 2004. Modelling a rail/road intermodal transportation system. Transportation Research Part E, Vol. 40, No. 3, pp. 255–270.

6. Assad, A. A., 1980. Models for rail transportation. Transportation Research Part A, Vol. 14, pp. 205-220.

7. Ball, M. O., Hoffman, R. L., Knorr, D., Wetherly, J., 2000. Assessing the benefits of collaborative decision making in air traffic management. Proceedings of 3rd USA/Europe Air Traffic Management R & D Seminar.

8. BANVERKET, 2006. Swedish Network Statement.

9. Barnhart, C., Hane, C. A., Vance, P. H., 2000. Using branch-and-price-and-cut to solve origin-destination integer multicommodity flow problems. Operations Research, Vol. 48, No. 2, pp. 318-326.

10. Berzins, W. E., Dhavala, M. D., 1998. Trust: impact upon collaborative decision making. Journal of Management Engineering, Vol. 4, No. 4, pp. 320-324.

11. Bourbeau, B., Crainic, G. T., Gendreau, M., Robert, J., 2005. Design for optimized multi-lateral multi-commodity markets. European Journal of Operational Research, Vol. 163, No. 2, pp. 503-529.

12. Brannlund, U., Lindberg, P. O., Nou, A., Nillson, J. E., 1998. Railway timetabling using Lagrangian relaxation. Transportation Science, Vol. 32, No.4, pp. 358-369.

13. Brewer, P. J., Plott, C. R., 1996. A binary conflict ascending (BICAP) mechanism for the decentralized allocation of the right to use railroad tracks. International Journal of Industrial Organization, Vol. 14, pp. 857-886.

14. Brunekreeft, G., 2001. A multiple-unit, multiple-period auction in the British electricity spot market. Energy Economics, Vol. 23, pp. 99-118.

15. Brunetta, L., Conforti, M., Fischetti, M., 2000. A polyhedral approach to an integer multicommodity flow problem. Discrete Applied Mathematics, Vol. 101, pp. 13-36.

16. Carey, M., 1994a. A model and strategy for train pathing with choice of lines, platforms and routes. Transportation Research Part B, Vol. 28, No. 5, pp. 333-353.

17. Carey, M., 1994b. Extending a train path model from one-way to two-way track. Transportation Research Part B, Vol. 28, No. 5, pp. 395-400.

18. Carey, M., Lockwood, D., 1995. A model, algorithms and strategy for train pathing. Journal of the Operational Research Society, Vol. 46, No.8, pp. 988-1005.

19. Carlson, P. M., 2000. Exploiting the opportunities of collaborative decision making: a model and efficient Solution algorithm for airline use. Transportation Science, Vol. 34, No. 4, pp. 381-393.

20. Cattrysse, D., Salomon, M., Kuik, R., Wassenhove, L. N. V., 1993. A dual ascent and column generation heuristic for the discrete lot sizing and scheduling problem with setup times. Management Science, Vol. 39, No.4, pp. 477-487.

21. Chang, K., Howard, K., Oiesen, R., Shiler, L., 2001. Enhancements to the FAA ground delay program under collaborative decision making. Interfaces, Vol. 31, No. 1, pp. 57-76.

22. Chen, J., Chen, X., Song, X., 2002. Bidder's strategy under group-buying auction on the internet. IEEE Transactions on Systems, Man, and Cybernetics, Part A, Vol. 32, No. 6, pp. 680-690.

23. Chen, J., Chen, X., Kauffman, R. J., Song, X., 2006. Cooperation in group-buying auctions. Proceedings of the 39th Annual Hawaii International Conference on System Sciences.

24. Cheng, Y., 1998. Hybrid simulation for resolving resource conflicts in train traffic rescheduling. Computers in Industry, Vol. 35, No. 3, pp. 233-246.

25. Chih, C. K., Bodden M. P., Hornung M., Kornhauser, A. L., 1990. Routing and inventory logistic system (RAILS): a heuristic model for optimally managing intermodal double stack trains. Journal of Transportation Research Forum, Vol. 31, No. 1, pp. 50-62.

26. Choi, J. H., Han, I., 2007. Combinatorial auction based collaborative procurement. Journal of Computer Information Systems, Vol.47, No.3, pp.118-127.

27. Choong, S. T., Cole, M. H., Kutanoglu, E., 2000. Empty container management for intermodal transportation networks. Transportation Research Part E, Vol. 38, No. 6, pp. 423–438.

28. Cordeau, J. F., Toth, P., Vigo, D., 1998. A survey of optimization models for train routing and scheduling. Transportation Science, Vol. 32, No. 4, pp. 380-404.

29. Corry, P., Kozan, E., 2006. An assignment model for dynamic load planning of intermodal trains. Computers & Operations Research Vol. 33, pp. 1-17.

30. Costa, M. C., Monclar, F.-R., Zrikem, M., 2002. Variable neighborhood decomposition search for the optimization of power plant cable layout. Journal of Intelligent Manufacturing, Vol. 13, Iss. 5, pp. 353-365.

31. Crainic, T. G., 1984. A comparison of two methods for the tactical planning in rail freight transportation, Operational Research'84, J.P. Brans (Ed.), North-Holland, Amsterdam, pp. 707-720.

32. Crainic, T. G., Gendreau, M., Farvoloden, J. M., 2000. A simplex-based tabu search method for capacitated network design. Informs Journal on Computing, Vol. 12, No. 3, pp. 223-236.

33. Cramton, P., Shoham, Y., Steinberg, R., 2006. Combinatorial Auctions, MIT Press.

34. Desrosiers, J., Soumis, F., Desrochers, M., 1984. Routing with time windows by column generation. Networks, Vol. 14, No. 4, pp. 545-565.

35. Dorfman, M. J., Medanic, J., 2004. Scheduling trains on a railway network using a discrete event model of railway traffic. Transportation Research Part B, Vol. 38, No. 1, pp. 81-98.

36. Elmaghraby, W. J., 2005. Multi-unit auctions with complementarities: issue of efficiency in electricity auctions. European Journal of Operational Research, Vol. 166, No. 2, pp. 430-448.

37. EUROSTAT, 2007. Road share of inland freight transport. Available online: http://ec.europa.eu/eurostat.

38. Evans, J. R., 1980. Reducing computational effort in detecting integral multicommodity networks. Computers & Operations Research, Vol. 7, No. 4, pp. 261-265.

39. Ferreira, L., 1997. Planning Australian freight transport rail operations: an overview. Transportation Research Part A, Vol. 31, No. 4, pp. 335-348.

40. Fundenberg, D., Tirole, J., 1991. Game Theory, MIT Press.

41. Ghoseiri, K., Szidarovszky, F., Asgharpour, M. J., 2004. A multi-objective train scheduling model and solution. Transportation Research Part B, Vol. 38, No. 10, pp. 927-952.

42. Gibson, S., 2003. Allocation of capacity in the rail industry. Utilities Policy, Vol. 11, pp. 39-42.

43. Godfrey, G. A., Hellings, C., Knutsen, A., 2004. A multiagent framework for collaborative airlift planning using commercial air assets. Mathematical and Computer Modeling, Vol. 39, pp. 885-896.

44. Gorman, M. F., 1998. An application of genetic and tabu searches to the freight railroad operating plan problem. Annals of Operation Research, Vol. 78, pp. 51-69.

45. Groothedde, B., Ruijgrok, C., Tavasszy, L., 2005. Towards collaborative, intermodal hub networks: a case study in the fast moving consumer goods market. Transportation Research Part E, Vol. 41, No. 6, pp. 567–583.

46. Haghani, A. E., 1987. Rail freight transportation: a review of recent optimization models for train routing and empty car distribution. Journal of Advanced Transportation, Vol. 21, No. 2, pp. 147-172.

47. Hallowell, F., Harker, P. T., 1996. Predicting on-time line-haul performance in scheduled railroad operations. Transportation Science, Vol. 30, No. 4, pp. 364-378.

48. Higgins, A., Kozan, E., Ferreira, L., 1996. Optimal scheduling of trains on a single line track. Transportation Research Part B, Vol. 30, No. 2, pp. 147-161.

49. Hogan, W. W., 1992. An efficient concurrent auction model for firm natural gas transportation capacity. Information Systems & Operational Research, Vol. 30, No. 3, pp. 240-256.

50. Holmberg, K., Yuan, D., 2000. A lagrangian heuristic based branch-and-bound approach for the capacitated network design problem. Operations Research, Vol. 48, No. 3, pp. 461-481.

51. Horty, J., Pollack, M., 2001. Evaluating new options in the context of existing plans. Artificial Intelligence, Vol. 127, pp. 199 – 220.

52. Hu, T. C., 1963. Multicommodity network flows. Operations Research, Vol. 11, pp. 344-360.

53. Hunsberger. L., 2001. Generating bids for group-related actions in the context of prior commitments. In Proceedings of the Agent Theories, Architectures and Languages.

54. Hunsberger, L., Grosz, B., 2000. A CA for collaborative planning. Proceedings of the Fourth International Conference on Multi-Agent Systems.

55. Huntley, C. L., Brown, D. E., Sappington, D. E., Markowicz, B. P., 1995. Fright routing and scheduling at CSX transportation. Interfaces, Vol. 25, No. 3, pp. 58-71.

56. Ito, T., Fukuta, N., Shintani, T., Sycara, K., 2000. A multi-agent support system for cooperative bidding in multiple auctions.

57. Jain, R. S., 2001. Spectrum auctions in India: lessons from experience. Telecommunications Policy, Vol. 25. pp. 671–688.

58. Jovanovic, D., Harker, P. T., 1991. Tactical scheduling of rail operations: the scan I system. Transportation Science, Vol. 25, No. 1, pp. 46-64.

59. Kennington J. L., 1978. A survey of linear cost multicommodity network flows. Operations Research, Vol. 26, No. 2, pp. 209-236.

60. Klemperer, P., 2002. How (not) to run auctions: The European 3G telecom auctions. European Economic Review, Vol. 46, pp, 829-845.

61. Kraay, D., Harker, P., Chen, B., 1991. Optimal pacing of trains in freight railroads: model, formulation and optimization. Operations Research, Vol. 39, No.1, pp. 82-99.

62. Kraay, D. R., Harker, P., 1995. Real-time scheduling of freight railroads. Transportation Research Part B, Vol. 29, No. 3, pp. 213-229.

63. Kraft, E. R., 1987. A branch and bound procedure for optimal train dispatching. Journal of Transportation Research Forum, Vol. 28, pp. 263-276.

64. Krajewska, M. A., Kopfer, H., 2006. Collaborating freight forwarding enterprises. OR Spectrum, Vol. 28, pp. 301-317.

65. Krishna V., 2002. Auction theory. Academic Press.

66. Kwon, O. H., Martland, C. D., Sussman, J. M., 1998. Routing and scheduling temporal and heterogeneous freight car traffic on rail networks. Transportation Research Part E, Vol. 34, No. 2, pp. 101-115.

67. Lavoie, S. M., Minoux, M., Odier, E., 1988. A new approach for crew pairing problems by column generation with an application to air transport. European Journal of Operational Research, Vol. 35, No. 1, pp. 45-58.

68. Li, L., Tayur, S., 2005. Medium-term pricing and operations planning in intermodal Transportation. Transportation Science, Vol. 39, No. 1, pp. 73-86.

69. Liu, C. M., 1997. Network dual steepest-edge methods for solving capacitated multicommodity network problems. Computers Industrial Engineering, Vol. 33, pp. 697-694.

70. Macharis, C., Bontekoning, Y. M., 2004. Opportunities for OR in intermodal freight transport research: A review. European Journal of Operational Research, Vol. 153, No. 2, pp. 400–416.

71. Mahmassani, H. S., Zhang, K., Dong, J., Lu, C.-C., Arcot, V. C., Miller-Hooks, E., 2007. A dynamic network simulation-assignment platform for multi-product intermodal Freight Transportation Analysis, presented at the 86th Annual Meeting of the Transportation Research Board, and forthcoming in Transportation Research Record: Journal of the Transportation Research Board.

72. Martello, S., Toth, P., 1990. Knapsack problems: algorithms and computer implementations. John Wiley & Sons.

73. Mamer, J. W., McBride, R. D., 2000. A decomposition-based pricing procedure for large-scale linear programs: an application to the linear multicommodity flow problem. Management Science, Vol. 46, No. 5, pp. 693-709.

74. Matsuo, T., Ito, T., Shintani, T., 2005. A volume discount-based allocation mechanism in group buying. Proceedings of the 2005 International Workshop on Data Engineering Issues in E-commerce, pp. 59-70.

75. Mees, A. I., 1991. Railway scheduling by network optimization. Mathematical and Computer Modeling, Vol. 15, No. 1, pp. 33-41.

76. Melody, W. H., 2001. Spectrum auction and efficient resource allocation: Learning from the 3G experience in Europe. Info, Vol. 3, No. 3, pp. 5–10.

77. Milgrom, P., 2004. Putting auction theory to work. Cambridge.

78. Nair, R., Miller-Hooks, E., Mahmassani, H., Arcot, V., Kuo, A., Zhang, K., Kozuki, A., 2007. From sea to shining sea: assessing the market potential for international rail-based intermodal services in Europe, forthcoming in Transportation Research Record.

79. Newbery, D. M., 2003. Network capacity auctions: promise and problems. Utilities Policy, Vol. 11, pp. 27-32.

80. Newman, A. L., Yano, C. A., 2000. Scheduling direct and indirect trains and containers in an intermodal setting. Transportation Science, Vol. 34, No. 3, pp. 256-270.

81. Nilsson, J. E., 1999. Allocation of track capacity: Experimental evidence on the use of priority auctioning in the railway industry. International Journal of Industrial Organization, Vol. 17, pp.1139–1162.

82. Nisan, N., 2000. Bidding and allocation in combinatorial auctions. Working paper. Presented at the NWU Microeconomics Workshop.

83. Nou, A., 1997. Railway timetabling - Lagrangian Heuristics, Technical report TRITA/MAT-97-OS12, Royal Institute of Technology, Stockholm, Sweden.

84. Nozick, L. K., Morlok, E. K., 1997. A model for medium-term operations planning in an intermodal rail-track service Transportation Research Part A, Vol. 31, No. 2, pp. 91-107.

85. Odijk, M. A., 1996. A constraint generation algorithm for the construction of periodic railway timetables. Transportation Research Part B, Vol. 30, No. 6, pp. 455-464.

86. Parkes, D. C., 2001. Iterative Combinatorial Auctions: Achieving Economic and Computational Efficiency, Ph.D. dissertation, University of Pennsylvania.

87. Petersen, E. R., Taylor, A. J., 1982. A structured model for rail line simulation and optimization. Transportation Science, Vol. 16, No. 2, pp. 192-206.

88. Petersen, E. R., Taylor, A. J., Martland, C. D., 1986. An introduction to computer assisted train dispatch. Journal of Advanced Transportation, Vol. 20, No. 1, pp. 63-72.

89. Plott, C. R., Salmon, T. C., 2004. The simultaneous, ascending auction: dynamics of price adjustment in experiments and in the UK3G spectrum auction. Journal of Economic Behavior & Organization, Vol. 53, pp. 353–383.

90. REORIENT Consortium, 2007. Demand and supply structures for intermodal (rail-based) and single modal (all truck) freight supply solutions, Deliverable 5.1, https://www.reorient.org.uk/.

91. REORIENT Consortium, 2007. Demand and supply structures for intermodal (rail-based) and single modal (all truck) freight supply solutions, Deliverable 6.1, https://www.reorient.org.uk/.

92. RHK, 2006. Finnish Network Statement, Finnish rail administration.

93. Schneur, R. A., Orlin, J. B., 1998. A scaling algorithm for multicommodity flow problems. Operations Research, Vol. 46, No. 2, pp. 231-246.

94. Sherali, H. D., Staats, R. W., Trani, A. A., 2003. An airspace planning and collaborative decision-making model: part I- probabilistic conflicts, workload, and equity considerations.

95. Shetty, B. R., Muthukrishnan, R., 1990. A parallel projection for the multicommodity network model. Journal of Operation Research Society, Vol. 41, No, 9, pp. 837-842.

96. Song, J. Regan, A. C., 2004. An auction based collaborative carrier network. Transportation Research, Part E, under review.

97. Szpigel, B., 1973. Optimal train scheduling on a single line railway. Operations Research, 72, 344-351.

98. Strandenes, P. S., Wolfstetter, E., 2005. Efficient (re-)scheduling: An auction approach. Economics Letters, Vol. 89, Issue, 2, pp. 187-192.

99. TRANSTOOLS Deliverable 2: WP 2. Available online: http://www.inro.tno.nl/.

100. Triki, O., Beraldi, P., Gross, G., 2005. Optimal capacity allocation in multi-auction electricity markets under uncertainty. Computers & Operations Research, Vol. 32, pp.201–217.

101. Vemuganti, R. R., Oblak, M., Aggarwal, A., 1989. Network models for fleet management. Decision Science, Vol. 20, pp. 182-197.

102. Von Hohenbalken, B., 1977. Simplicial decomposition in nonlinear programming algorithms. Mathematical Programming, Vol. 13, pp. 49-68.

103. Vossen, T., Ball, M.O., 2005. Slot trading opportunities in collaborative ground delay programs. Transportation Science, Vol. 40, No.1, pp. 29-43.

104. Vossen, T., Ball, M.O., 2006. Optimization and mediated bartering models for ground delay programs. Naval Research Logistics, Vol. 53, No. 1, pp. 75-90.

105. Xia, M., Koehler, G. J., Whinston, A. B., 2004. Pricing combinatorial auctions. European Journal of Operational Research, Vol. 154, No. 1, pp. 251–270.

106. Xia, M., Stallaer, J., Whinston, A. B., 2005. Solving the combinatorial double auction problem, European Journal of Operational Research, Vol. 164, No. 1, pp. 239–251.

107. Zhang, K., Nair, R., Mahmassani, H., Miller-Hooks, E., Kuo, A. Y., Arcot, V. C., Dong, J., Lu, C.-C., Kozuki, A., 2007. Application and validation of a dynamic freight simulation-assignment model to a large-scale intermodal rail network: the Pan-European case. Transportation Research Board, accepted for publication.