THE HONG KONG POLYTECHNIC UNIVERSITY

Department of Electrical Engineering

Modelling Negotiations in Open Railway Access Market for Resource Allocation

by

Chi Wai Tsang, B.Eng (Hons)

A thesis submitted in partial fulfilment of the requirements for the Degree of Doctor of Philosophy

September 2006

Certificate of Originality

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgement has been made in the text.

Signature: ____________________

Name: Chi Wai Tsang

Synopsis

An open railway access market usually consists of an infrastructure provider (IP) and a group of train service providers (TSPs). Through disintegration and distribution, the managerial responsibilities on railway resources are allocated to these stakeholders. To harmonise the interrelated resource allocation processes, negotiation among the stakeholders is an important and inevitable process to resolve the operational conflicts. This study aims to develop a software platform to enable such negotiations and investigate the behaviour of the stakeholders in negotiations.

To address the distributed nature of the stakeholders and their behaviour during negotiation, a Multi-Agent System for Open Railway Access Market (MAS-ORAM) is established. MAS-ORAM is a virtual market where each stakeholder is represented by a software agent. With this system, the study focuses on modelling three major negotiations among the stakeholders (i.e. IP vs. TSP, IP vs. multiple TSPs, and TSP vs. TSP). Such representations of the open market and the subsequent study on the interactions between the stakeholders in railway management, particularly for open access markets, are the novelty of this research work.

To facilitate rational decision-making by agents, not only has the study employed algorithms of standard optimisation techniques, but also of artificial intelligence approaches. The former includes Branch-and-Bound and Lemke’s Complementary Pivoting algorithms, while the latter involves the solving of a Prioritised Fuzzy

Constraint Satisfaction problem. In addition, two policies on sequencing multiple negotiations by an IP and three negotiation strategies for TSPs have been devised. The performances of these algorithms and negotiation behaviours have been thoroughly examined through extensive simulation studies and statistical analysis.

Simulation results have shown that software agents can be set up to represent railway stakeholders of different operation objectives. During negotiation, these agents exhibit rational behaviour in making concession. Results have also confirmed that the setup of MAS-ORAM is able to derive Pareto-optimal resource plans. The system, with proper enhancements and adaptation to a specific railway market, may thus be deployed by the open railway access market stakeholders as an analytical tool before the actual negotiations are conducted in practice. The research also demonstrates the feasibility of applying agent modelling in railway management.

Publications Arising from the Thesis

Refereed International Journal Publications

Tsang, C.W. and Ho, T.K. (2006) ‘Conflict resolution through negotiation in railway open access market - a multi-agent system approach’, *Transportation Planning and Technology*, **29**, (3), pp. 157-182

Tsang, C.W. and Ho, T.K. (2006) ‘Optimal track access right allocation for agent negotiation in open railway market’, *IEEE Transactions on Intelligent Transport Systems*, (Under Revision)

Tsang, C.W. and Ho, T.K. (2006) ‘Train schedule coordination at interchange station in an open railway market through agent negotiation’, *IEEE Transactions on Intelligent Transport Systems*, (Under Revision)

Refereed International Conference Publications

Tsang, C.W. and Ho, T.K. (2004) ‘A prioritised fuzzy constraint satisfaction approach to model agent negotiation for railway scheduling’, *Proceedings in IEEE International Conference on Machine Learning and Cybernetics 2004* (ICMLC 2004), Shanghai, China, [CD-ROM]

Tsang, C.W. and Ho, T.K. (2004) ‘Modelling issues on the railway resource management process using multi-agent system’, *Proceedings in the 27th Australasian Transport Research Forum* (ATRF’04), Adelaide, Australia, [CD-ROM]

Tsang, C.W. and Ho, T.K. (2006) ‘Schedule coordination through negotiation between train service providers in an open railway access market’, *Proceedings in 2006 IEEE International Conference on Systems, Man, and Cybernetics* (SMC 2006), Taipei, Taiwan, [CD-ROM]

Acknowledgements

I would like to sincerely express my great appreciation to my supervisor, Dr Mark T. K. Ho, for his constant guidance and indispensable advice. His kind consideration, support and supervision throughout the study have made the research a fruitful and wonderful learning experience.

I would like to thank Prof Luis Ferreira at the Queensland University of Technology and all the anonymous referees for their invaluable comments and suggestions. Their keen interest on the work has provided much encouragement for the pursuit of the research.

Additionally, I would like to thank my colleagues, Dr K. K. Wong and Mr S. K. Chen, for their willingness to share their expertise in railway engineering and computer simulation. I would also like to acknowledge the continuous warm care and concern received from my family, brothers and sisters at church, and friends.

Contents

List of Figures

List of Tables

List of Abbreviations

ACC	Agent Communication Channel
AI	Artificial Intelligence
AMS	Agent Management System
ANN	Artificial Neural Network
BDI	Belief-Desire-Intention
BNB	Branch-and-Bound
BSBP	Buyer and Seller Behaviour Protocol
CGC	Congestion Charge
CPU	Capacity Utilisation
CNP	Contract Net Protocol
CTP	Commencement Time Period
DAI	Distributed Artificial Intelligence
DF	Directory Facilitator
DPS	Distributed Problem Solving
FCFS	First-Come-First-Serve
FIPA	Foundation for Intelligent Physical Agents
FIPA ACL	FIPA Agent Communication Language
FLBC	Formal Language for Business Communication
FIPA-OS	FIPA Open Source
FSM	Finite State Machine
GA	Genetic Algorithm
GUI	Graphical User Interface
HW2PF	Highest-Willingness-to-Pay-First
IP	Infrastructure Provider
JADE	Java Agent DEvelopment Framework

JVM	Java Virtual Machine
KIF	Knowledge Interchange Format
KQML	Knowledge Query and Manipulation Language
LCPA	Lemke’s Complementary Pivoting Algorithm
LGPL	Lesser General Public License
MAS	Multi-agent Systems
MAS-ORAM	Multi-agent System for Open Railway Access Market
MBSBP	Multiple Buyers and Seller Behaviour Protocol
MSP	Maintenance Service Provider
PDC	Peak Demand Charge
PFCS	Prioritised Fuzzy Constraint Satisfaction
RSP	Rolling Stock Provider
SL	Semantic Language
TAC	Track Access Charge
TEC	Traction Energy Charge
TUC	Track Usage Charge
TSP	Train Service Provider
WAN	Wide Area Network
XML	Extensible Markup Language

Chapter 1

Introduction

1.1. Railway Resource Allocation in Open Access Markets

Modern railway resource management has been embracing new opportunities and challenges ever since the introduction of open access markets. In conventional railway markets, a railway is solely owned by a single entity. The owner of the railway is responsible for both the management of infrastructure and the operation of train services. This kind of integrated management usually gives rise to railway monopoly which demands strong interference from local governments for regulations. The burden of heavy regulations on pricing and service provision usually leads to low efficiency in operation, resulting in poor adaptation to market demand. However, in open access markets, the responsibilities of infrastructure provision and train operation are distributed to independent stakeholders. This has led to an infrastructure provider (IP) selling track capacity to a group of competing train service providers (TSPs). By restructuring the conventional railway markets through disintegration, the regulatory agencies anticipate improvement on the operational efficiency in their railway markets so that rail transportation is more responsive to market demand.

Railway restructuring has created at least two new and important problems for railway resource management. Firstly, since the TSPs can only operate train services on the permanent ways with the permission from the IP, it is necessary for them to negotiate with the IP to obtain the appropriate track access rights. Secondly, the independently managed TSPs may also wish to coordinate their train schedules to facilitate passenger transfer at an interchange station so that they reap the benefit of an increase in passenger demand. In both problems, negotiation plays a key role in steering the stakeholders with various objectives towards a mutually acceptable agreement. In other words, whether railway resources can be allocated in an efficient manner depends heavily on the negotiation behaviour of stakeholders.

1.2. Objectives of Thesis

Currently, most studies aim to analyse the existing open markets according to the observed financial outcomes or train operational performances (Rothengatter, 1991; Godward, 1998; Harris, 1999; Mizutani, 1999; Shaw 2001; Watson, 2001; Crompton & Jupe, 2003). However, there has been little research devoted to performance evaluation of the stakeholders by appropriate modelling and simulation. In fact, there is even a lack of study to examine the requirements and feasibility of adopting a simulation approach. As a consequence, the main objectives of this thesis are to develop a software tool that is capable of simulating negotiation and examine the applicability of such tool in assisting decision-making in railway resource planning.

The goals of this study and their relationships are depicted in Fig. 1.1. Having identified the key modelling issues on the negotiations involved in open railway access markets, a Multi-Agent System for Open Railway Access Market (MAS-ORAM) is established. MAS-ORAM is a virtual market where each stakeholder is represented by a software agent. The virtual market enables the modelling of three negotiations (but not restricted to three) in open railway markets. The first negotiation, IP-TSP, is the core transaction in an open market and it occurs between an IP agent and a TSP agent

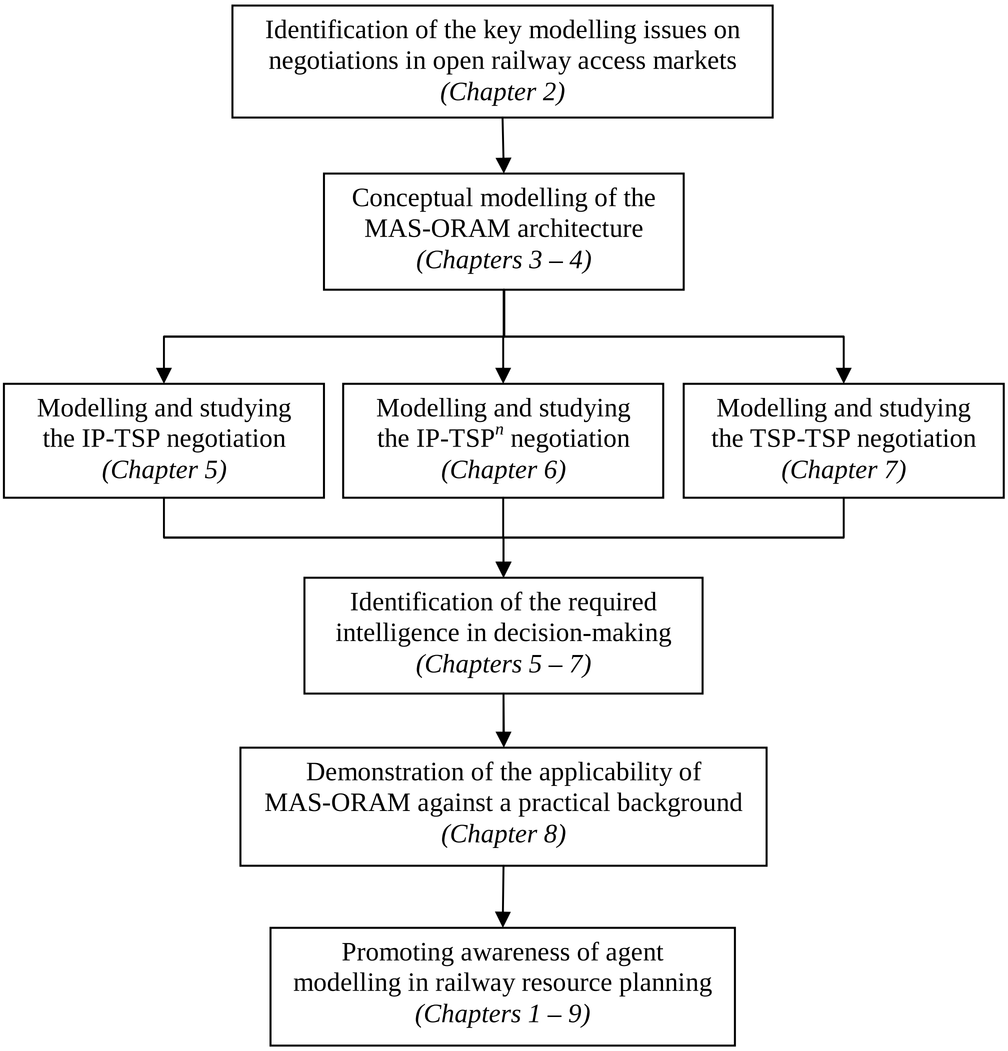


Fig. 1.1. Objectives and structure of thesis

for reaching an agreement on track access rights. The second one, IP-TSP^n, is an extension to the first negotiation and it requires the IP agent to grant track access rights to a set of n TSP agents. The last negotiation, TSP-TSP, involves two TSP agents exploring the possibility on coordinating their train schedules at an interchange station. During the course of modelling and studying these negotiations, there is a further need to identify the intelligence in making rational decisions so that the software agents are enabled to accomplish their designated tasks. In addition to these objectives, by demonstrating the usefulness of MAS-ORAM in assisting resource planning in open access markets, it is also intended to promote better understanding and enhance awareness of agent modelling in railway resource planning and management.

1.3. Contributions to Research

The contributions of this work to research are in twofold. Firstly, the study has associated two relatively isolated research topics together. Railway resource management problems are traditionally studied by central decision-making models (Ghosh, 1999) while distributed models, such as agent modelling, are seldom employed. On the other hand, most of the applications involving multi-agent systems are found in distributed sensing, data mining and e-commerce (Wooldridge, 2002). The idea of employing agent modelling in open railway access markets therefore initiates a new and valuable topic for research.

Secondly, the work in the thesis has demonstrated the usefulness of a range of decision-making techniques and negotiation protocols in modelling the agent transactions conducted on MAS-ORAM. In an IP-TSP negotiation, the decision-making problem of the TSP is modelled as a Prioritised Fuzzy Constraint Satisfaction (PFCS) problem, and a rule-based resolution technique is used to optimise the TSP's satisfaction on track access rights. On the other hand, a constrained optimisation problem is formulated to represent the objective of the IP, and a Branch-and-Bound (BNB) algorithm is proposed to derive the optimal resource plan for the IP. This negotiation is enabled by the Buyer and Seller Behaviour Protocol (BSBP) which is capable of steering the agents to reach the Pareto-optimal deal. To allow multilateral negotiation in the IP-TSP^n negotiation, BSBP is extended to Multiple Buyers and Seller Behaviour Protocol (MBSBP). The IP agent handles the multiple negotiations in a sequential manner by employing either First-Come-First-Serve (FCFS) or Highest-Willingness-to-Pay-First (HW2PF) as the sequencing policy. In order to evaluate the performance of the two sequencing policies with respect to the benefits of the IP and the quality of train services, a statistical analysis based on simulation is conducted. In the third negotiation problem of the TSP-TSP transaction, a simple protocol between two TSP agents is introduced, allowing the agents to propose, accept or reject offers. Three negotiation strategies (S_{po}, S_{min} and S_{max}) are proposed for the TSP agent. These strategies involve using the Lemke's Complementary Pivoting Algorithm (LCPA) incorporated with their respective local searching techniques to generate potential offers.

1.4. Contributions to Railway Resource Management

The simulation of negotiation activities on MAS-ORAM is intended to be adopted for strategic and tactical planning in open railway access markets. Resource planning in railways is often divided into three time horizons: strategic, tactical and operational (Watson, 2001). In strategic planning, decision makers concentrate on whether changes to the current infrastructure is necessary (e.g. adding sidings at the bottlenecks, constructing a new extension line, etc.). In tactical and operational planning, the infrastructure is assumed to be unchanged. Specifically, tactical planning involves decisions on the daily frequency and duration of train services and the subsequent allocation of staff members on board the trains. On the other hand, operational planning focuses on making prompt arrangement on the real-time perturbations to the daily schedule (e.g. recovery of train services due to train delays).

With respect to strategic planning, the negotiation system may be used by the IP to evaluate various means of capacity allocation. For example, simulation can be conducted to compare different infrastructure pricing regimes or internal scheduling policies. The IP may determine whether there are better ways to improve the current capacity utilisation before considering new infrastructure constructions. The IP can therefore benefit from the reduction in the number of unnecessary and costly constructions.

Regarding to tactical planning, MAS-ORAM may be used by the stakeholders (i.e. both TSP and IP) to determine the potential gain or loss prior to commencing the actual negotiation. According to the simulation results, the stakeholders can adjust their operational objectives or requirements in order to enhance their negotiation power and enable better possibility of striking a good deal. In some circumstances, it may also allow the stakeholders to withhold from initiating a negotiation in order to eliminate the unnecessary transaction costs involved in the negotiation process.

Apart from providing a useful evaluation tool for the TSPs and IPs, MAS-ORAM can be employed by the railway regulatory authority to simulate the possible effects of any new regulation on the railway markets prior to the implementation. This reduces the risk of imposing poor or ineffective regulations that would otherwise have caused an adverse impact to the operations of the railway markets.

All contributions described above assume MAS-ORAM functions as an evaluation tool in railway resource planning. Since decision-making in practice is undoubtedly complex, the work in the thesis does not intend to replace the current human-to-human interactions with automated negotiation. However, with proper considerations on the assumptions made in the negotiation models, the proposed MAS-ORAM is expected to generate useful results to the stakeholders.

1.5. Thesis Structure

Chapter 2 reviews the background on open railway access markets. In particular, the importance of negotiation between the various railways stakeholders will be discussed. The chapter also generalises the problem nature associated with modelling open railway access markets and the negotiation between different parties.

Chapter 3 reviews the background on multi-agent systems. In addition, the common standards and software agent development toolkits for multi-agent systems are highlighted. The related agent applications in e-commerce, electricity markets and railways are also discussed.

Chapter 4 puts forward the MAS-ORAM architecture. Since the architecture is implemented using an agent development toolkit JADE, the key functions and capability of the toolkit is further elaborated. With the proposed MAS-ORAM architecture, the three negotiation problems of IP-TSP, IP-TSP^n, and TSP-TSP are formally specified.

Chapter 5 focuses on the modelling and the subsequent evaluation of the IP-TSP transaction. The mathematical models for the negotiation protocol and the decision-making processes in the IP and TSP agents are formulated. A comprehensive simulation study is set up to determine whether the agents are rational in their negotiation behaviour.

Chapter 6 extends the IP-TSP transaction model derived in Chapter 5 to model the IP-TSP^n negotiation. A statistical evaluation by simulation is proposed and conducted to compare the consequences of employing different negotiation sequences, in terms of the benefits of the IP and the quality of train services.

Chapter 7 aims at modelling and studying the TSP-TSP negotiation concerning the schedule coordination problem between two TSPs. Similar to the study in Chapter 5, mathematical models are constructed for the negotiation process and the TSP agents. This is then followed by simulation studies on a set of extreme cases to show the rationality exhibited by the agents.

Chapter 8 contains two feasibility studies on MAS-ORAM to demonstrate its potential applicability against a practical background in assisting railway resource planning in open access markets. The first hypothetical study concerns the track access rights allocation problem at the Hunter Valley rail network in Australia while the second one examines the schedule coordination at the Liverpool station in the UK. Finally, Chapter 9 summarises the findings and contributions of the thesis and suggests further works for research.

1.6. Remarks

This chapter introduces the concept of open railway access markets and the role of negotiation in reaching agreements between railway stakeholders. In order to improve the negotiation outcomes and avoid irreversible changes to open markets, there is a need of developing a software evaluation tool that enables railway stakeholders and regulatory bodies to conduct critical cost-and-benefit analysis prior to the actual negotiation. Therefore, this thesis aims to examine the modelling challenges for negotiations in open railway access markets. Based on the multi-agent system (MAS) paradigm, a virtual negotiation market, called MAS-ORAM, is proposed and the study focuses on modelling three major negotiations (IP-TSP, IP-TSP^n and TSP-TSP) among the railway stakeholders. In addition, the intelligence required by the stakeholders in making concession during negotiation is examined and the feasibility of representing the stakeholders as software agents is investigated.

Chapter 2

Railway Open Access Markets

In recent years, extensive regulatory reforms in railways have been implemented in many countries where the primary objective is to introduce intra-modal competition in rail transportation. The successful reform precedents from utilities, such as gas, electricity and telecommunication, have encouraged the formation of open railway access markets. Under this new market structure, negotiation becomes an important process for stakeholders to allocate railway resources. In this chapter, the background on the evolution of railway markets is first introduced, followed by the description of the major negotiations conducted in open access markets. Then, the benefits and obstacles for a computer simulation approach to resolve the problems generated in open railway access markets are discussed.

2.1. Types of Railway Markets

Conventional railway markets are often referred as integrated railways because the ownership of infrastructure and the operation of train services belong to the same entity. Some of these conventional railway markets have recently restructured into open access markets or third-party access markets. The background and relationship between these railway markets are discussed below.

2.1.1. Conventional Structure and Organisation

Many railways, particularly those originated from the 19th century, were established by a set of private companies. These companies constructed and possessed their own infrastructure, over which they also had the exclusive rights to operate train services. At the beginning, these railways usually served the purpose of transporting freight from its source to destination. For instance, coal might be carried from mines to ports on the coast for exports. Passenger transportation was later introduced between major cities to facilitate long-distance travel. Although these railways might be in close proximity with one another, the companies often developed their systems individually and had little intention to connect them together to form a large network (ECMT, 2001; DOTARS, 2003).

These independent developments gradually caused duplication of tracks and stations. Different railways also employed different track gauges and signalling systems. Thus, not only did railway transportation suffer from inefficiencies due to excessive capacity, but it was also in lack of system integration and interoperability. These problems eventually prompted for the nationalisation of railways, either within a country or a state, so that the fragmented systems become the property of the government or a private corporation. Control and management of the entire railway operations and developments were then centralised to a single integrated organisation (ECMT, 2001; DOTARS, 2003).

The emergence of such an integrated structure has been explained formally by transportation economic principles (Boyer, 1998; Campos & Cantos, 1999). Firstly, railway asset is considered to be indivisible as a result of the lumpy change in supply being significantly greater than the fluctuations in demand. When a railway is constructed or expanded, capacity is often created in large incremental step (or lump) relative to the unit consumption by the train services. In other words, the newly built infrastructure is capable of supporting a large increase in traffic volume. Since the fixed investment cost is recovered from the actual rise in demand, the average cost of transportation declines when traffic volume increases over the range of additional capacity (Fig. 2.1). Consequently, it is usually cheaper ($P_A < P_B$) to provide train services when the entire demand is captured by a single operator.

Secondly, railway costs are sub-additive which refers to the improved efficiency when train services are provided by a single railway rather than sharing the services by different ones. Since it is possible for many railway services (e.g. passenger and freight) to use a common infrastructure, the addition of a train on permanent ways only requires a relatively small increase in expenditure when compared with the vast investment in establishing a new system. Therefore, railways are naturally less expensive to operate by a single entity and the industry is not easily subject to intra-modal competition.

Owing to asset indivisibility and cost sub-additivity, railway was commonly considered as an incontestable (non-competitive) business (Campos & Cantos, 1999; ECMT, 2001). With the absence of competition, it was not surprising to discover that local jurisdictions had imposed regulations on pricing regimes so as to protect the

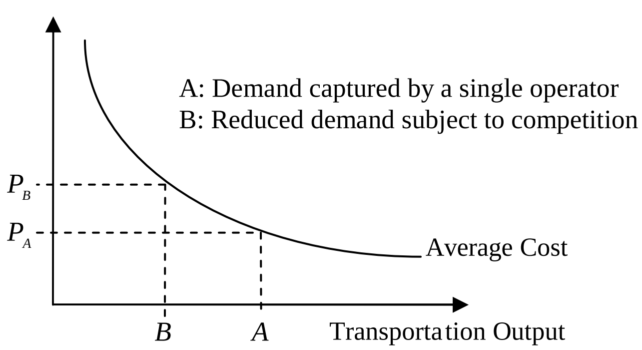


Fig. 2.1. Average cost curve

general interest of the community from excessive monopoly manipulations. Moreover, there were additional regulations on train operation as a result of the prevailing social perceptions that railways should provide a minimum transportation supply to the population and assist the urbanisation of rural areas. These social responsibilities often forced the industry to operate some unprofitable routes which greatly discouraged efficient development and provision of rail transportation.

The progressive burdening from heavy regulations eventually resulted in poor adaptability to market demands which in turn led to the degradation of quality of service and the requirement of considerable subsidisation from local governments (Boyer, 1998; Crompton & Jupe, 2003). Since the 1950s, railways have also been challenged by road transportation due to the rapid development of road networks which have provided reliable and easily accessible infrastructures for automobiles. On the other hand, similar growth in railways was largely hindered by excessive regulations. With such rapid improvements in road networks, railway was severely losing the market share to road transportation (ECMT, 1998; BTRE, 2003; EC, 2006). Railway was thus in desperate need of countermeasures to revitalise the industry, making it an efficient market-driven transportation.

2.1.2. Open Access Markets

In response to the fierce inter-modal competition, many countries have restructured their railways through deregulations (Table 2.1). These reforms were also encouraged

Table 2.1. Worldwide Deregulations in Railways

Year	Country	Principal Legislation
1987	Japan	Railway Enterprise Law
1988	Sweden	Transport Policy Act
1992	Argentina	State Reform and Public Enterprises Restructuring Law
1993	New Zealand	State-Owned Enterprises Act 1986
1994	UK	Railways Act 1993
1995	Australia	National Competition Policy

by an adjustment in economic principles on railways (Campos & Cantos, 1999; BTRE, 2003). Despite the conventional view that railway is incontestable, the recent perspective argues that competition is possible for train operations (above-rail activities) even though infrastructure provisions (below-rail activities) may prove more elusive. The barrier to intra-modal competition may therefore be lowered by allowing multiple train service operators to gain access to the infrastructure from a common provider. This forms the basis of open railway access markets.

An open railway market, in its simplest form, consists of a group of train service providers (TSPs) and an independent infrastructure provider (IP). In the UK, the ancillary services of rolling stock and maintenance provisions are also separately offered by the rolling stock leasing providers (RSPs) and the maintenance service providers (MSPs) respectively (Shaw, 2000). An open railway market therefore involves multiple stakeholders arranged as a supply-chain through which railway resources (e.g. track capacity and rolling stock) are supplied to the TSPs to allow the ultimate train service provisions to the end-consumers (Fig. 2.2).

The competing TSPs can be classified by their types of service provisions. At the

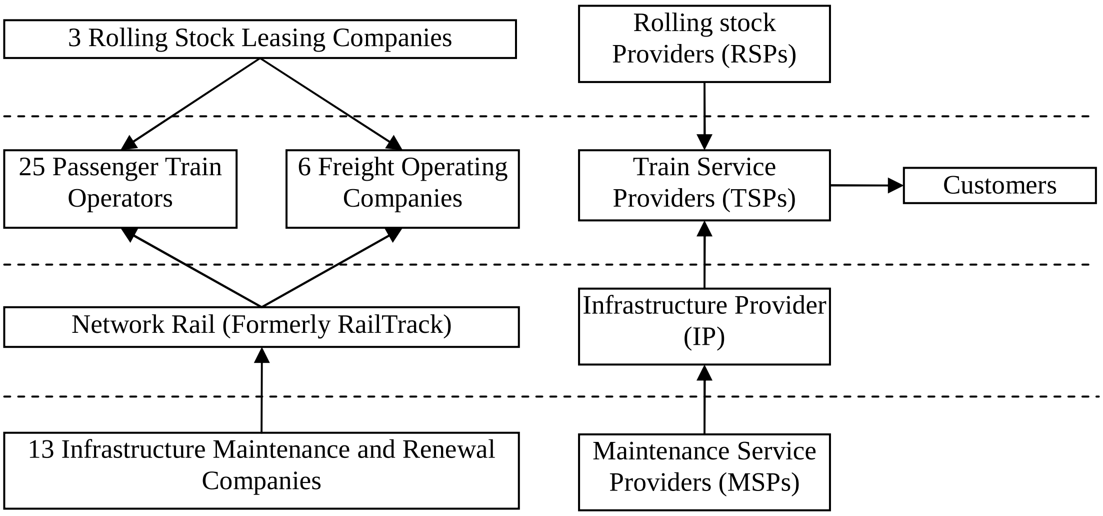


Fig. 2.2. Open access market in the UK upon privatisation

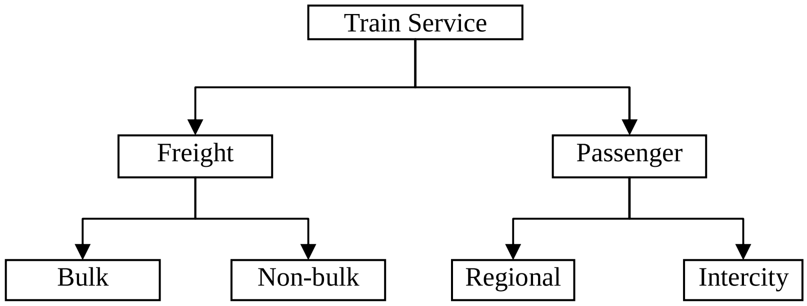


Fig. 2.3. Possible types of railway services

first level, train services can be regarded as freight or passenger (Fig. 2.3). Freight services can be further grouped by the nature of commodities being bulk (e.g. coal, petrochemicals) or non-bulk (e.g. foodstuffs, postal, parcels) (Network Rail, 2004a). On the other hand, passenger services are classified as regional or intercity according to the distances travelled. In open access markets, these rail services are now operated by different stakeholders, and on-rail and off-rail competitions are introduced into the markets.

On-rail competition refers to the competition of capacity and customers (Shaw, 2001). Regardless of the types of rail services, all TSPs are required to obtain track and station capacity from the IP as a result of their common operation on the same infrastructure. Such competition is anticipated to improve network utilisation. On the other hand, direct competition for customers between train operators usually occurs in the freight market. Its benefits are derived from offering more choices to consumers which creates pressure on the train operators to reduce expenditure and increase revenue (Jensen, 1998). However, similar competition is less apparent in passenger services. It is rare to find two passenger operators competing for identical routes because of limited transportation demand. Nevertheless, a moderate competition is still possible between partially overlapping routes and between services of different train speeds (e.g. regular and express trains).

On the contrary, off-rail competition is created from the social demand for benchmarking the quality of services among the existing operators, even if their services are running at different regions of the network (Shaw, 2001). If the operators fail to provide an acceptable level of service, the rights of operation in the network may be terminated by the local jurisdiction and acquired by other potential operators. The existing operators are therefore pressed to respond to the market demand, and preferably, develop innovative plans to explore new demand.

Despite these ideal advantages of competition, whether the widespread access reforms have resulted in a net benefit to the railways is still debatable because these reforms have concurrently generated several new challenges for the industry. Firstly, the resource allocation problem now involves multiple parties whose objectives are likely to be in conflicts. This causes addition complexity in making decisions in train planning (Watson, 2001; Gibson *et al.*, 2002; Nilsson, 2002; Gibson, 2003). Secondly, with the increased number of stakeholders, there is a corresponding rise in transaction costs (in terms of time and money) in forming contractual agreements for train operation (Campos & Cantos, 1999; ECMT, 2001). Moreover, the new market structure may create obstacles in providing seamless services and coordinated services at interchange stations due to the possible conflicts between different stakeholders (Campos & Cantos, 1999; BTRE, 2003). Finally, in regions where competition is limited, the infrastructure owner may have low investment incentives on facilities because the IP is in a weak position to bargain for a favourable rate of return (Campos & Cantos, 1999). All these problems are yet to be resolved.

2.1.3. Third-party Access Markets

Currently, only the British and Swedish systems have sought for full-scale open access reforms (ECMT, 1998; Campos & Cantos, 1999). Since open markets require the complete separation of train operations from infrastructure provision, some countries have preferred a less radical approach by allowing third-party access to an integrated network. Under this kind of reform, the incumbent owner of the infrastructure also operates train on the tracks, but independent train operators are eligible to gain access to the fixed facilities to enter the competition.

To ensure that the incumbent owner will allocate capacity to external train operators on a fair and equitable basis, the incumbent owner is required to implement a mandatory process called ring fencing. Ring fencing involves the development of an internal organisation structure so that it prevents the flows of information, personnel and inappropriate transferring of costs and revenues within the organisation of the incumbent owner (BTRE, 2003). A typical realisation of ring fencing is the formation of separate departments responsible for infrastructure provision and train service operation. Despite belonging to the same company, the train operation department is required to obtain track capacity from the infrastructure department as if it were an external operator.

Most countries given in Table 2.1, including Argentina, Australia, Japan and New Zealand, have employed the third-party access approach (Campos & Cantos, 1999; ECMT, 2001). Despite the difference in definition between open access and third-party access, these railway markets have key similarities in both market structures and operational activities. The former involves multiple stakeholders of different managerial authorities, while the latter requires the external train operators to interact with the IP for track access allocation.

2.2. Negotiations

A simple open access market composes of an IP and a set of TSPs. In addition, ancillary service providers such as RSPs and MSPs may also be present. The efficiency of railway operations greatly depends on how these stakeholders interact to produce the resource plans. In this section, the main contents and the parties involved in these interactions are described, along with the related issues and problems.

2.2.1. Infrastructure Provider vs. Train Service Provider

The interaction between an IP and a TSP is the core interaction in an open or third-party access market. The main objective of this interaction is to form a track access rights agreement, which specifies the access price and capacity allocated to the TSP. However, owing to the presence of different managerial authorities, the requirements of the stakeholders are likely to be in conflicts and it is essential to resolve these disputes during the formation of an agreement.

2.2.1.1. Railway Access Pricing

There are three principal pricing mechanisms in open access markets. Posted pricing, direct negotiation and auctioning are the proposed mechanisms to access charge setting (BTRE, 2003; ECMT, 2005). In posted pricing, charge rates are established in advance and published to the access seekers. The tariff is often composed of a basic charge in terms of the vehicle-kilometre or gross tonne-kilometre transported, and an uplift cost that is levied according to the operating characteristics (e.g. freight/passenger services and types of rolling stock). In direct negotiation, the IP and the TSPs take turns to make concessions on issues including access charge, train schedule and operating characteristics until both stakeholders agree on the terms of usage. For auctioning, capacity is pre-packaged into various sets of non-conflicting train paths to allow interested seekers to bid at their most preferable prices. The operator with the highest bid will obtain the train paths under a set of restrictions.

Apparently, posted pricing provides train operators with more certainty in managing their businesses, but the IP may fail to discriminate TSPs with different operating requirements effectively. For example, trains travelling at different speeds may be charged identically even though they have different traction energy and peak demand requirements. Conversely, services with identical speed specifications but scheduled on different traffic conditions might also be charged at the same price despite having different capacity consumption (see Section 2.2.1.2). On the other hand, direct negotiation and auctioning have better capability to distinguish operators with respect to their willingness-to-pay for rights-of-ways. Nevertheless, experience has suggested that negotiated pricing can sometimes require time-consuming and costly transactions, while auctioning has never been employed in practice because of the difficulty in devising train paths that simultaneously suit the requirements of several train operators (BTRE, 2003). These existing regimes have their merits and limitations, and the railway regulators and stakeholders are still striving for better alternatives whenever possible.

2.2.1.2. Capacity Management

Along with determining a suitable pricing regime, the IP also needs to formulate a conflict-free and preferably efficient resource allocation plan for the access seekers. Since the train operators are independently managed, they will occasionally request conflicting train paths. The IP then has the responsibility to resolve their disputes in rights-of-way.

Efficient allocation is complicated by heterogeneous traffic condition (i.e. when trains are operating with a wide range of speeds). Fig. 2.4 illustrates the effect on capacity utilisation when the traffic demand is homogeneous or otherwise. Capacity utilisation is defined as the ratio of the time taken in operating a set of trains with their minimum headways (i.e. A and B) to the time taken in travelling at their actual timetables (i.e. W) (Gibson *et al.*, 2002). Clearly, when trains running at different speeds are scheduled together, more capacity is needed to generate the same number of services ($B/W > A/W$).

In principle, the cost of additional capacity consumption may be recovered from the access charge. However, the predefined tariffs in posted pricing are unlikely to respond to the ongoing changes in relative train speeds in the competitive market. On the contrary, direct negotiation is able to provide a means to dynamically compute the associated costs of capacity utilisation and traction power supply. Therefore, the access charge can be more appropriately recovered by negotiation if a high transaction speed is available. In addition, negotiation allows the operational train speeds to be

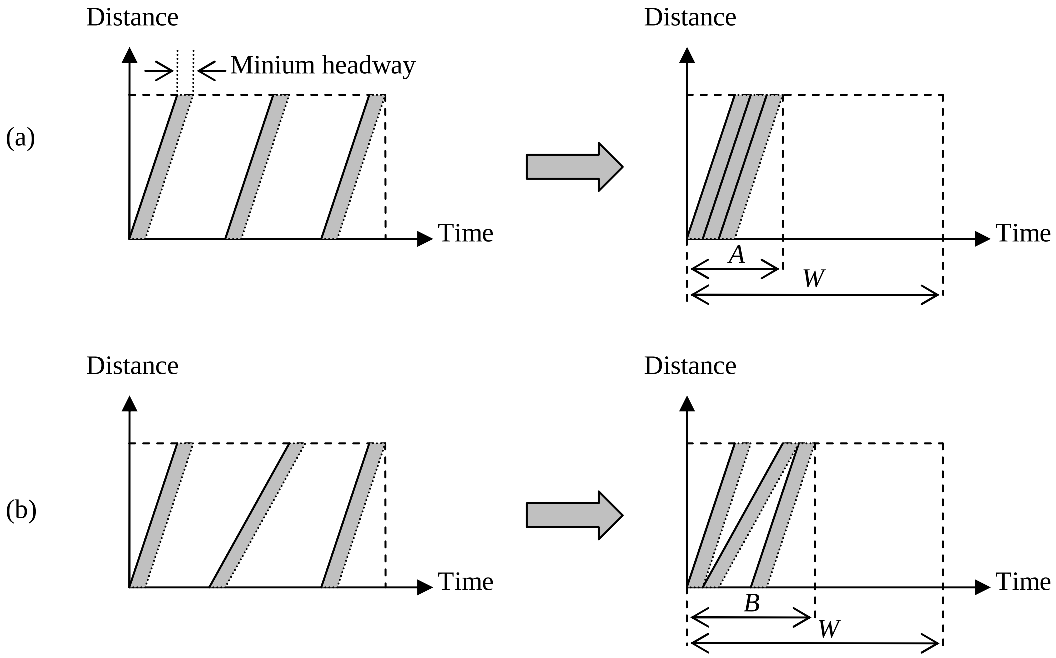


Fig. 2.4. Capacity utilisation of (a) homogeneous traffic (b) heterogeneous traffic

determined by the requirements of the access seekers. If the service providers are willing to afford a higher tariff, heterogeneous traffic may be allowed, otherwise the infrastructure provider may offer a cheaper access charge for capacity saving.

2.2.1.3. Wheel-Rail Maintenance

Disputes in maintenance costs of track and rolling stock are another type of potential conflict to be resolved (Johansson & Nilsson, 2004; Grassie, 2005). There is an indivisible relationship between rails and wheels. Poor rail quality will induce an increased rolling stock maintenance cost and vice versa. In an integrated railway, the maintenance cost on rails can be balanced with the investments on rolling stock's quality. However, with the separation of responsibilities, service providers may tend to keep maintenance on vehicles minimum so as to reduce their operating costs. In order to recover the imposed maintenance fee, the IP has to decide whether to raise the access price to reflect the actual damages on track or to restrict the use of track to better maintained rolling stock.

2.2.1.4. Coordination of Multiple Negotiations

The one-to-one negotiation between an IP and a TSP forms the core interaction in open access markets. However, since there are multiple train operators in open access markets, the IP is thus required to conduct multiple transactions before the entire network timetable can be produced. In other words, in addition to handling the individual negotiations of the TSPs, the IP needs to decide how these multiple negotiations may be managed to increase its potential benefits. This problem may be reflected by the timetabling production process in the UK (Network Rail, 2004b).

The process begins with the TSPs devising their requirements on track access (Fig.

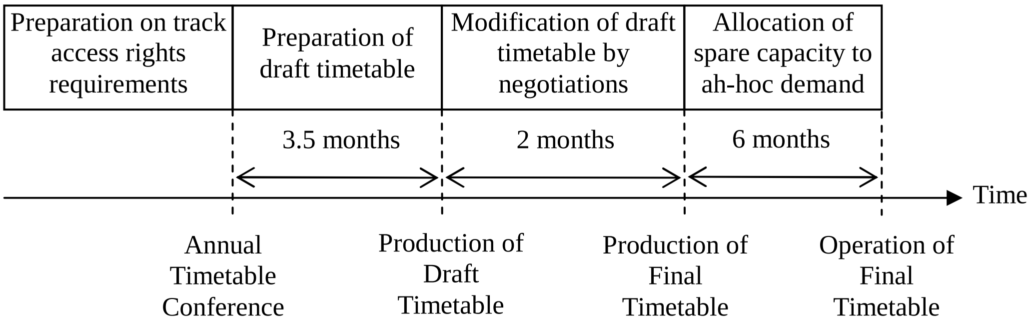
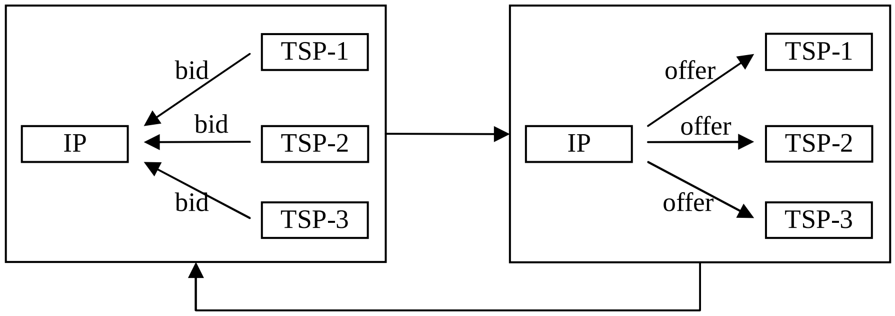


Fig. 2.5. Timetable production process in the UK

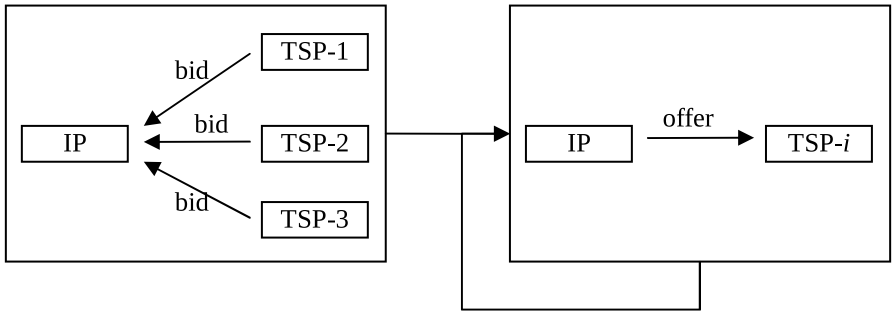
2.5). At the annual timetable conference, the train operators submit these requirements to the IP. After collecting the requests, the IP resolves the operational conflicts and produces a draft timetable. When the timetable is available, the stakeholders begin a negotiation period during which the IP and TSPs take turns to resolve further operational differences. Owing to the substantial amount of time required for formulating the new offers, the negotiation process may only have a few rounds before the final timetable is produced. Afterwards, there is a period of time before the timetable is actually put into operation. During this period, any ad-hoc services (particularly for those operated by freight operators) may fill up the spare capacity.

During the modification of the draft timetable, the IP may conduct the track capacity allocation in at least two different ways (Fig. 2.6). In the first approach (combinatorial generation), the IP may collect all offers from the TSPs and determine the optimal allocations for all access seekers simultaneously. If the TSPs decide to reject the offers produced by the IP, they can revise their bids and submit them to the IP in the next round of negotiation. The process iterates until the access agreements are secured, or the stakeholders withdraw from the negotiation.

Since train timetabling is a complex and time-consuming process (Watson, 2001), combinatorial generation may be infeasible when considering the deadline of the final timetable production. Therefore, in the second approach (sequence generation), the IP



(a) Combinatorial generation



(b) Sequence generation

Fig. 2.6. Offer generations by IP in multiple negotiations

conducts the individual negotiations in a sequential manner. This significantly reduces the complexity in decision-making. Nevertheless, the IP is required to determine the order in which the individual negotiations are to be conducted.

2.2.2. Train Service Provider vs. Train Service Provider

Another interaction in open access markets occurs between two TSPs when they decide to explore the possibility of coordinating their services at an interchange station. Schedule coordination can reduce the waiting time at an interchange station and thereby improves the attractiveness of the train services offered by both operators. The negotiation may thus lead to mutual benefits on increased transportation demand for both parties.

2.2.2.1. Seamless Services and Interoperability

As one objective of railway restructuring is to revitalise the railway industry from the continuous loss of market share to road transportation, there have been concerns to both maintain and promote further seamless services in open access markets. The availability of a direct transportation from source to destination is essential to compete with the door-to-door and just-in-time services offered by road transportation. Removing the barriers for seamless services is therefore another key issue in modern railway markets.

In the context of railway reform, seamless services can be regarded as the operation of train services involving multiple systems. This issue is of more concern in Australia and European countries, where trains are required to cross boundaries of different jurisdictions. The attention of seamless service provision can be realised from the National Competition Policy (of Australia) (BTRE, 2003) and the European Rail Directive 91/440/EEC (EC, 2006) which stated the importance of achieving interoperability between railways in different states/countries. However, four barriers have been identified in achieving full-scale interoperability between railways of different jurisdictions (Mulley & Nelson, 1999).

Firstly, technical interoperability must be resolved for interconnectivity of the physical infrastructure. This includes the implementation of a compatible track gauge, signalling system and power system. Secondly, the railways must encourage corporate interoperability which refers to the willingness and the ability of different organisations to cooperate for providing train services. The third barrier is juridical, which concerns the different legislations between states/countries. In addition, cultural differences in languages and attitudes to quality of service may also impede interoperability.

Providing solutions for these barriers is a long-term process. For example in Europe, the initial directive announced in 1991 was followed by updates and amendments in Directives 95/18/EC and 95/19/EC announced in 1995, and Directives 2001/12/EC, 2001/13/EC and 2001/14/EC published in 2001 (EC, 2006). As a result, where immediate interoperability is not feasible, the availability of coordinated train services can facilitate the transportation across regions. In addition, even if seamless services become available, coordinated services may still introduce intra-modal competition by being an alternative choice for consumers.

2.2.2.2. Schedule Coordination at Interchange Stations

A schedule coordination problem involves the adjustment of arrival and departure times of a set of trains serving different routes but sharing a common (interchange) station. When the services are coordinated, transit time of passengers is shortened. This reduces the impediment to consumers using rail transportation, hence improves revenue collection from the increased demand.

Prior to railway restructuring, the train coordination problem in an integrated railway was decided internally by train planners. At the beginning, all train services possessed their earliest commencement times and preferred journey times (station dwell times and inter-station runtimes). Schedule coordination was achieved by offsetting the commencement times of the services while the preferred journey time was kept undisturbed (Nachtigall & Voget, 1996). As the train planners had the exclusive rights to modify the commencement times of all services, minimisation of passenger waiting time for a group of train services could easily be achieved.

However, in open access markets, train coordination is complicated by the different managerial authorities. As the stakeholders only look after their internal benefits, the commencement times may not be adjusted as handily as in the integrated railways. Since both operators desire to have a schedule with minimum deviation from the earliest commencement time to reduce the idle cost on rolling stock, the stakeholders are required to conduct a negotiation to resolve the conflict. A successful deal may become possible when the revenue increased from the transport demand is greater than the idle cost generated.

2.2.2.3. Prohibition to Collusions

Besides the possible benefits of schedule coordination, it is also worth noting that during the formulation of coordinated schedules, the TSPs are not allowed to negotiate for any share of revenue generated from the increased in transportation demand (ECMT, 2001; BTRE, 2003). Otherwise the natural advantage of cost sub-additivity of railways is likely to force the weaker operators out of the competition. Regulations are therefore available to prevent collusions so as to facilitate a contestable market environment.

2.2.3. Interactions with Ancillary Service Providers

Apart from the core negotiations conducted among the IPs and TSPs, there are additional interactions in the railway market when ancillary service providers are present. For examples, MSPs may interact with the IP to acquire timeslots for performing the maintenance work, and RSPs may interact with TSP for leasing rolling stock to the TSPs.

However, these interactions are of different natures from the negotiations described thus far. Ancillary service providers are usually readily capable to satisfy the demand from their clients. For instance, a MSP will be able to conduct the maintenance work on permanent ways at the specified time defined by the IP. Similarly, a RSP will deliver the rolling stock to the TSP at the required scheduled time. As a result, these interactions are more suitably regard as off-the-shelf transactions and they involve little need of intensive negotiations to resolve their conflicts.

2.3. Research Opportunities

From the above discussions, three major problems regarding to the interactions between the stakeholders have been identified. Firstly, between an IP and a TSP, there is a need to resolve the pricing and capacity allocation problem. Next, the IP is required to develop proper means to handle the multiple transactions with the TSPs in order to establish a complete network timetable. Lastly, the TSPs may negotiate on the commencement times of their services to improve transportation demand.

In all these interactions, there is a need to derive effective mechanisms to resolve the disputes between the stakeholders. Various post-evaluations on the performance of the current railway markets have been conducted with respect to regulatory efficiency (Rothengatter, 1991; Campos & Cantos, 1999; Cantos & Maudos, 2001; ECMT, 2001), train planning process (Watson, 2001; Gibson *et al.*, 2002; Nilsson, 2002; Gibson, 2003) and accounting profiles of stakeholders (Dodgson, 1994; Godward, 1998; Harris, 1999; Mizutani, 1999; Shaw, 2001; Crompton & Jupe, 2003). Findings derived from these studies can be applied in future improvements in resource management in the open markets, which is followed by a new cycle of post-evaluations. Unfortunately, the time-consuming and costly execution of changes often hinders the actual implementation of these new findings.

2.3.1. A Computer Simulation Approach

With the advance of fast computing technologies, computer simulation is a cost-effective means to evaluate any change, however hypothetical, in a system. For example, simulation suites have been developed to study a variety of traffic control strategies according to sophisticated models of train dynamics, traction systems and power systems (Goodman *et al.*, 1998). Simulation therefore allows pre-evaluation studies and avoids irreversible changes to the physical system. It is beneficial if open markets can be modelled in an appropriate manner to assist studies in future improvements on railway operation.

A simulation model is a representation of the behaviour of a system that can be executed in a computer. Most simulation models are devised and implemented as a single (or central) computation unit which derives the expected system outputs by processing the user-specified inputs with an algorithm. However, the idea of central evaluation is inappropriate for open railway access markets due to the obstacles described as follows.

2.3.2. Major Obstacles in Central Decision-making Models

2.3.2.1. Distributed Self-interested Entities

As a result of railway reforms, resource planning is now a distributed rather than centralised problem and different stakeholders will inevitably attempt to optimise their internal benefits. In fact, these optimisations are likely to involve multiple attributes such as cost and travelling times, and they are subject to constraints derived from business (e.g. availability of rolling stock supply), engineering (e.g. maximum line speeds) and regulatory (e.g. regulated ceiling and floor prices) causes. Some of these constraints and business objectives are not revealed to other stakeholders to avoid possible loss of advantage during the business transactions.

A suitable modelling framework should therefore enable the representation of the stakeholders as separated entities with individual control over their information, decisions and actions. The framework should also allow separate local simulation models for solving the distributed multi-dimensional constrained optimisation problems.

2.3.2.2. Behavioural Modelling

Despite the isolated control over their activities, the stakeholders are still interdependent during the formulation of train timetables. In the case of negotiation, the stakeholders will attempt to persuade their negotiating partner to align with their operational objectives through bargaining. To resemble the natural process of resolving conflicts, the modelling of the coordination activities is of utmost importance. As a result, in addition to the distributed framework, there are additional requirements on modelling the interactions between the local entities. However, classical simulation models are generally not designed to capture these behaviours.

2.3.2.3. Rationality of Local Entities

In principle, solving the local optimisation problems may take the advantage of the classical simulation models. The rational decisions may be the outputs of the local models while the operational objectives of the stakeholder can be considered as the user-specified inputs. However, the choice of algorithm is complicated by the additional inputs from the responses of its interacting entities which can only be determined during runtime. These algorithms therefore require a certain degree of flexibility so that the distributed entities may decide the best actions dynamically without human interferences.

Although this study will only focus on developing agents that react to the changing environment, proactiveness is another important dimension of rationality that requires attention. If such capability is required, the local entities should be able to inspect their internal status and initiate activities (e.g. promotion of idle resources) which are consistent with their business goals in order to enhance the competitiveness of the stakeholders.

2.4. Remarks

This chapter has reviewed the background of open and third-party access markets in railways. To respond to the increasing inter-modal competition from road transportation, the structure and organisation of the railway markets in many countries have been transformed from integration to separation. Through the separation of train service provision from infrastructure provision, intra-modal competition has been introduced in the railway markets.

Under the new market structure, there are three important interactions demanding for resolutions on operational conflicts among the stakeholders. The first one occurs between a single IP and a single TSP which involves the determination of access price and capacity allocation. The second one also concerns the same issues, but the IP is required to handle multiple transactions efficiently. In the third interaction, two TSPs have to decide whether coordinating the schedules of their train services is beneficial.

Most studies on open railway access markets have aimed to analyse the current situation according to the observed outcomes from the existing reforms, but there has been little research devoted to evaluate the performance of the stakeholders by the use of modelling and simulation. In fact, there has been even a lack of study to examine the requirements and feasibility of adopting a simulation approach.

Having considered the characteristics of open markets, the idea of centralised computation may be used to capture the rationality for individual stakeholders, but it is ineffective to represent the distributed nature of open markets and the interactions conducted between the stakeholders. In Chapter 3, a different approach, modelling with Multi-agent Systems (MAS), is introduced. In Chapter 4, a Multi-agent System for Open Railway Access Market (MAS-ORAM) is proposed as a plausible means to model the interacting behaviour between the railway stakeholders.

Chapter 3

Multi-agent Systems

The last chapter has identified the difficulties of employing the concept of centralised computation in capturing the distributed nature of the railway stakeholders and their behaviour during negotiation. In this chapter, the concept of Multi-agent Systems (MAS), which fundamentally differs from centralised computation, is introduced. The background of MAS and the common misconceptions about agents are first given, followed by the implementation issues on developing multi-agent systems. The applications of agent modelling in e-commerce, electricity markets, road transportation and railways are then reviewed briefly.

3.1. Background

The origin of MAS took root in the research disciplines of Artificial Intelligence (AI) and Distributed Artificial Intelligence (DAI). The emergence of MAS from the two disciplines is discussed below.

3.1.1. Artificial Intelligence

The research on AI began in the 1950s. By definition, AI is a branch in science and engineering whose goal is to understand and mimic the intelligence of human beings in machines, especially in computer programs (McCarthy, 2004). The scope of intelligence was later broadened to include other living organisms such as insects.

Despite the clear definition of AI, the meaning of intelligence has been ambiguous because it is difficult to develop an objective distinction between being intelligent and otherwise.

Nevertheless, most researchers consider that intelligence is closely related to problem-solving abilities. In other words, one may regard an entity to possess certain extent of intelligence if it can solve problems. With such a perception on intelligence, the aim of AI is then to develop machines or programs that can tackle problems by employing observable mechanisms from living organisms. Since the beginning of AI, researchers have been studying these observable processes mainly from the philosophical and biological perspectives.

From the philosophical view, part of the human intelligence is contributed from the ability to reason with knowledge. Therefore, the studies on the representation of knowledge in machines, and the validation of inference rules in proposition logic, first-order logic and probability theory contributed to the major works in the early AI research. Other logical reasoning approaches such as temporal logic, fuzzy logic and case-based heuristics rules were also developed and examined. The realisation of these reasoning logics in software programs has gradually led to the development of knowledge-based systems such as expert systems (Russel & Norvig, 1994).

On the other hand, some biological processes have also been identified to exhibit problem-solving abilities. Two examples are Artificial Neural Network (ANN) and Genetic Algorithm (GA). In ANN (Minsky & Papert, 1988), the logical relationships between a set of inputs and outputs are modelled as a collection of interconnected artificial neurons. By properly deciding the number, type and topology of neurons, ANN provides a means to make inferences from a set of input stimuli similar to the biological neural system. In GA (Holland, 1992), the decision-making process is inspired by genetic evolution. A feasible solution to a problem is represented by the encodings in a chromosome. By carefully adapting the evolutionary operations of crossover, mutation and selection, a relatively good decision is anticipated to emerge from a set of less favourable ones.

3.1.2. Distributed Artificial Intelligence

While classical AI has focused mainly on mimicking the intelligence exhibited by a single individual, a subfield of AI emerged in the late-1970s whose focus has been the study of intelligence derived from a group of interacting organisms. This subfield, known as Distributed Artificial Intelligence (DAI), mainly attempts to solve problems which are inherently distributed, where knowledge and activities are separated naturally in space (Bond & Gasser, 1988). While each entity in the group has local problem solving capabilities, the entire problem is too complex for any single entity to handle. Only by proper coordination can these entities resolve the complex problem in a timely and an efficient manner (Durfee *et al.*, 1989).

Distributed Problem Solving (DPS) and Multi-agent Systems (MAS) are two branches of research under DAI (Bond & Gasser, 1988). DPS considers the decomposition of a problem and the allocation of these sub-problems to a set of loosely-coupled problem-solvers, which cooperate through the direct sharing of information (e.g. intermediate solutions). By being loosely-coupled, the problem-solvers are assumed to have little knowledge about the internal structure and status of each other (Bellifemine *et al.*, 2003a), but they may form temporal communication links to exchange the required information. Through such cooperative activities, these problem-solvers are therefore working towards a common goal. On the other hand, apart from allowing cooperation between the problem-solvers, MAS also permits them to possess individual goals that may be in conflict. Therefore, research in MAS tends to focus on devising mechanisms for conflict resolutions, which is essential in studying competitive interactions.

3.1.3. Multi-agent Systems

The study on MAS began in the 1980s (Wooldridge, 2002). The problem-solving entities in MAS are called agents. Similar to the difficulties in defining intelligence, a commonly accepted definition of agent is not available. However, it is generally recognised that an agent is at least an encapsulated computer system, situates in some environment, and can act flexibly and autonomously in that environment to meet its design objectives (Wooldridge, 2002; Jennings & Bussmann, 2003). In other words, an agent is a piece of software-driven hardware that can only work in some predetermined application domains (i.e. they are not ‘super-agents’), but provided that it is operating in these domains, it can handle its designated tasks rationally and adapt to changes in a flexible manner without human interventions. In MAS, the environment contains more than one agent, and they are required to exhibit social interactions such as negotiations or conducting auctions in the progress of pursuing their goals.

Agents are therefore characterised by being autonomous, rational (reactive and proactive) and able to conduct social interactions (Wooldridge, 1997). The advantages of conceptualising a system that composes of such entities have been summarised as decomposition, abstraction and organisation (Jennings & Bussmann, 2003). Firstly, the decomposition of a complex system into a group of agents allows software engineers to better manage and develop the program by handling the sub-problems in relative isolation. Secondly, through the abstraction of an agent community, some properties of the problem are emphasised while others are suppressed. For example, the programmer of agent A only needs to assume that the other agents will interact with A through some specified (social-like) interaction protocols, but not to concern with their detailed implementation. Finally, with respect to organisation, MAS allows a hierarchical structure of agents, where an agent can consist of several sub-agents. These advantages are particularly apparent when the complex system is naturally separated in space and groups, and the ultimate system behaviour emerges from the interactions among the local entities. The notion of a system composing of autonomous agents allows these complex systems to be studied from a ‘bottom-up’ approach (Teodorvić, 2003; Yaskawa & Sakata, 2003).

3.1.4. Common Misconceptions about Agents

Agent modelling is a comparatively new area for research which has attracted much attention from both academics and industries. The growing interest and applications in this topic have resulted in two extreme perceptions on MAS. On one hand, the extreme optimists may believe that agent modelling is the solution to all problems. On the other hand, the pessimists are criticising that agent-based software is identical to conventional programs, or programs derived from object-oriented programming. The following discussions clarify these common misconceptions about agent modelling.

Although the study on MAS is originated from AI, it has been pointed out that the involvement of AI and/or agent-specific techniques in developing a multi-agent system only contributes to a portion of effort in most applications (Wooldridge & Jennings, 1999). Since it is often the autonomy and social ability exhibited in a distributed problem that prompt the use of agent modelling, the rationality required from performing the atomic tasks of agents can still be greatly benefited from the adoption of classical (i.e. non-AI) techniques. This follows that agent modelling is not the solution to all problems because a decision problem that is not solvable by current techniques will not become solvable using agent modelling.

On the other hand, conventional software programs are different from agents, albeit agents are indeed software programs (Franklin & Graesser, 1996). An important property of agents is their ability to sense and act on the (physical or virtual) environment where they are situated. In general, the actions of an agent usually affect what it will sense later. For example, during a negotiation between agents A and B, when A proposes an offer O_{A1} to B, B may reject the offer and counter-propose O_{B1} to A, but if A proposes an alternative offer O_{A2} to B, then B may accept the offer. This is regarded as temporal continuity in agent modelling. By contrast, in conventional programs, their outputs will not usually have direct consequences on the environment. As a result, the requirement for being a software agent is more demanding than a simple program.

Similarly, objects are not necessarily agents (Wooldridge, 2002; Jennings & Bussmann, 2003). In the conventional definition, objects are computational entities that encapsulate states, perform actions and communicate by message passing (Wooldridge, 2002). While this bears much similarity to the definition of agents, software agents require a higher degree of autonomy and a more complex behaviour that are not considered in object-oriented programming. Although objects have a certain degree of autonomy through the encapsulation of internal states, objects have little autonomy regarding behaviour. For instance, object-oriented programming allows object A to directly invoke a function in object B because the programmers usually consider objects as cooperative entities to solve a common problem. However, in agent modelling, it is essential for an agent to possess the exclusive rights to decide and perform its actions because the requests from other agents may not be consistent with its best interest. With the consideration of such requirement, not only does agent-oriented programming need to prohibit the direct control of an agent over the others, but it also needs to provide the required freedom for an agent to respond (or not to respond) to the others. This is often achieved by the implementation of an agent interaction protocol that allows the act of bargaining. By employing such protocol, agents are also enabled to satisfy the requirement on temporal continuity.

3.2. Agent Development

From the above discussions, agent modelling provides a potentially useful means to solve distributed problems involving social-like interactions. To understand how multi-agent systems can be realised in practice, the important issues related to determining the size of an agent system, the current standards governing agent systems and the available middleware for agent development are further discussed.

3.2.1. Size of Agent Systems

At the beginning of developing a multi-agent system, software engineers have to decide a proper size of the agent community (i.e. number of agents) and the organisation of these agents. For example, a system could have only one layer of agents to represent the companies along a supply-chain, but it is equally feasible to expand two more layers for the departments within the companies and the staff working in the departments. In such cases, the agents in the department and staff layers will be sub-agents of the one at the company level (Fig. 3.1).

The one-layered approach offers a simple configuration and modelling of the interactions between agents. However, it may not be fully benefited from the decomposition provided by agent modelling. On the other hand, the three-layered

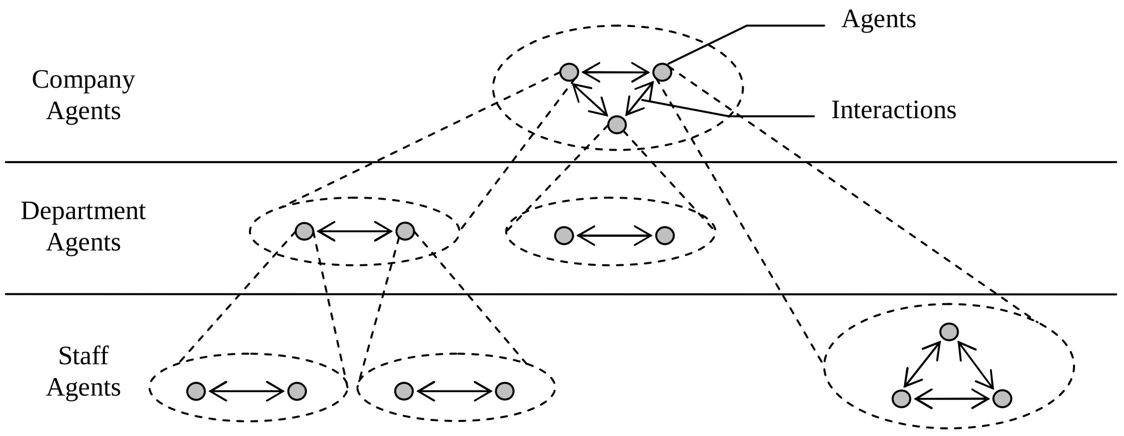


Fig. 3.1. Organisation of agents in layers

approach clearly requires a higher demand in modelling the interactions between agents, but more detailed studies can be performed at the department level and the personnel level. In fact, there is often a tension between the degree of decomposition and the level of organisation (Wooldridge & Jennings, 1999; Zambonelli & Luck, 2006). As noted in Section 3.1.3, decomposition helps the developers to distribute the sub-problems to different agents. The finer is the separation, the more manageable are the sub-problems. Nevertheless, when the number of agent increases (i.e. the organisation becomes complex), there will be a growing demand on message exchange between agents. This may result in a substantial amount of communication overheads (in the form of bargaining) which greatly reduces the efficiency on problem-solving using the multi-agent system. In other words, software engineers often need to balance the costs and benefits between decomposition and organisation. On one hand, this decision depends on the sophistication of the current computing technologies. On the other, it depends on the depth of study required. For example, if the study is focused on the interaction between companies, then the one-layered approach will usually be adequate.

3.2.2. Standards for Agent Systems

During the early years of research in MAS, most practitioners developed and studied their multi-agent systems independently. There were no common standards available for constructing agent systems. This resulted in impediments in applying agent modelling in solving practical problems. Being aware of the problem, a non-profit association called Foundation for Intelligent Physical Agents (FIPA) was established in 1996 (FIPA, 2005). Since then, FIPA has developed a set of specifications for agent systems to promote interoperability between agents. In 2005, FIPA also became an IEEE Computer Society standards organisation (FIPA, 2005).

One of the key specifications produced by FIPA is the FIPA reference model of an agent platform depicted in Fig. 3.2 (Bellifemine *et al.*, 1999; Poslad *et al.*, 2000). An agent platform is a virtual environment that allows software agents to operate on different operating systems and hardware architectures. As a consequence, agent developers need not be concerned about the diversity of the software and hardware configurations on which their agents are implemented. On this platform, FIPA specifies the responsibilities of three mandatory services. Firstly, the Agent Management System (AMS) provides a white page service (i.e. maintains valid agent

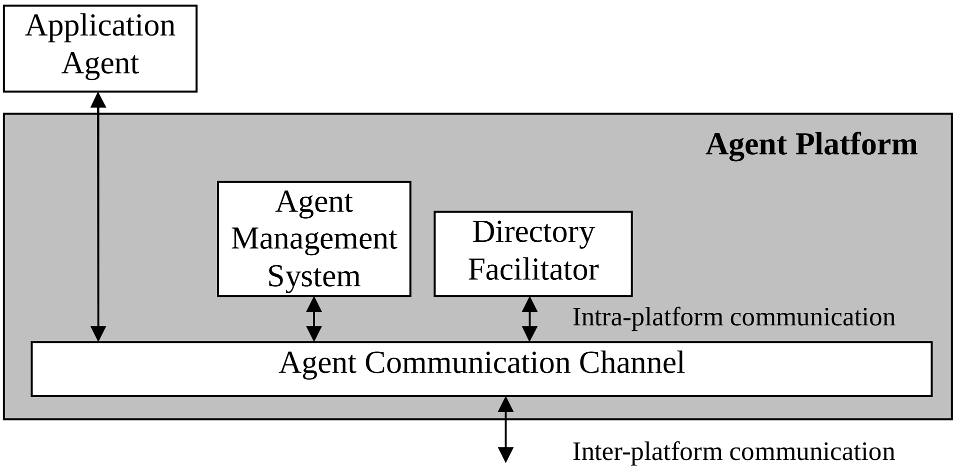


Fig. 3.2. FIPA reference model of an agent platform

IDs and states) to all agents on the platform. Secondly, the Directory Facilitator (DF) provides a yellow page service (i.e. maintains a record of agent services) on the agent platform. Finally, the Agent Communication Channel (ACC) controls all messages exchanged within the platform and across different platforms.

While the FIPA reference model has provided a standard for the composition of an agent system, the communication standards for agents are all derived from speech act theory (Covington, 1998). In speech act theory, the declaration of a sentence in human communication is considered as an action to influence the audiences. For example, when Bob says to Mary, ‘I need a pen’, Bob is not only delivering the statement to Mary, but he is also attempting to persuade Mary to give him a pen. Every speech act therefore consists of an illocutionary force F applied to a proposition P, which is formally known as the $F(P)$ hypothesis (Covington, 1998). In the example, F = ‘*request*’, while P = ‘Mary gives Bob a pen’.

To express the illocutionary force, Knowledge Query and Manipulation Language (KQML) (Finin *et al.*, 1994), FIPA Agent Communication Language (FIPA ACL) (Bellifemine *et al.*, 1999), and Formal Language for Business Communication (FLBC) (Moore, 1999) are three commonly used agent communication languages. These languages define a set of illocutionary forces called performatives such as ‘*request*’, ‘*inform*’, ‘*propose*’, ‘*reject*’, and ‘*accept*’. On the other hand, to represent the proposition, various types of knowledge representation techniques called content languages have been employed. These include Knowledge Interchange Format (KIF) (Wooldridge, 2002), Semantic Language (SL) (Caire, 2001), Constraint Choice Language (CCL) (Willmott *et al.*, 2000), and Extensible Markup Language (XML) (Bellifemine *et al.*, 2003a). These content languages differ from the nature of their knowledge domain. For example, SL is a general content language which supports expressions of proposition, action, and identifying reference expression (i.e. an embedded ACL message). On the other hand, CCL is specially designed to represent knowledge associated with a constraint satisfaction problem, which requires expressions of proposition, action, and object. Fig. 3.3 illustrates an example of their differences.

In addition to the need of an agent communication language and a content language, agent communication also requires a common ontology definition. In the example of borrowing a pen, Bob and Mary must have a common definition on ‘pen’, otherwise the communication will be meaningless. For example, if Bob is referring a pen as a

SL

Proposition

(inform
sender: Bob
receiver: Mary
content: **(is Tom handsome)**
language: SL)

Bob informs Mary (or wants her to believe) that Tom is handsome.

Action

(Request
sender: Bob
receiver: Mary
content: **(action (Mary) (give Pen (Bob))**
language: SL)

Bob requests Mary to give Bob a pen.

Identifying reference expression

(Request
sender: Bob
receiver: Mary
content:
(action (Mary)
(*inform-ref*
sender: Mary
receiver: Bob
content: (iota ?x LocatePen ?x)
language: SL))
language: SL)

Bob requests Mary to inform him the location of his pen.

CCL

Proposition

(inform
sender: Consultant
receiver: Client
content: **(CSP-insoluble (CSP-ref “P-1”)**
language: CCL)

Consultant informs Client that the CSP named “P-1” is not solvable.

Action

(Request
sender: Client
receiver: Consultant
content:
(action: CSP-give values
(object: CSP
CSP-ref: “P-1”
Var{ A, B, C}
language: CCL)

Client requests Consultant to solve the CSP named “P-1” consisting of 3 variables.

Object

(Inform
sender: Consultant
receiver: Client
content:
(object: CSP-solution
CSP-ref: “P-1”
Ans{ (A 30, B 45, C60) }
language: CCL)

Consultant informs Client the CSP-solution of “P-1”, with A=30, B=45.

Fig. 3.3. Knowledge representation in SL and CCL

stationery, but Mary is considering it as a ring of fence, then the communication will not lead to the desired action (i.e. Mary gives Bob a writing pen). In agent communication, there are no standards governing ontology definitions and programmers are required to create their own ontology by defining a set of vocabularies (i.e. terminologies used in agent communications, e.g. pen, give, etc.) and their relationships (i.e. structure and semantic, e.g. pen has the attributes of colour, size and price).

3.2.3. Middleware and Agent Shell

To develop the agent platform and implement the agent communication languages and content languages, a substantial amount of knowledge and efforts in computer science are required. To reduce the development time of these software infrastructures, a number of agent software packages have been introduced. In particular, JADE (Java Agent DEvelopment Framework) (Bellifemine *et al.*, 1999), and FIPA-OS (FIPA Open Source) (Poslad *et al.*, 2000) are two commonly employed agent development toolkits which are designed to comply with the FIPA standards.

These software packages are often referred as middleware. In general, middleware is a class of software that contains higher-level libraries which capture the generic functions or services that are used by the most applications (Bellifemine *et al.*, 2003a). This encourages the reusability of codes and simplifies the software development so that agent developers may concentrate on design and implementation of the application logic. In agent-oriented programming, apart from the creation of the agent platform and the implementation of agent communication languages and content languages, the generic functions also include procedures to handle asynchronous agent communication, timeout detection of agent activities and agent lifecycle control (Emorphia, 2001; Bellifemine *et al.*, 2003a).

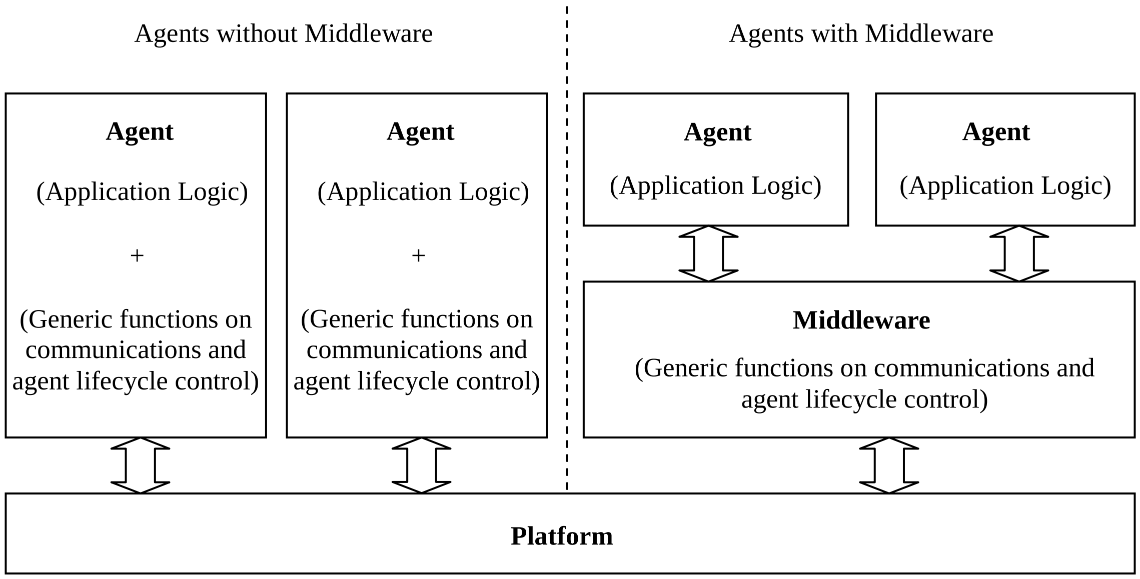


Fig. 3.4. Role of middleware in agent development

The middleware therefore provides a programming environment for agent developers to design their applications within an agent shell (Fig. 3.4). With the aid of the middleware, agent developers are only required to devise the local agent models and their interaction protocols. For example, the rational decision-making process of an agent may employ classical techniques such as nonlinear programming (Bazaraa *et al.*, 1993) or agent-specific techniques such as the BDI (Belief-Desire-Intention) model (Rao & Georgeff, 1995). For the interaction protocol, a popular choice may be the contract net protocol (CNP) (Smith, 1980).

3.3. Applications

Agent modelling has found a wide range of applications in many distributed problems. Described below is a selection of agent applications in e-commerce, electricity markets, road transportation and railways.

3.3.1. E-commerce

The increasing utilisation of the Internet through both fixed and mobile computer terminals has encouraged the rapid development of e-commerce (Shaw *et al.*, 1997; Liu & You, 2003). Apart from the conventional approaches of purchasing merchandises via personal visits and mail ordering, e-commerce has offered an alternative means of shopping behind computers. Consumers may now obtain their desired products conveniently through online ordering, negotiation and e-auctions.

Considering the vast number of suppliers available on the Internet and the substantial time required to monitor the bid prices of products in e-auctions, a number of research has been conducted to devise more efficient technologies for e-shoppers. The ideas of having some intelligent decision support systems which help shoppers to shortlist the desired items and automatically perform biddings on their behalf have prompted the use of agent modelling in e-commerce. There are several reasons of applying agent modelling in these applications (Jennings *et al.*, 2000a, b). Firstly, the buyers and sellers are physically distributed and they may form a network of temporal connections via the Internet. In other words, the e-shopping problem is naturally a distributed problem, composing of a network of loosely-coupled entities. Secondly, these entities require a high degree of autonomy in decision-making and need to conduct social-like behaviour in the case of negotiation and e-auctions. In addition, with such a diverse spectrum of evolving choices and requirements, the entities also need to handle the dynamism and uncertainty in a sensible manner.

One major contribution generated from agent research in e-commerce is the variety of mechanisms in making concessions during agent negotiation and auctions. At the beginning, concession (i.e. the reduction in cost of an attribute such as price) made between successive rounds of negotiation is based on primitive functions such as linear, quadratic or exponential ones (Faratin *et al.*, 1998, 2002). Later on, more intelligent methods, such as using GA to determine the choice of concession function according to the proponent's behaviour (Krovi, 1999) and fuzzy-logic based negotiation functions (Luo *et al.*, 2003), have also been introduced. In addition, by employing iterative functions, agents with reactive property are also devised (Sim & Wong, 2001; Sim & Choi, 2003; Sim & Wang, 2004).

3.3.2. Electricity Markets

Electricity markets have been deregulated in many countries similar to the open railway markets. In order to introduce competition in open electricity markets, the provision of services in electricity generation, transmission and distribution are allocated to different stakeholders. There is usually only one transmission company (TC) which is responsible to provide the transmission grid to a set of generation companies (GCs) to inject electrical power into the grid, and a set of distribution companies (DCs) to receive power from the grid (Fig. 3.5).

In an electricity market, GCs sell electricity to DCs via an auctioning approach. According to the marginal cost curves for power generation, the GCs will derive a set of bid prices and the quantities of electricity generation at those prices. Having taken consideration of the forecast demand and prices, the DCs then select a subset of these bids to purchase the required electricity for distribution (Brazier *et al.*, 2002; Al-Agtash

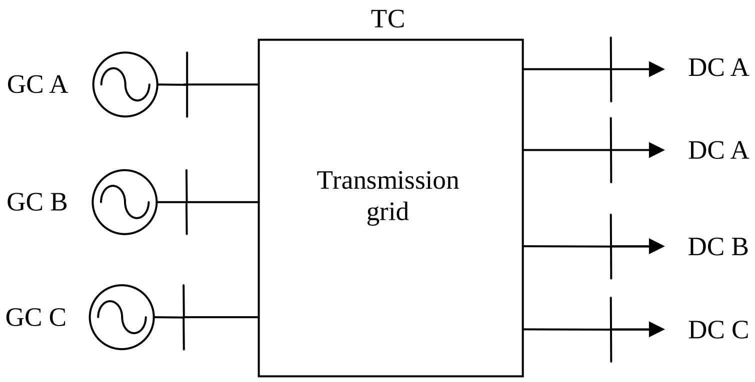


Fig. 3.5. An open electricity market

& Al-Fahoum, 2005; Bagnall & Smith, 2005).

When setting the bid prices, the GCs may consider forming coalition with one another (Krishna & Ramesh, 1998a, b). When a coalition is formed, two GCs modify their bid prices in return of a possible gain through the sharing of revenue obtained from their power generation. In such cases, the two GCs will need to conduct a negotiation in order to settle at a mutually agreeable proportion of revenue share.

Apart from paying a premium to the GCs, the DCs also need to settle a transmission service charge to the TC for the use of the grid (Ilic *et al.*, 2003). The charge is levied by the TC considering the constraints on maximum power flow in the network. Nevertheless, the DCs have the rights to reject the offer if they are not satisfied. Under such circumstance, the DCs will need to participate in the auction with the GCs again before a new transmission service charge may be derived.

Owing to the distribution of stakeholders and the social interactions exhibited between them, agent modelling has been employed to study on the decision-making requirements of the stakeholders in open electricity markets (Krishna & Ramesh, 1998a, b; Brazier *et al.*, 2002; Ilic *et al.*, 2003; Al-Agtash & Al-Fahoum, 2005; Bagnall & Smith, 2005). While the negotiation and auction behaviour have been benefited from the works in e-commerce, decision-making problems in electricity markets are more complex owing to the constraints imposed by the physical infrastructure (e.g. the maximum rating of generators and the maximum power flow of the grid). As a consequence, specially designed algorithms have been devised to enable the agents to perform rationally in these applications.

3.3.3. Road Transportation

Apart from the applications in e-commerce and electricity markets, agent modelling has also found many applications in road transportation (Dia, 2002; Rossetti *et al.* 2002; Hallé & Chaib-draa, 2005; Ossowski *et al.*, 2005). With the continuing increase in the number of automobiles, road congestion is now a major problem in many countries. Transport authorities have introduced means, such as ramp metering (Zhang & Levinson, 2005), to relieve traffic congestion by diverting traffic to less congested areas. As a result, drivers often need to adjust their route choices as they acquire new information from the current traffic conditions and road signs.

It is under this context that agent modelling has been employed in many applications in road transportation to represent drivers as software agents (Dia, 2002; Rossetti *et al.* 2002). In these studies, the development of driver agents enables the evaluation of new traffic control strategies and proper urban planning. These agents are also able to mimic the proactive behaviour of drivers by employing BDI-model (Rao & Georgeff, 1995).

3.3.4. Railways

Despite the substantial efforts in MAS applications in e-commerce, electricity markets and road transportation, there are relatively few applications in railways. To the author’s best knowledge, there is an absence of research conducted on open railway access markets using agent modelling. The following is a brief description on two limited agent applications in railways.

A multi-agent system was developed for dispatching freight trains on a single railway line (Cuppari *et al.*, 1999). Traffic congestion on the line was traditionally handled by human traffic coordinators in different sections along the railway line. They cooperatively determine an efficient real-time schedule for the freight trains. In the multi-agent system, software agents are developed to act on behalf of these coordinators so that the dynamism (e.g. the changing congestion level) and uncertainties (e.g. stochastic delays) can be handled at high speed. The system includes six agents. Three agents are responsible for regulating the congestion level along three sections of the line, and one agent oversees the overall traffic condition. The remaining two agents are responsible for efficient management of wagon-loading at the terminal and train-dispatching at the yard respectively.

Another application was on the train coupling and sharing problem in freight rail transportation (Böcker *et al.*, 2001). Since track allocation cost is calculated with respect to the number but not the length of trains, in order to reduce the cost, different wagons may be coupled together at terminus to form a longer unit so that the wagon loads may be transported on the track at the same time. The wagons may then be uncoupled when the unit reaches the marshalling yard at the destination. In the study, the wagons of the same destination are grouped together as an agent. To explore the possible reduction in track allocation cost, agents need to interact with each other to determine whether the other agents may share the use of track, given the deadline for transportation. Through the resolution of agent interaction, a global train timetable is derived from the local activities of the agents.

3.4. Remarks

From the discussions in this chapter, agent modelling is clearly more suitable than the concept of centralised computation to resolve distributed problems involving behavioural interactions. Agent modelling therefore offers a promising means to handle the resource management problems in open railway access markets. With the availability of the agent standards and development toolkits, the technological impediment to construct a multi-agent system has been reduced. In addition, the research experiences in e-commerce, electricity markets and road transportation have provided the necessary footprints for modelling agent behaviours and tackling the domain specific tasks encountered. In the next chapter, a Multi-agent System for Open Railway Access Market (MAS-ORAM) is proposed.

Chapter 4

Multi-agent System for Open Railway Access Market

In Chapter 2, the background on open railway access markets was examined. An open market basically consists of a number of stakeholders including an infrastructure provider (IP) and a group of train service providers (TSPs). In some cases, ancillary service providers such as rolling stock providers (RSPs) and maintenance service providers (MSPs) are also present. With the separation and distribution of managerial responsibilities, on-rail and off-rail competitions have been introduced in modern rail markets, which have greatly changed the resource management processes in railways. In particular, the stakeholders have to negotiate with each other to resolve the possible conflicts in cost recovery or capacity allocation.

Post-evaluation studies on the efficiency in railway resource management are usually time-consuming and costly. Instead, a computer simulation approach is conceived to be a cost-effective means to study the conflict resolutions between the stakeholders. Unfortunately, open markets are distributed systems characterised by the composition of self-interested parties and the involvement of social-like interactions, which have posed major difficulties in employing the conceptual view of regarding a system as a central decision-making unit.

From the review on the research conducted on Multi-agent Systems (MAS) in

Chapter 3, agent modelling differs from the classical concept in that the system consists of a group of self-interested entities capable of performing cooperative and competitive activities. Comparing to the characteristics of open railway access markets, agent modelling seems to offer a potentially feasible means to study the resource management problems of the stakeholders. With the availability of the FIPA standards and the FIPA-compliant middleware for agent development, the involvement in time and technical knowledge on devising agent-based software have been reduced. Therefore, this chapter proposes a Multi-agent System for Open Railway Access Market (MAS-ORAM) using a popular middleware package called JADE (Java Agent DEvelopment Framework).

4.1. Agent Development in JADE

JADE is a middleware for developing FIPA-compliant agents (Bellifemine *et al.*, 2003a). The source code of JADE is distributed under the GNU Lesser General Public License (LGPL). Under the permission of this licence, JADE users have no restrictions on developing their agent-based systems provided that the included software libraries are not modified. By entrusting the freedom of using the software, the licence intends to promote the widespread development and application of agent-based technologies.

As a middleware, JADE provides the essential software components for agent development. Firstly, it enables the creation of an agent platform on which user-developed agents may attach to cooperate or compete. It also offers a generic agent class structure (i.e. agent shell) through which agent programmers can focus on devising the application logic. Moreover, JADE has simplified the creation of agent messages by offering the use of FIPA ACL and SL, and allowing users to define their own ontology. In addition, it provides a rich set of agent behaviours so that programmers can devise their tailor-made agents to perform a variety of different functions.

4.1.1. Agent Platform

The JADE agent platform is composed of a collection of Java Virtual Machines (JVMs). The platform can be operated on a single host machine or distributed among several ones as shown in Fig. 4.1 (Bellifemine *et al.*, 2003b). Each host executes only one JVM, and the JVMs on separated hosts are interconnected together through a communication network (e.g. Bluetooth, GPRS, Wireless-LAN and the Internet). Each JVM represents a basic container in which multiple agents are allowed to operate concurrently on a host. In JADE, the tasks of AMS (white-page) and DF (yellow-page) services required by the FIPA standards are provided by specific agents. The container where the AMS and DF agents are situated is referred as the main container. The other

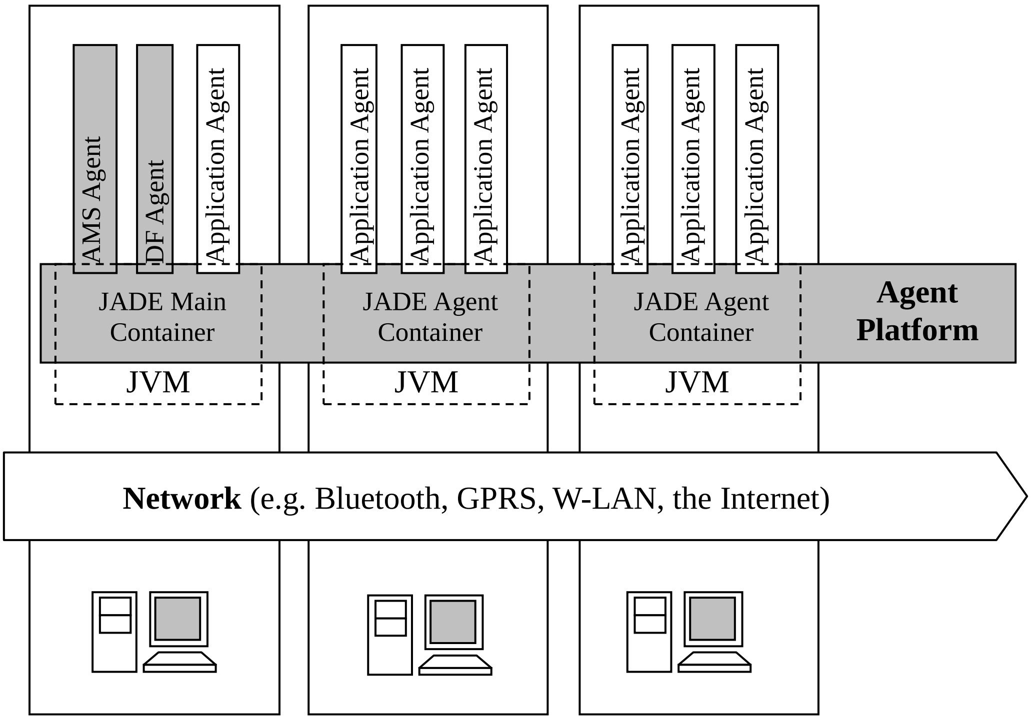


Fig. 4.1. JADE agent platform distributed over several hosts

containers are required to share the use of AMS and DF services with the main container so that there are only one AMS agent and one DF agent present on an agent platform.

4.1.2. Agent Class Structure

JADE provides two classes, *GuiAgent* and *Agent*, so as to enable the programmers to develop their agents with and without the support of a Graphical User Interface (GUI) respectively. These classes effectively provide an agent shell through which the application agents will possess (inherit) the basic properties of a JADE agent. These properties include the basic interaction with the agent platform (i.e. accessing AMS and DF services) and a set of functional calls that may be used to customise the agents (e.g. sending and receiving messages, using standard interaction protocols such as the contract-net protocol). The availability of these properties hides the complex implementation of the platform and functions from the programmers, making the development of agents easier and less time-consuming.

4.1.3. Agent Messages

As explained in Chapter 3, in order to enable social-like interactions among agents, communication between agents is based on speech act theory, or the $F(P)$ hypothesis. JADE employs FIPA ACL as the standard agent communication language to represent the illocutionary force F. A complete list of FIPA ACL performatives and their definitions is shown in Table 4.1 (Wooldridge, 2002). The proposition P is then expressed using a content language selected by the programmers. By default, JADE employs SL which is a versatile language suitable for many generic applications. Other languages such as KIF, CCL and XML may also be imported from external libraries.

Table 4.1. Definitions FIPA ACL Performatives

Performatives	Definitions
accept-proposal	Accept a proposal made by the recipient
agree	Agree to perform an action previously requested by the recipient
cancel	Ask the recipient to stop executing an action previously requested by the sender
cfp	Call-For-Proposals: Initiate negotiation between agents
confirm	Inform the recipient that the sender believes a given proposition is true
disconfirm	Inform the recipient that the sender believes a given proposition is false
failure	Notify that an action requested by the recipient was attempted but failed
inform	Persuade the recipient to believe the content message
inform-if	Persuade the recipient to believe a given proposition is true
inform-ref	Persuade the recipient to believe for a referential expression
not-understood	Inform the recipient that sender did not understand why an action have been performed
propagate	Ask the recipient to forward the embedded message to a set of agents satisfying some defined criteria
propose	Submit a proposal for consideration by the recipient
proxy	Inform the recipient that the sender is treating the agent as a proxy for a set of agents
query-if	Ask the recipient whether a given proposition is true
query-ref	Ask the recipient for a referential expression
refuse	Refuse to perform an action previously requested by the recipient
reject-proposal	Reject a proposal made by the recipient
request	Request the recipient to perform some action
request-when	Request the recipient to perform some action once when some proposition becomes true
request-whenever	Request the recipient to perform some action whenever some proposition becomes true
subscribe	Ask the recipient to notify the sender whenever the value of something changes

JADE also simplifies ontology creation by providing a base class called *Ontology*, which contains the essential functions to interact with the *ContentManager*, another important class for manipulating agent messages when programming in an agent shell. To completely set up a user-specific ontology, the programmers should first define the structure of the domain knowledge as Java objects which are extended from one of the classes: *Concept*, *AgentAction* and *Predicate*. *Concept* is usually used to create knowledge representation of tangible or intangible objects (e.g. pens, railway stations, time, etc.). *AgentAction*, as its name implies, is used to represent agent actions (e.g. give, find solution, etc.). *Predicate* is used to devise expressions that can be either true or false (e.g. whether a problem is solvable). Having performed the first step, the schemas (a concise summary) of the relationship between the knowledge structures should be defined in a class extended from *Ontology*. In the final step, the set of vocabulary (i.e. member variables) that has been used in the definitions of the knowledge structures should be listed in a Java interface.

4.1.4. Agent Behaviours

To allow the software agents to perform the necessary tasks within the agent shells, JADE provides a rich set of classes for constructing these activities called behaviours. A class hierarchy of these behaviours is shown in Fig. 4.2. These behaviours are different from normal procedural calls in classical programming. When a set of agent tasks derived from these basic behaviours are added to an agent, they will not be necessarily performed in a sequential manner. Instead, an internal scheduler is available to arrange the order and interleave between the tasks automatically so that these tasks appear as if they were performing concurrently.

Among the set of JADE behaviours, an agent task may be regarded as *SimpleBehaviour* or *CompositeBehaviour*. *SimpleBehaviour* is a generic behaviour where programmers have the highest freedom to define the task content, in addition to the starting and stopping criteria of the behaviour. Four additional behaviours, *OneShotBehaviour*, *WakerBehaviour*, *CyclicBehaviour* and *TickerBehaviour* are extended from *SimpleBehaviour* by JADE. *OneShotBehaviour* will be performed once, and its contained task will begin as soon as the behaviour is loaded on the scheduler. *WakerBehaviour* is identical to *OneShotBehaviour* except that the behaviour will not

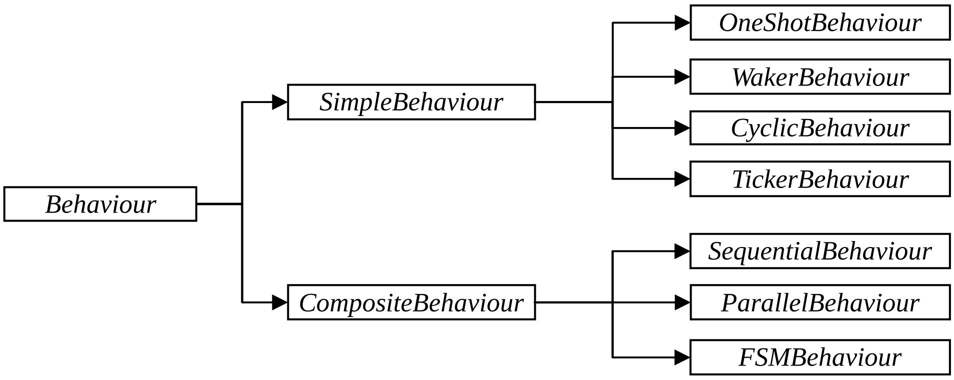


Fig. 4.2. Class hierarchy of JADE agent behaviours

start until a given time is elapsed. In *CyclicBehaviour* and *TickerBehaviour*, the task will be carried out repeatedly, but the latter has a given time lag between successive actions.

CompositeBehaviour is composed of a collection of sub-behaviours. In *SequentialBehaviour*, the sub-behaviours will be executed one after another. On the other hand, the sub-behaviours in *ParallelBehaviour* will be scheduled to interleave among themselves and there is an option to terminate the entire set of behaviours when either one or all the sub-behaviours have finished. *FSMBehaviour* (FSM stands for Finite State Machine) resembles the operation of a finite state machine which requires programmers to define the initial, intermediate and terminal states (or tasks), in addition to their possible transitions between the states.

Although simple and composite behaviours are primitive and they do not directly mimic the decision-making behaviours exhibited by human beings, they can be used as elementary building blocks to derive more complex behaviours. For instance, Distributed Systems Group at University of Hamburg has extended these behaviours in a project called JADE-X (Pokahr & Braubach, 2005) so that JADE agents may incorporate the BDI-behaviour proposed by Rao and Georgeff (1995). In the BDI agent modelling paradigm, agents are given a set of capabilities known as plans at design stage, but the decision on when these plans are executed is determined by the real time commitment to one or more goals, which in turn is affected by the runtime conditions that are acquired as beliefs. As a result, BDI-agents can exhibit human-like decision-making behaviours where the commitment to a goal can be adjusted dynamically. This is particular useful in developing agents that are required to change objective at runtime (e.g. proactive agents).

4.2. MAS-ORAM Architecture

The concepts of distributed and self-interested entities, coordinated behaviours and local intelligence in agent modelling have motivated the use of MAS in studying the resource management problems in open railway access markets. A high-level conceptual framework of a MAS-ORAM employed in this study is illustrated in Fig. 4.3.

This framework contains only one level of agents to represent the stakeholders in an open access market because the current study is intended to focus on the conflict resolutions between the stakeholders, rather than those between departments and staff members. Each stakeholder is represented by one agent. In addition to the IP and TSP agents, the framework may also be expanded to include the ancillary service provider agents of RSP and MSP. Prior to operations, the stakeholders assign their confidential information such as cost curves and operational tactics to their corresponding software agents before they are connected to a common agent platform. When an agent joins the platform, it is registered to the DF agent whose function is to maintain an updated record of the agent addresses and the services they provide. A stakeholder agent may therefore recognise the existence of other agents by performing a query to the DF agent.

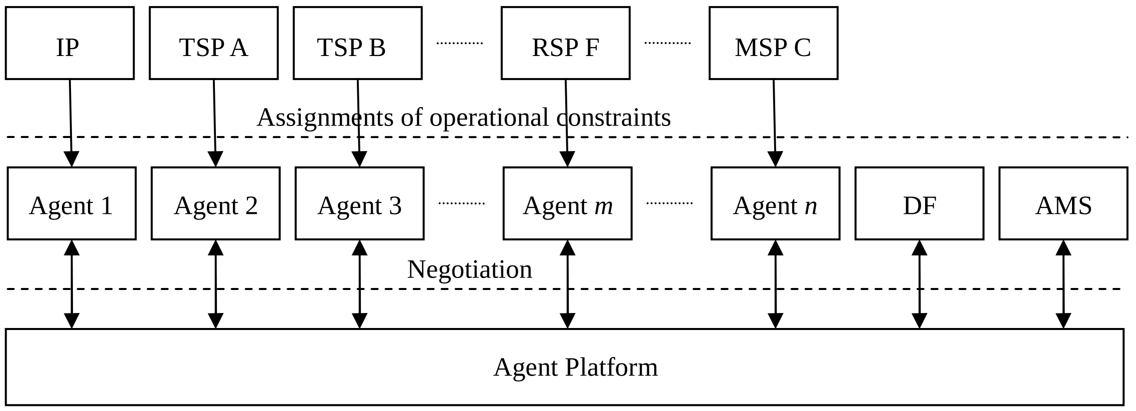


Fig. 4.3. A high-level conceptual framework of a MAS-ORAM

An agent on the platform will be perceived as either a resource provider or a purchaser without disclosing its internal status to the other agents. The agents on the platform are not expected to share a common goal, but they may form temporal association to examine whether a sale of resource is feasible and beneficial according to the pre-assigned criteria of the stakeholders.

It is also valuable to note that the proposed framework can also be applied in third-party access markets. In this case, one of the TSP agents will be possessed by the same stakeholder that owns the IP agent. However, as most regulations will prevent unfair gain of track access by the incumbent owner, the above and below railway activities will still be separately managed by different departments within the company. As a result, the two departments can still be modelled as separate agents.

4.2.1. Single Platform

With the incorporation of the JADE agent platform, Fig. 4.4 shows a possible realisation of agent architecture for the proposed MAS-ORAM framework using a single platform. Despite the adoption of the one-layer representation of a stakeholder agent in the study, the realisation illustrated in the figure (also in Figs. 4.5 and 4.6) displays a multi-layer representation with the inclusion of auxiliary agents. The reason for such generalisation is in twofold. Firstly, the realisation is prepared for future expansion of the MAS-ORAM to allow the examination of interactions in the department and staff levels. Secondly, in addition to the purpose of conducting simulation studies on open access markets, the realisation is devised with the considerations of using the multi-agent system to perform the real transactions, even if such implementation may not be feasible in the near future. Nevertheless, the elimination of the auxiliary agents from the figure will directly reduce the MAS-ORAM

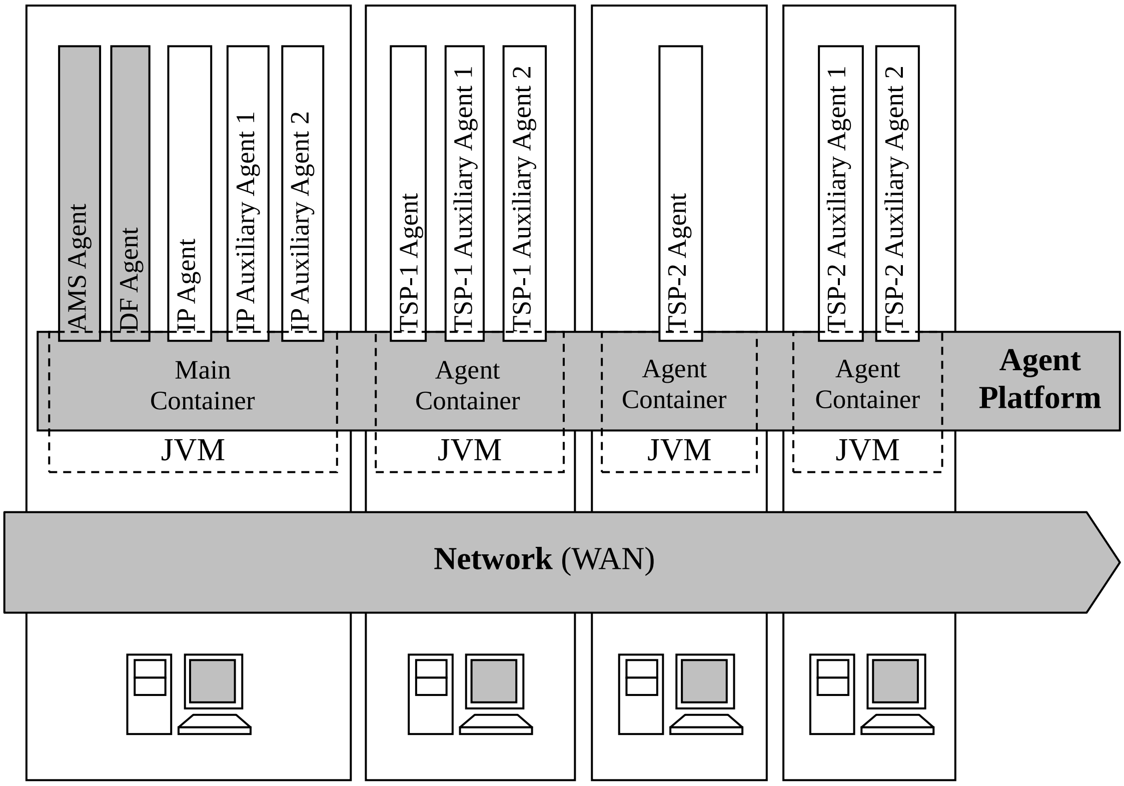


Fig. 4.4. A realisation of MAS-ORAM: IP agent on main container

to the one-layer representation.

In the single platform architecture shown in Fig. 4.4, the stakeholders are located on separate hosts and the IP agent is situated on the main container. These hosts are interconnected via a Wide Area Network (WAN) so that the agents are supported with a network capable of handling large volume of data transfer.

There are two benefits of this arrangement. Firstly, since stakeholders are managed by different authorities, it is better to physically distribute their agents on different hosts to allow local management and configurations of hardware, software and data. For example, the IP agent may require a high computation demand so that the host machine belonging to the IP is allocated with equipment of higher processing speed and memory requirements. On the other hand, the TSP agents may not need such powerful configurations and they can opt for less expensive settings. Using this architecture, the stakeholders can tailor-design their host machines according to their local requirements.

The second benefit is expandability. As the railway agents are developed and enhanced, the tasks performed by an agent may become increasingly complex. Eventually, the need of auxiliary agents may arise from the demand for more efficient management of the subtasks. More importantly, higher computational efficiency may be achieved by parallelism and concurrency. For instance, the main agent and the auxiliary agents of TSP-2 may be allocated in separate hosts so that the speed of computation is improved by simultaneous operations of two processors.

Despite these advantages, the architecture suffers from a lack of graceful degradation. As shown in Fig 4.4, the IP agent is situated on the main container, where AMS and DF services are also located. In other words, the host machine of the IP has the responsibility to provide AMS and DF services. If the host machine of IP fails, then the entire system will come to a halt, even for transactions between TSP agents that do not involve the IP agent.

An improved architecture is thus shown in Fig. 4.5, where the AMS and DF agents are isolated from all stakeholder agents. Even when the IP host machine fails, activities between the TSP agents can still be performed. If the main container failed, a redundant (or backup) system can be set up to replace the fault system. Such backup system does not require intensive computation power and set up cost. The host of this main container may be owned and operated by a neutral-party (e.g. the transport department).

4.2.2. Multiple Platforms

In the case where the open market spans across two or more jurisdiction areas, the multiple platforms architecture shown in Fig. 4.6 can be employed. Each platform is

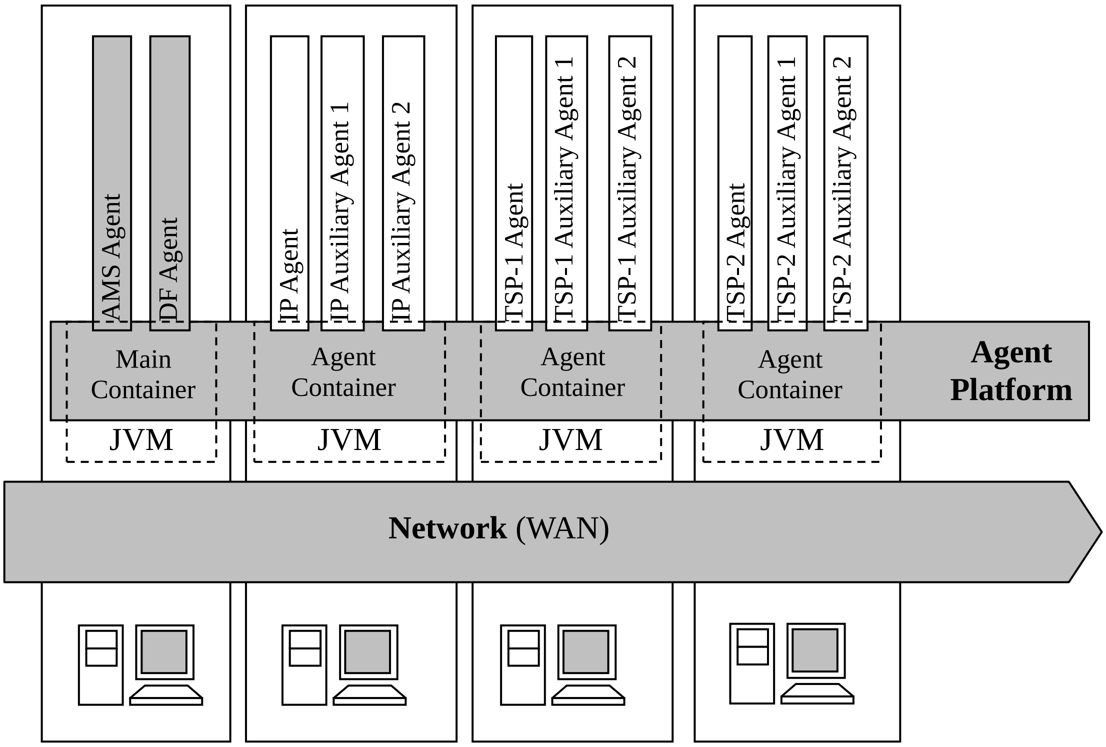


Fig. 4.5. A realisation of MAS-ORAM: Isolated main container

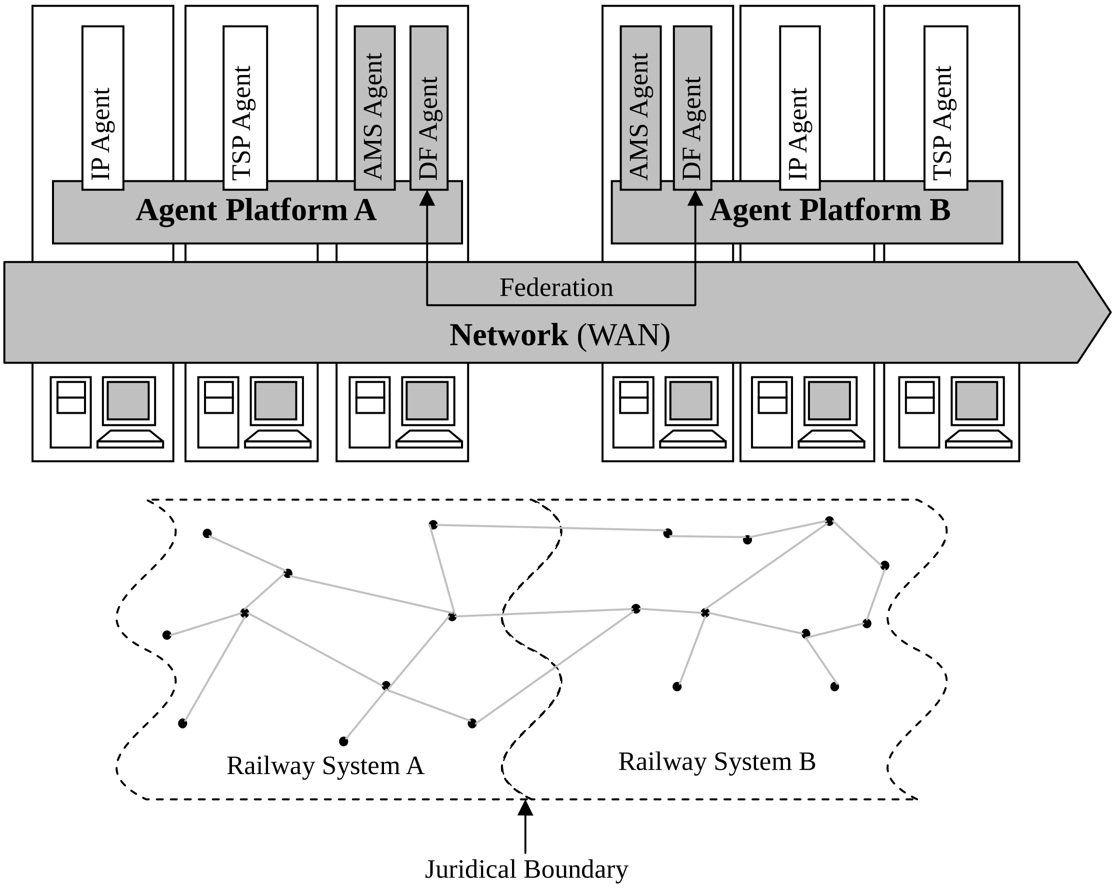


Fig. 4.6. Multiple platforms for railway open market across multi-juridical areas

mainly responsible for the transactions occurred within a single jurisdiction area. Platforms are interconnected by the federation of DF agents so that services provided by jurisdiction area A are available to jurisdiction B and vice versa. Transactions related to inter-juridical travel can thus be performed. Fault-tolerance associated with the breakdown of individual platform is also enhanced.

4.2.3. Single Container Platform

Fig. 4.7 shows the single container platform architecture used for testing and development in this study. As MAS-ORAM is still at the early stage of development, the complexity of agents is less demanding. In addition, since the needs for graceful degradation and flexible local configuration are only issues for the practical implementation, the single container architecture is sufficient for testing the behaviour of the agents in this study.

4.3. Problem Specifications

With the proposed MAS-ORAM framework and realisations in architectures, it is anticipated that a variety of simulation studies can be conducted on the resource

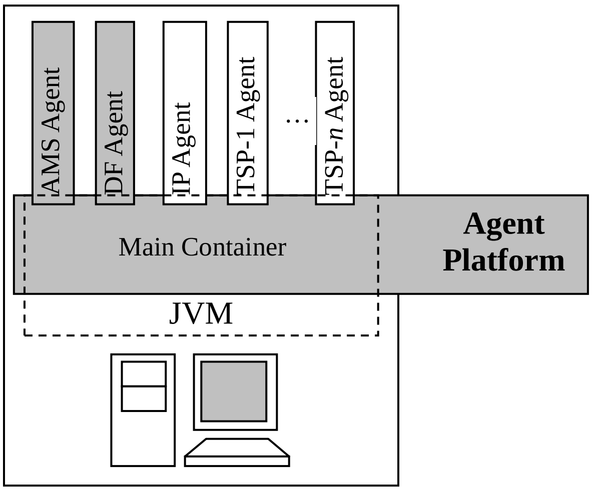


Fig. 4.7. Single container platform for testing and development

management problems in open access markets. As examined in Chapter 2, the major resource management problems include the track access allocation problem between an IP and a TSP (IP-TSP), the track access allocation problem between an IP and a group of TSPs (IP-TSP^n), and the schedule coordination problem at an interchange station between two TSPs (TSP-TSP) (Fig 4.8). These problems will be covered in detail in Chapters 5 to 7, but a summary of the methodologies employed is given below.

4.3.1. Track Access Allocation in Single Transaction

In an IP-TSP negotiation, the decision-making problem of the TSP agent is formulated as a Prioritised Fuzzy Constraint Satisfaction (PFCS) problem (Luo *et al.*, 2003). In this problem, the satisfactions of the TSP on a range of schedule times are represented by a set of fuzzy membership functions. Through a set of relaxation criteria on the fuzzy constraints, the TSP agent is enabled to generate a sequence of crisp constraints that will optimise the overall satisfaction on the track access rights. At the same time, the IP agent uses a Branch-and-Bound (BNB) algorithm (Papadimitriou & Steiglitz, 1998) to derive the optimal resource plan based on its charging functions and capacity constraints. Negotiation between the agents is enabled by the Buyer and Seller Behaviour Protocol (BSBP) (Luo *et al.*, 2003) which assists the negotiation to arrive at the Pareto-optimal deal. In a negotiation involving multiple

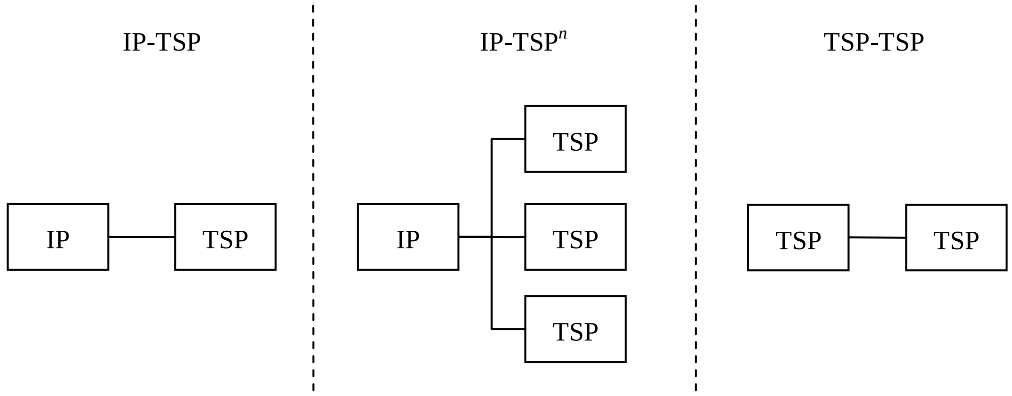


Fig. 4.8. Negotiating parties in an open railway access market

entities, a solution is Pareto-optimal if any deviations from this solution results in worse payoffs for at least one entity (Ehtamo *et al.*, 1996). In other words, the Pareto-optimal can be regarded as a ‘win-win’ solution.

4.3.2. Track Access Allocation in Multiple Transactions

To allow multilateral negotiation in an IP-TSP^n transaction, BSBP is extended to Multiple Buyers and Seller Behaviour Protocol (MBSBP). The IP agent handles the multiple negotiations in a sequential manner by employing either the First-Come-First-Serve (FCFS) or the Highest-Willingness-to-Pay-First (HW2PF) sequencing policies. The performance of these sequencing policies under different traffic conditions is studied by statistical analysis, such as *t*-test statistics and hypothesis testing (Walpole *et al.*, 1998).

4.3.3. Schedule Coordination at Interchange Stations

The TSP-TSP negotiation uses a simple protocol allowing the agents to propose, accept or reject offers. There are three negotiation strategies (S_{po}, S_{min} and S_{max}) proposed for the TSP agents to generate potential offers. These strategies involve the incorporation of Lemke’s Complementary Pivoting Algorithm (LCPA) in their respective local searching techniques. The performance of these strategies with respect to the quality of solution and negotiation time is evaluated by a comprehensive set of simulation case studies.

4.4. Remarks

This chapter has introduced a popular agent development middleware called JADE, which provides the essential tools for agent development, including the agent platform, the agent shell and a basic set of functions for creating agent messages and agent behaviours. In addition, a MAS-ORAM framework for modelling open access markets has been proposed. The framework is realised by incorporating the JADE agent platform, from which several MAS-architectures are proposed and their advantages on expandability, flexible local configuration and graceful degradation are discussed. The single platform architecture may be employed by an open access market within a juridical region, while the multiple platforms structure is used when stakeholders are separated by inter-juridical boundaries. However, the single-container platform is specially designed for the testing and development of MAS-ORAM in this study.

The three major resource management problems in open access markets (IP-TSP, IP-TSP^n, and TSP-TSP) are also briefly reviewed and their details and simulation results will be presented in Chapters 5 to 7. However, the application of the proposed MAS-ORAM is not restricted to these problems. For example, the framework also allows the study of the effects from different degrees of competition by altering the number of resource providers and/or purchasers. This may aid railway regulators to determine the suitable degree of competition in railways. Moreover, different transaction policies (e.g. posted pricing, negotiation and auctioning) can be formulated and tested to improve the charging regime. Further studies may also be performed to evaluate the impacts from any proposed changes in regulations, business objectives and engineering operations by modifying the rational behaviour of the agents. For instance, constraints as a result of regulatory changes can be added locally to the relevant agents, and modification on business objectives and scheduling mechanism may be achieved by adjusting the internal cost functions and implementing a proper mathematical model respectively. Results from these simulations are expected to improve capacity utilisation and competitiveness of the stakeholders.

Chapter 5

Bilateral Negotiation for Track Access Rights

Owing to the separation of train operations from infrastructure provision, one of the core activities in open access markets is the allocation of track access rights between an IP and a group of TSPs. As reviewed in Chapter 2, there are basically three access charging regimes, namely posted pricing, negotiation and auctioning. Since posted pricing usually fails to differentiate track access rights of different operating requirements, negotiation and auctioning are preferred so that costs may be recovered more effectively according to the actual utilisation of infrastructure. As auctioning is not yet a popular approach in the current open access markets, the allocation of track capacity through negotiation is modelled and examined in this study. In particular, this chapter focuses on the bilateral negotiation between a single IP and a single TSP (the IP-TSP transaction).

The formulation of the IP-TSP negotiation is first presented here. Based on the MAS-ORAM architecture proposed in Chapter 4, each railway stakeholder is represented by one agent. The formulation involves the definitions of the track access rights, the interaction protocol and the objectives of the stakeholders. Optimisation algorithms are then devised for the decision-making processes of the IP and TSP agents respectively, followed by a set of simulation case studies and the corresponding discussions on the results and findings.

5.1. Mathematical Modelling

An IP-TSP transaction is regarded as a one-to-one negotiation on a product between a buyer and a seller. The product under negotiation is the track access rights. The buyer of the track access rights is the TSP while the seller is the IP. Under this context, negotiation is an iterative process in which the two stakeholders take turns to express their requirements on a track access rights until a mutually acceptable agreement is reached, or one of them withdraws from the process. With this description, there are four components in an IP-TSP transaction, namely the track access rights, negotiation protocol, TSP-model and IP-model.

5.1.1. Track Access Rights

A track access rights specifies the conditions for track usage by a TSP. It consists of a schedule describing the train movement in space and time. As a result of the different engineering specifications such as gauge widths and energy consumption, a track access rights also identifies the type of rolling stock to be operated on rails. In addition, during the negotiation, a parameter called flex is established in some countries, such as the UK, to denote the time flexibility with which the IP can revise the train schedule when track or station capacity becomes scarce (Gibson *et al.*, 2002). Flex may be defined as a set of discrete levels where the lowest and highest levels refer to the minimum (0 min) and maximum (say 10 mins) flexibilities to shift a schedule profile respectively. The TSP also has to agree on a payment of track access charge (TAC) in order to obtain the permission for train operation.

A track access rights P is defined in (5.1), where $c \in \{1, 2, ..., \infty\}$ is the TAC (in

$ or other currencies); Ψ is the train schedule as defined in (5.2); $\omega \in \{ \omega_i \mid i = 1, ..., n_\omega \}$ is the rolling stock selected for operation (n_ω is the total number of types of rolling stock); and $\phi \in \{ \phi_i \mid i = 1, ..., n_\phi \}$ is the chosen flex level (n_ϕ is the total number of available flex levels).

$$P = \langle c, \Psi, \omega, \phi \rangle \tag{5.1}$$

A train schedule Ψ consists of a set of IDs $S = \{ s_i \mid i = 1, ..., n_s \}$ identifying the sequence of visited stations (n_s is the total number of train stations). The movement of train in time is described by the service commencement time (i.e. the arrival time at the first station) ζ (in hh:mm), the dwell times at each station $T_D = \{ t_{Di} \mid i = 1, ..., n_s \}$ (in min), and the inter-station runtimes $T_R = \{ t_{Ri} \mid i = 1, ..., n_s - 1 \}$ (in min) between adjacent stations. Hence, Ψ is formally defined as a 4-duple in (5.2).

$$\Psi = \langle S, \zeta, T_D, T_R \rangle \tag{5.2}$$

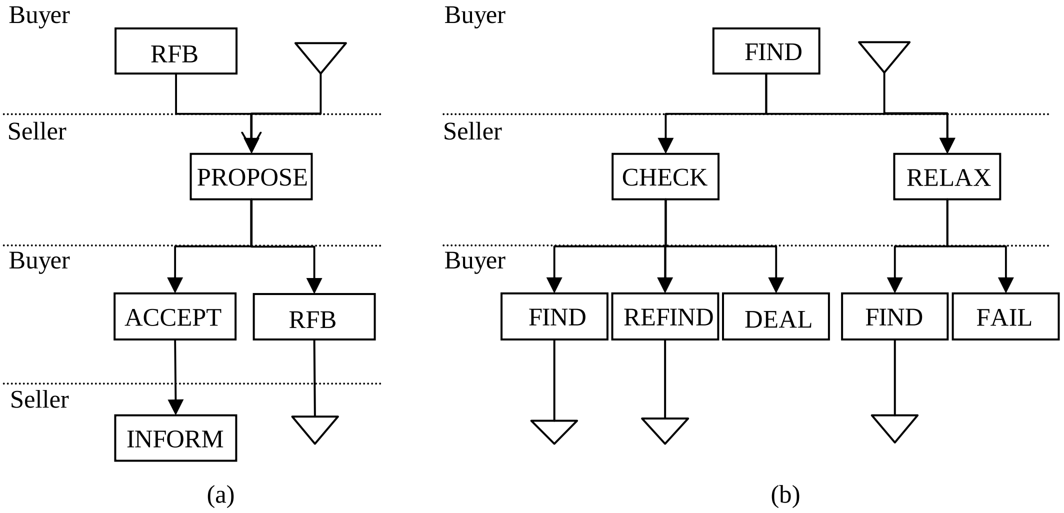
5.1.2. Negotiation Protocol

One approach to classify the various types of negotiation is by the number of parties involved (Luo *et al.*, 2003). Negotiation is regarded as multilateral when there are more than two parties participating in the bargaining process. When only two agents are involved, the negotiation is bilateral. In either case, an interaction protocol is required in agent modelling to specify the actions available to the parties during their communication. The following protocols have been considered for the modelling of the bilateral IP-TSP transactions.

5.1.2.1. Contract Net Protocol

Contact Net Protocol (CNP) (Smith, 1980) is widely used in agent negotiation. This protocol provides a simple yet robust communication procedure to allow the buyer agent to select an appropriate seller agent in a multi-agent system. At the beginning, the buyer sends a request-for-bid (RFB) message to the potential sellers in order to seek for the desired product. This message contains a user-specific description of the product and a deadline for receiving the replies. Upon the arrival of the RFB message, the sellers construct their individual bids and submit (PROPOSE) them to the buyer. After evaluating the received bids, the buyer may award (ACCEPT) the contract to the most acceptable bidder, or refine the requirements on the product and initiate another RFB message. The seller that has been awarded the bid is required to send a confirmation message (INFORM) to secure the contract. The process is summarised in Fig. 5.1a.

CNP is usually applied to multilateral negotiation, where the buyer agent broadcasts the RFB messages to multiple sellers. However, if the buyer targets the message to a specific seller, the negotiation reduces to bilateral.



(a) (b)

Fig. 5.1. Negotiation protocols (a) CNP (b) BSBP

5.1.2.2. Buyer-and-Seller-Behaviour-Protocol

Buyer-and-Seller-Behaviour-Protocol (BSBP) (Luo *et al.*, 2003) is another agent interaction protocol specially designed to model bilateral negotiations on a product possessing multiple attributes. The procedure is depicted in Fig. 5.1b. Initially, the buyer agent expresses its partial requirements using a crisp constraint (inequality), which is enveloped in a FIND message. The message is sent to a specific seller agent whose responsibility is to generate a feasible offer and submit back via a CHECK reply. The offer is accepted (DEAL) if it satisfies the buyer's reserved requirements, and the buyer is willing to comply with (or obey) the restrictions associated with the offer. Otherwise, the offer is rejected. In case of violations of the requirements, a FIND message enveloping a new additional constraint is supplied to the seller agent. In case of unacceptable restrictions, a REFIND message is sent to the seller to ask for a new offer while the original requirements remain.

However, if no feasible offer can be generated in response to a FIND/REFIND request, a RELAX message is issued by the seller in order to prompt the buyer to modify one of the submitted constraints. The buyer may then revise its requirements (FIND) or withdraw from the negotiation (FAIL).

In CNP and BSBP, the negotiation power of the buyer resides in the possibility of refining its product requirements so that the buyer agent is not necessarily confined by the seller's proposed offer. Similarly, the seller has the rights to optimise the offer according not only to the buyer's requirements, but also its internal benefits. This allows both agents to make concessions during the negotiation until their expectations coincide. Otherwise, the negotiation is terminated without any commitment made. Despite the deficiency in modelling multilateral negotiation, BSBP has the advantage of allowing the seller to explore different alternatives under the same set of buyer's requirements before making concessions. This is particularly useful in applications where there often exist multiple optimal solutions. BSBP thus provides a greater negotiation space for these applications.

In fact, the IP-TSP transaction is likely to possess several track access rights that are equally beneficial to the IP. For example, suppose there are two potential offers, P' and P''. When P' has a high access charge with high capacity consumption, P'' may be considered as favourable as P' if it has a cheaper access charge with lower capacity consumption. Despite the lower revenue collection, P'' may allow the IP to utilise the track capacity more efficiently to support more frequent train services. As both offers are considered equally favourable, either P' or P'' may be proposed to the TSP. If the TSP rejects the offer, the IP can propose the alternative in the next negotiation round. Owing to its flexibility in negotiation, BSBP is employed in the IP-TSP transaction.

5.1.3. TSP-Model

According to BSBP, the two main tasks of a TSP agent are to derive the requirements (i.e. a set of crisp constraints) on track access rights and decide how to make concession (i.e. relax the constraints) during negotiation. As pointed out in Chapter 3, most of the current concession-making mechanisms incorporate either a simple or an iterative function (Faratin *et al.*, 1998, 2002; Sim & Wong, 2001; Sim & Choi, 2003; Sim & Wang, 2004) to generate a series of numerical values on an attribute, such as price. These modelling techniques, however, do not generate the crisp constraint as required by the BSBP, but only a set of finite numerical values, which causes difficulties in employing these mechanisms in the TSP-model. Moreover, since the concession functions are monotonic, they are only suitable for applications where the satisfaction of the product varies monotonically with the value of the attribute (e.g. the lower is the price, the higher is the satisfaction). Unfortunately, in the TSP's decision-making process, attributes such as the arrival and departure times of train services do not necessarily vary monotonically, and attributes such as the types of rolling stock and flex cannot be represented as a function at all.

These problems are resolved by modelling the objectives of the TSP agent as a Prioritised Fuzzy Constraint Satisfaction (PFCS) problem (Luo *et al.*, 2003). Instead of bargaining with numerical values, the model enables the derivation of a set of crisp constraints to represent the requirements on an attribute. This introduces the flexibility of employing non-monotonic concession functions to model the satisfaction of the product.

5.1.3.1. Prioritised Fuzzy Constraint Satisfaction Problem

A standard PFCS problem (Luo *et al.*, 2003) is defined as a 4-tuple (X, D, C, ρ) where $X = \{x_i \mid i = 1, ..., n\}$ is a finite set of n variables; $D = \{d_i \mid i = 1, ..., n\}$ is the set of domains; $C = \{R_i \mid \mu_{R_i} : \prod_{x_j \in \mathrm{var}(R_i)} d_j \to [0,1],\ i = 1, ..., m\}$ is a set of m fuzzy constraints ($\mathrm{var}(R_i)$ denotes the set of variables in the fuzzy constraint R_i, and μ_{R_i} is the membership function of R_i); and $\rho : C \to [0, \infty)$ is a priority function.

Given a feasible assignment $v_X = \{v_i \mid i = 1, ..., n\}$ to X, the overall satisfaction is defined by (5.3), where the operator $\lozenge : [0,1] \times [0,1] \to [0,1]$ has the property of $a_1 \lozenge a_2 = (a_2 - 1)a_1 + 1$.

$$\alpha(v_X) = \min_{1 \le i \le m} \left\{ \frac{\rho(R_i)}{\max\limits_{1 \le j \le m}(\rho(R_i))} \lozenge \mu_{R_i}(v_{\mathrm{var}(R_i)}) \right\} \tag{5.3}$$

When the objective of the buyer agent is modelled as a PFCS problem, the buyer is required to express its criteria on a product as a set of fuzzy constraints. The overall satisfaction $\alpha(v_X)$ models the degree of acceptance of the offer with respect to the set of fuzzy constraints. However, the decision on accepting an offer from a seller does not solely depends on $\alpha(v_X)$. In fact, an offer will be rejected if it violates the constraint below, where τ is the accepting threshold.

$$\min\{\alpha(v_X), \beta\} \geq \tau \tag{5.4}$$

$\beta \in [0,1]$ is the degree of obedience on the restrictions of the product imposed by the seller. Even if an offer satisfies the buyer's requirements (i.e. $\alpha(v_X) \geq \tau$), the restrictions may discourage the buyer to accept the offer. As a result, the buyer agent reserves a set of fuzzy propositions $F = \{f_i \mid i = 1, ..., l\}$, where f_i is the degree of obedience on a restriction i, and a truth function $t : F \rightarrow [0,1]$, which are used to determine the overall obedience level of the buyer as shown in (5.5).

$$\beta = t(f_1', ..., f_s') \tag{5.5}$$

5.1.3.2. Decision Variables

When modelling the TSP objectives as a PFCS problem, the quality of the track access charge and schedule times are represented by a set of fuzzy membership functions $\mu_i(x_i) \in [0,1]$, $i = 1, ..., m$, and $x_i \in \{c, \zeta, t_{D1}, ..., t_{Dn_s}, t_{R1}, ..., t_{Rn_s - 1}\}$. A crisp constraint $x_i^a \leq x_i \leq x_i^b$ on an attribute x_i is denoted by the bounds $[x_i^a \mid x_i^b]$. At the beginning of negotiation, the constraints are set at the most preferable values $\hat{x}_i = \arg_{x_i}\{\mu_i(x_i) = 1\}$, $\forall x_i$. A reduction of x_i^a or an increase in x_i^b corresponds to a concession on the attribute. Moreover, a priority value $\rho_i \in [0,1]$, $i = 1, ..., m$ is associated with each attribute to indicate their relative importance to the TSP. Given an offer P' received from the IP agent, the satisfaction of the product is expressed as follows:

$$\alpha(P') = \min_{1 \le i \le m} \left\{ \frac{\rho_i}{\max\limits_{1 \le j \le m}(\rho_j)} \Diamond \mu_i(x_i') \right\}, \text{ for } x_i' \in \{c', \zeta' \, t_{D1}', ..., t_{Dn_s}', t_{R1}', ..., t_{Rn_s-1}'\} \quad (5.6)$$

Rolling stock and flex are modelled as restrictions imposed by the IP agent. If the TSP agent is not willing to comply with the imposed restrictions, the IP is requested to suggest an alternative. To determine whether the TSP should obey the restrictions, two sets of fuzzy values $F_\omega = \{ f_{\omega_i} \mid i = 1, ..., n_\omega \}$ and $F_\phi = \{ f_{\phi_i} \mid i = 1, ..., n_\phi \}$, for $f_{\omega_i}, f_{\phi_i} \in [0, 1]$, are used to indicate the degree of obedience on rolling stock and flex respectively. The overall obedience level of P' is given in (5.7).

$$\beta = \min\{ f_{\omega'}, f_{\phi'} \} \quad (5.7)$$

The objective of the TSP agent is to maximise the satisfaction of the track access rights, subject to (5.8), where $\tau \in [0, 1]$ is the accepting threshold to denote the minimum target satisfaction. $\tau = 0$ gives the highest possibility for successful negotiation because the TSP agent may concede over the entire range specified by the fuzzy membership functions. On the other hand, when $\tau = 1$, the TSP agent will only accept the most preferable schedule defined by the user. Since different TSPs may have different limitations in making concessions, the value of τ should be calibrated according to the requirements of the train planners of the TSP.

$$\min\{ \alpha(P'), \beta \} \ge \tau \quad (5.8)$$

5.1.3.3. Membership Functions

The membership functions on attributes $x_i \in \{ c, \zeta, t_{D1}, \ldots, t_{Dn_s}, t_{R1}, \ldots, t_{Rn_s-1} \}$ are defined by (5.9), where $\hat{x}_i \in \{ \hat{c}, \hat{\zeta}, \hat{t}_{D1}, \ldots, \hat{t}_{Dn_s}, \hat{t}_{R1}, \ldots, \hat{t}_{Rn_s-1} \}$ are the most preferable value of the attributes. $\mu_i^R(x_i)$ is further modelled in (5.10), where x_i^R is the lowest value of x_i that satisfies $x_i^R > \hat{x}_i$ and $\mu_i^R(x_i^R) = 0$. $\mu_i^L(x_i)$ is defined separately for different attributes. $\mu_i^L(x_i)$ for access charge, commencement time and schedule times are modelled by (5.11), (5.12) and (5.13) respectively, where x_i^L is the largest value of x_i that satisfies $x_i^L < \hat{x}_i$ and $\mu_i^L(x_i^L) = 0$.

$$\mu_i(x_i) = \begin{cases} \mu_i^L(x_i) & \text{for } x < \hat{x} \\ \mu_i^R(x_i) & \text{for } x \geq \hat{x} \end{cases} \tag{5.9}$$

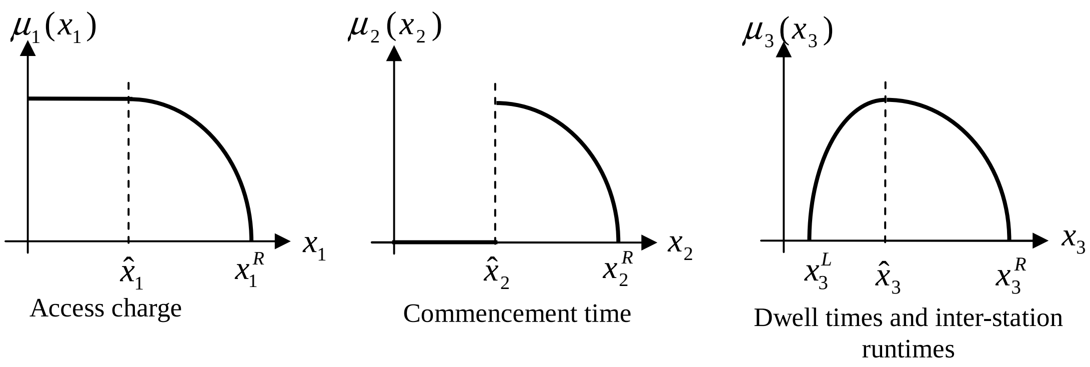
$$\mu_i^R(x_i) = \begin{cases} 1 - \left(\dfrac{x_i - \hat{x}_i}{x_i^R - \hat{x}_i} \right)^2 & \text{if } \hat{x}_i \leq x_i \leq x_i^R \\ 0 & \text{if } x_i > x_i^R \end{cases} \tag{5.10}$$

$$\mu_i^L(x_i) = 1 \tag{5.11}$$

$$\mu_i^L(x_i) = 0 \tag{5.12}$$

$$\mu_i^L(x_i) = \begin{cases} 1 - \left(\dfrac{x_i - \hat{x}_i}{x_i^L - \hat{x}_i} \right)^2 & \text{if } x_i^L \leq x_i \leq \hat{x}_i \\ 0 & \text{if } x_i < x_i^L \end{cases} \tag{5.13}$$

These functions are illustrated in Fig. 5.2. The implications of these definitions are described as follows. For track access charge, $\hat{c}$ is a reasonably low value at which the TSP considers the forthcoming track access rights to be value-for-money. This may be estimated from previous negotiation experience, or deduced from a conservative estimation. Any prices offered lower than this value are considered



Access charge

Commencement time

Dwell times and inter-station runtimes

Fig. 5.2. Illustration on membership functions

equally satisfied. For prices larger than this value, the decrease in satisfaction is modelled as a quadratic function, indicating that the larger is the deviation from the most preferable value, the greater is the drop of satisfaction. Other functions, such as exponential ones, are also feasible but quadratic function is employed for the purpose of simple demonstration.

The preference on the commencement time depends on the earliest dispatching time of the rolling stock. In scheduling, this is often known as the release date. Since the TSP cannot make agreement with the IP if the commencement time of the service is earlier than the release date, the satisfaction is zero when $\zeta < \hat{\zeta}$. In addition, the satisfaction decreases when the commencement time is larger than the release date because the rolling stock will then be idle, implying wastage of resource.

For dwell times, the most preferable values represent the average expectation of the waiting times experienced by passengers. When the actual dwell times exceed these values, the passengers may be annoyed by the additional time required (Murata & Goodman, 1998). The degree of annoyance is also modelled as quadratic equations. Similarly, when the dwell times are less than the most preferable values, the passengers may not have adequate time for alighting and boarding the trains, hence the satisfaction decreases.

In addition to the passenger expectation, the satisfaction of inter-station runtimes is also dependent on energy consumption. It is likely that passengers demand for shorter journey time but this implies higher acceleration and speed, which requires higher power consumption. Preferably, the TSP, which is subject to the maximum permissible speed, tends not to operate their trains too fast to avoid high electricity charge but also not too slow to avoid complaints from passengers.

While it may be argued that other functions are feasible and perhaps more accurate to represent the observed trends on the attributes, quadratic functions are simple and they have been already employed to model passenger expectation on waiting and travelling times in railways (Murata & Goodman, 1998). In addition, these quadratic functions can be easily generated from regression techniques using data collected in surveys on the expectation from passengers or railway planners.

To determine which attribute should be conceded during the negotiation, the TSP agent maintains a set of bounds $\Delta^i_j = [l^i_j, u^i_j]$ in negotiation round j and a constant step-size S_i for each attribute. At the beginning of negotiation, the bounds are set at the most preferable values of each constraint. When a constraint is relaxed in round j, the bounds are updated according to (5.14).

$$\Delta^i_j = \begin{cases} [(l^i_{j-1} - S_i), u^i_{j-1}] & \text{if } \mu^L_i(l^i_{j-1} - S_i) > \mu^R_i(u^i_{j-1} + S_i) \\ [l^i_{j-1}, (u^i_{j-1} + S_i)] & \text{if } \mu^L_i(l^i_{j-1} - S_i) < \mu^R_i(u^i_{j-1} + S_i) \\ [(l^i_{j-1} - S_i), (u^i_{j-1} + S_i)] & \text{if } \mu^L_i(l^i_{j-1} - S_i) = \mu^R_i(u^i_{j-1} + S_i) \end{cases} \tag{5.14}$$

In simple words, when a constraint is relaxed by the TSP, the stakeholder will concede by relaxing the side that yields the least drop in satisfaction. Since the two sides of the membership functions are monotonically decreasing, (5.14) will minimise the loss in TSP's satisfaction when making concession.

5.1.4. IP-Model

In the IP-model, the stakeholder is assumed to maximise the overall track capacity utilisation and revenue collection from all TSPs. The utility function employed by the IP agent is given in (5.15), where U is the utility value (in \$) from the perspective of the IP agent, c is the track access charge (in \$); w_η is the unit valuation of capacity consumption (in \$) and $\Delta\eta$ is the capacity consumed by the train service (no unit). The term $-w_\eta \Delta\eta$ implies a minimisation of capacity usage by the TSP's train service.

$$\max U = c - w_\eta \Delta\eta \tag{5.15}$$

When there is a lack of demand on track capacity, it is reasonable that the IP should sell the capacity at the highest price (i.e. maximise $+w_\eta \Delta\eta$) because the IP may assume that there are no more potential buyers in the worst scenario. When the IP attempts to sell the capacity at the highest cost, it may gain the highest revenue. This situation is likely to occur in non-competitive railway lines.

However, if the IP foresees that other potential buyers are interested in the capacity (as in the case of competitive markets), then the IP should reserve capacity for these buyers. The essence is that if the capacity allocated in a negotiation is minimised, the available capacity remained will be maximised, which allows the IP to negotiate more deals at later stages. For simplicity, this study assumes a constant value of w_η. Nevertheless, since demand changes with time, the value of w_η should also vary accordingly.

This problem in (5.15) is subject to the constraint set Ξ composing of:

i) *Basic Domains of Variables*: $c \in \{1, 2, ..., \infty\}$, $\omega \in \{\, \omega_i \,|\, i = 1, ..., n_\omega \,\}$, $\phi \in \{\, \phi_i \,|$

$i = 1, ..., n_{\phi}\}$, $\zeta \in \{00{:}00, \ldots, 23{:}59\}$, $t_{Dj} \in \{1, 2, ..., \infty\}$, $t_{Rk} \in \{1, 2, ..., \infty\}$, $\forall j, k$.

ii) *Submitted TSP Constraints*: $c^{a} \leq c \leq c^{b}$, $\zeta^{a} \leq \zeta \leq \zeta^{b}$, $t_{Dj}^{a} \leq t_{Dj} \leq t_{Dj}^{b}$, $t_{Rk}^{a} \leq t_{Rk} \leq t_{Rk}^{b}$, $\forall j, k$; and

iii) *Headway Requirements*: $h_{\min} \leq h_{d}$, where $h_{\min}$ and h_{d} are the minimum and actual headway time respectively. In conventional train operation, the actual headway time refers to the time taken for a train to arrive at a certain point along a track (e.g. a station) after the train in front (leading train) has reached the same point. On the other hand, the minimum headway time is the total sum of the minimum braking time, reaction time of driver and equipment in response to a stop signal, and the time taken by the leading train to move by its train length (Hill, 1995). When the actual headway time is larger than the minimum headway time, the train behind is prevented from colliding to the rear end of the leading train. These two terms are usually measured in seconds (in metro systems) or in minutes (in mainline systems). In order to maintain consistency with the resolution of schedule times in this study, these terms are approximated by their ceiling values measured in minutes.

TAC is derived from the sub-charges on track usage, traction energy, peak power demand and congestion. The derivations of these charges and capacity utilisation are described as follows.

5.1.4.1. Track Usage Charge

Track Usage Charge (TUC) recovers the costs of using the track facilities. The charge varies with the amount of maintenance required if the service is allowed to run on the track. In general, the relationship is complex and non-linear as the level of maintenance depends on a number of factors such as the type of rolling stock, the number of vehicles or the weight of train and the maximum allowable speed of train (Dodgson, 1994). To simplify the charging regime, TUC is simply calculated on the total vehicle-kilometre travelled (for passenger services) or the total gross-ton-kilometre travelled (for freight services) in many railway systems. The charge rates vary with different types of rolling stock and they are determined by simulation software, such as mini-MARPAS in the UK (Dodgson, 1994). Having adopted the current charging practice, TUC is defined by (5.16) where c_1^{ω} is the charge rate (in \$/veh·km) for rolling stock ω; n_v^{ω} is the number of vehicles; L_i is the length of track (in km) in inter-station run i.

$$TUC = c_1^{\omega} n_v^{\omega} \sum_{i=1}^{n_s - 1} L_i \tag{5.16}$$

In an IP-TSP transaction, it is assumed that the available types of rolling stock are commonly known by both agents. Each type of rolling stock has a predefined number of vehicles and length. The charge rates are predetermined and are available to the IP agent only.

5.1.4.2. Traction Energy Charge

A power utility company charges the IP according to the units of energy consumed and the peak demand (neglecting the charges to voltage regulation and current distortion due to harmonic effects). Traction Energy Charge (TEC) is levied to recover the units of electricity consumed by a train service. If c_2 is the charge rate (in \$/kWh) for the electricity provision and $E(\omega, t_{Ri})$ is the unit of energy consumed (in kWh) during inter-station run i when rolling stock ω completes inter-station run i at t_{Ri}, TEC is computed by (5.17).

$$TEC = c_2 \sum_{i=1}^{n_s - 1} E(\omega, t_{Ri}) \tag{5.17}$$

For each type of rolling stock, the IP reserves a lookup table in which the energy consumption can be obtained according to its runtimes over a specific inter-station run. This table and the charge rate are available solely to the IP agent.

5.1.4.3. Peak Demand Charge

Peak Demand Charge (PDC) denotes the second component of the electricity tariff. If c_3 is the charge rate (in \$/MW) for the increase in peak power demand at the substation, and $\Delta P(\omega, \Psi)$ is the increase in such demand (in MW) when rolling stock ω is running at schedule Ψ, PDC is calculated by (5.18).

$$PDC = c_3 \Delta P(\omega, \Psi) \tag{5.18}$$

A typical power-demand graph of a train is shown in Fig. 5.3. The peaks correspond to the instants when the train accelerates to a particular speed leading to the highest power. The train speed continues to rise until it reaches the maximum allowable speed at which the power demand becomes relatively constant. At times,

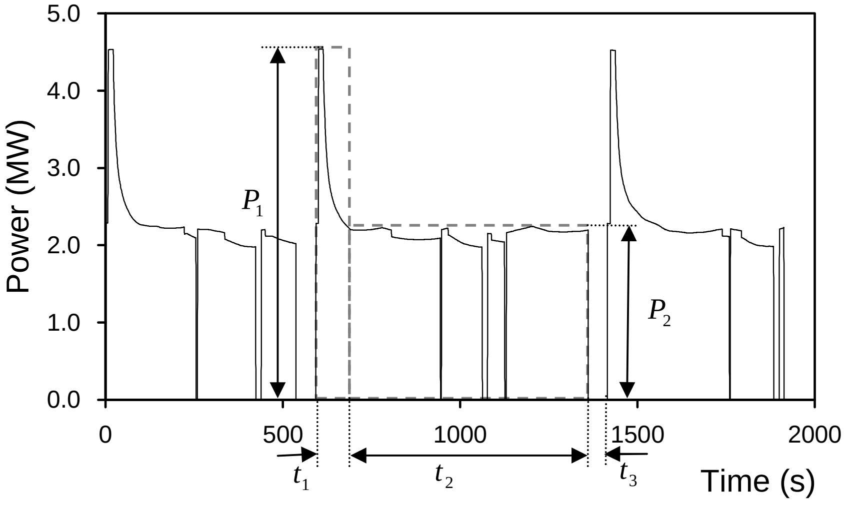


Fig. 5.3. Typical traction power graph for three inter-station runs

the train may be switched to coasting mode, during which the traction motor is turned off and no energy is consumed. Such a demand profile is simplified and modelled as a 5-tuple by (5.19), where t_1 is the time (in min) required for the train to accelerate from stationary to full speed; t_2 is the time (in min) between the first instance of full speed to the instance of braking before the next station; t_3 is the time (in min) required to brake from the maximum speed to a complete halt; P_1 is the maximum power demand attained (in MW) during t_1; P_2 is the maximum demand (in MW) during t_2.

$$\Lambda = \left\langle t_1, t_2, t_3, P_1, P_2 \right\rangle \tag{5.19}$$

Unlike the derivation of energy consumption in TEC, the change in peak demand requires additional information of the other train schedules. With the simplifications in (5.19), the peak demand is calculated by the superposition of demand profiles from all existing scheduled services, as shown in Fig. 5.4.

5.1.4.4. Congestion Charge

Congestion Charge (CGC) is used to recover the expected costs that the IP is

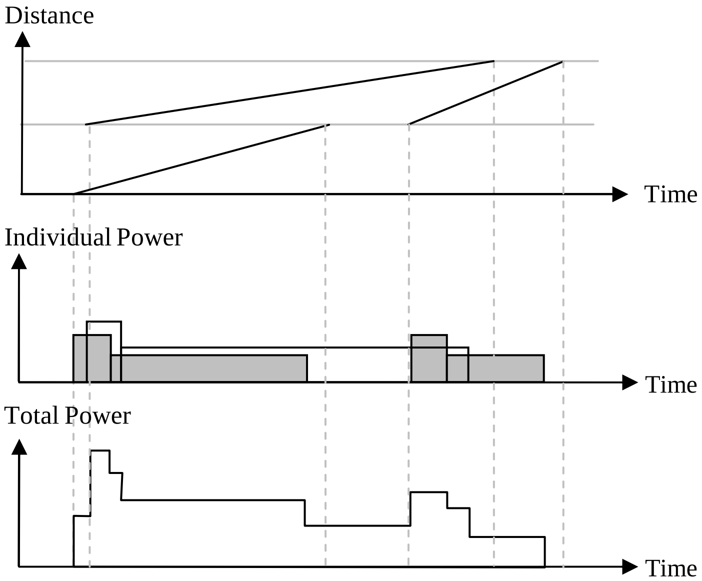


Fig. 5.4. Superposition of peak demand graphs

required to pay the other TSPs when the network becomes congested. In UK, this charge is related to the expected reaction delay resulting from the train service, which is modelled as an exponential function of capacity utilisation (Gibson *et al.*, 2002). Moreover, a TSP is entitled to receive a discount on CGC if it agrees on certain flex levels. If c_4 is the charge rate (in \$/min) for the expected delay caused in the network; d_ϕ is the discount factor associated with flex ϕ; A_i is the track specific constant (in min) at section i; η_i is the resultant capacity utilisation at section i, CGC is computed by (5.20).

$$CGC = c_4 d_\phi \sum_{i=1}^{n_s-1} A_i \exp(\eta_i) \tag{5.20}$$

All the charging factors in (5.20) are exclusive to the IP agent. Capacity utilisation for a single inter-station run is defined in (5.21) and computed iteratively for multiple inter-station runs using (5.22)-(5.24).

5.1.4.5. Capacity Utilisation

Capacity utilisation (CPU) is defined as the ratio of the time taken in operating a set of trains with their minimum headways to the time taken in travelling at their actual timetables (Gibson *et al.*, 2002). Fig. 5.5 illustrates the capacity utilisation for a single inter-station run i within a timeframe W_i (e.g. 30 min). The timetable of a train j is denoted by its departure time t_{Pi}^j at station i and arrival time t_{Ai+1}^j at station $i+1$. Associated with each train is the minimum headway time $h_{\min}$ which includes the time for braking the train from maximum speed to a complete halt, the time taken for the tail of the front train to clear its length and a safety margin for the reaction time of drivers and equipment (Hill, 1995). $h_{\min}$ is represented by the thickness of the parallelograms.

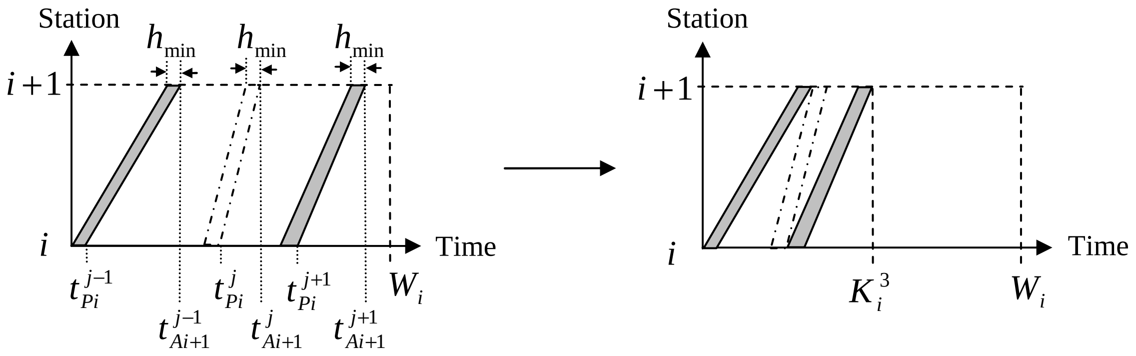


Fig. 5.5. Illustration of capacity utilisation

If these parallelograms are joined together by the vertices as shown in Fig. 5.5, the trains are operating at minimum headway and K_i^n yields the minimum possible time (in min) spanned by the n trains on the track along inter-station run i. Capacity utilisation at inter-station run i is thus defined in (5.21) and the cumulative capacity utilisation of all inter-station runs is defined in (5.22).

$$\eta_i = K_i^n / W_i \tag{5.21}$$

$$\eta = \sum_{i=1}^{n_s-1} K_i^n \Big/ \sum_{i=1}^{n_s-1} W_i \tag{5.22}$$

K_i^n at a particular inter-station run i may be evaluated iteratively for all trains as follows. In computing K_i^2 for two consecutive trains, there are two possibilities as depicted in Fig. 5.6. Case (a) refers to the situation when the train behind is faster and vice versa in case (b). Let t_{Ri}^{j*} be the inter-station runtime for the slower service. In both cases, K_i^2 is computed by (5.23).

$$K_i^2 = 2h_{\min} + t_{Ri}^{j*}, \quad j* = \arg\{ \max_{j=1,2} (t_{Ri}^{j}) \} \tag{5.23}$$

Fig. 5.7 shows the instance when an additional service is operated after the second train. K_i^3 now depends on the relative runtimes of the second and third trains. In

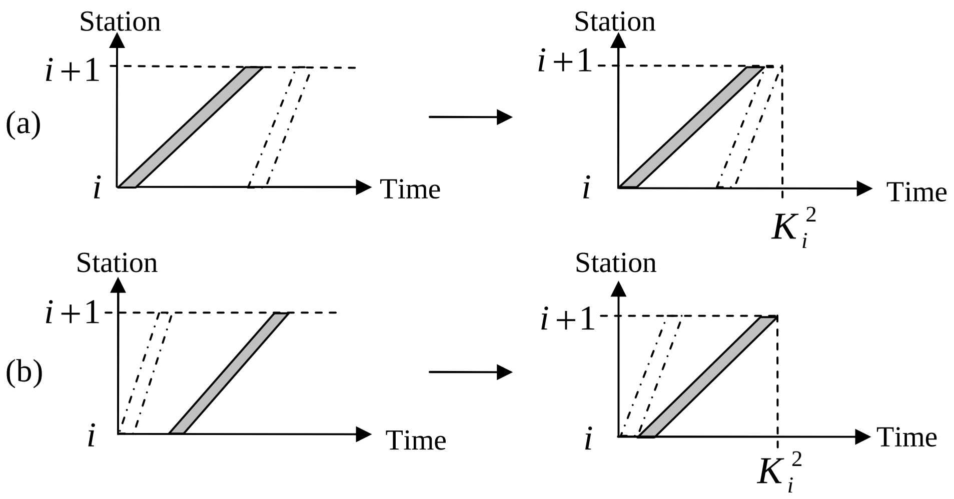


Fig. 5.6. Derivation of K_i^2: (a) train behind is faster; (b) train behind is slower

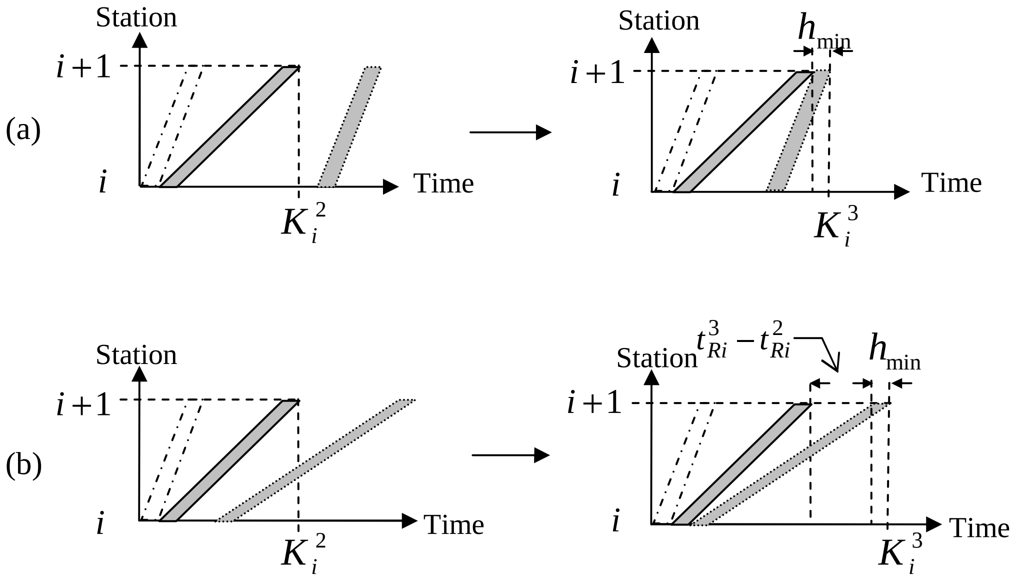


Fig. 5.7. Derivation of K_i^3: (a) train behind is faster; (b) train behind is slower

fact, for all other services in window W_i, K_i^n is computed iteratively by (5.24).

$$K_i^{j+1} = \begin{cases} K_i^j + h_{\min} & \text{if } t_{Ri}^{j+1} \le t_{Ri}^j \\ K_i^j + h_{\min} + t_{Ri}^{j+1} - t_{Ri}^j & \text{if } t_{Ri}^{j+1} > t_{Ri}^j \end{cases}, \text{ for } j \ge 2 \tag{5.24}$$

5.2. Optimisation Algorithms

With the TSP and IP objectives formulated as a PFCS problem and a combinatorial optimisation problem respectively, this section proposes the algorithms to generate the outputs of the agents upon the receipts of the proponent’s responses during the negotiation. In particular, the TSP agent needs to submit the appropriate constraints to the IP agent, and the IP agent should derive the optimal proposal to the TSP agent.

5.2.1. TSP-Model

5.2.1.1. Rules for Replying Behaviour

Based on the TSP-model, an algorithm is required to make inferences from the IP responses defined in the interaction protocol (i.e. CHECK and RELAX) and then deduce the best actions (i.e. FIND, REFIND, DEAL, and FAIL) to reply the IP agent. A rule-based approach has been proposed by Luo *et al.* (2003) and it has been proven that the algorithm is able to lead to a Pareto-optimal solution if the seller agent is proposing offers which maximise the seller’s benefits.

Therefore, the rule-based approach is also employed in this study for the TSP agents. These rules are summarised as follows.

i) *FIND*: A FIND message contains a crisp constraint $x_i^a \leq x_i \leq x_i^b$, where $x_i \in X = \{c, \zeta, t_{D1}, \ldots, t_{Dn_s}, t_{R1}, \ldots, t_{Rn_s-1}\}$. When the TSP agent receives a CHECK or RELAX message, the attribute causing the least reduction in potential satisfaction is selected for submission. Potential satisfaction is the minimum satisfaction achieved by fulfilling all crisp constraints in the submitted set, except the one governing the testing attribute $x_t \in X$. The constraint associated with x_t is replaced by an incremental relaxation defined by (5.14). Thus, potential satisfaction can be computed from (5.6) by creating an instance P' using the bounds x_i^a or x_i^b (whichever causes a lower satisfaction), $\forall x_i$.

ii) *REFIND*: A REFIND message is issued when the offer in a CHECK message has

$\beta < \tau$, where β is defined in (5.7), and τ is the accepting threshold.

iii) *DEAL*: An offer proposed in a CHECK message by the IP will be accepted if the offer satisfies (5.4).

iv) *FAIL*: A transaction will be terminated without any commitment made if the proposed offers in CHECK violate (5.4) and all possible relaxations on the prioritised fuzzy constraints have been exhausted.

5.2.1.2. Behaviour Settings using Effective Priority

Despite the adoption of the rules on selecting the replying behaviour, the issue on setting the priority value ρ_i associated with each attribute x_i has not been explained. This section aims to set up a series of propositions and their proofs so that the TSP stakeholders can set up their agents to perform the desired behaviours. Propositions P1 to P3 are constructed by assuming that the domain of attribute x_i is defined only on one side (left or right) of the most preferable value $\hat{x}_i$. These propositions are further used to construct P4 and P5 in which both sides of $\hat{x}_i$ are defined.

Preliminaries: A one-sided membership function is described by a quadratic function as shown in (5.25), where Δx_i is the deviation of x_i from $\hat{x}_i$, and $\Delta \hat{x}_i$ is the maximum possible deviation from $\hat{x}_i$.

$$\mu_i(\Delta x_i) = 1 - \left(\frac{\Delta x_i}{\Delta \hat{x}_i} \right)^2 \tag{5.25}$$

The satisfaction of x_i is denoted by $\alpha_i(\Delta x_i)$ defined in (5.26), where ρ_i is the priority associated with the attribute, and $\rho_{\max} = \max_{\forall i}\{\rho_i\}$. Substitution of (5.25) into (5.26) yields (5.27).

$$\alpha_i(\Delta x_i) = [\mu_i(\Delta x_i) - 1]\frac{\rho_i}{\rho_{\max}} + 1 \tag{5.26}$$

$$\alpha_i(\Delta x_i) = 1 - \frac{\rho_i}{\rho_{\max}}\left(\frac{\Delta x_i}{\Delta \hat{x}_i}\right)^2 \tag{5.27}$$

In this study, all attributes in a train schedule are discrete, and there is a regular step change S_i between the feasible values of the attributes. For example, dwell times and inter-station runtimes are valid for steps of 1 minute and the allowable change in track access charge is set to an integer, say \$50 or \$100. In other words, it is more appropriate to represent (5.27) with discrete variable $k = \{0, 1, 2, ...\}$ by (5.28).

$$\alpha_i(k) = 1 - \frac{\rho_i {S_i}^2}{\rho_{\max} {\Delta \hat{x}_i}^2} k^2 \tag{5.28}$$

Although (5.28) involves four control variables, the TSP operator may only use ρ_i, S_i and $\Delta \hat{x}_i$ to define the satisfaction because $\rho_{\max}$ is a dependent variable of ρ_i. Since the agent will use the satisfaction values to decide which attribute will be relaxed in the next round of negotiation, these settings will determine the corresponding agent behaviour. However, as $\Delta \hat{x}_i$ and S_i are already used to control the acceptable range of the attributes and the amount of their concession respectively, it is more appropriate to adjust ρ_i to obtain the desired agent response.

Proposition P1: The sequence of initial relaxation on the attributes does not necessarily follow the ascending order of ρ_i.

Proof: According to (5.26), the transformation of $\mu_i(\Delta x_i)$ to $\alpha_i(\Delta x_i)$ is illustrated in Fig. 5.8. The membership function is in fact rescaled vertically by the priority ratio. When a negotiating product possesses n attributes, the shapes of the

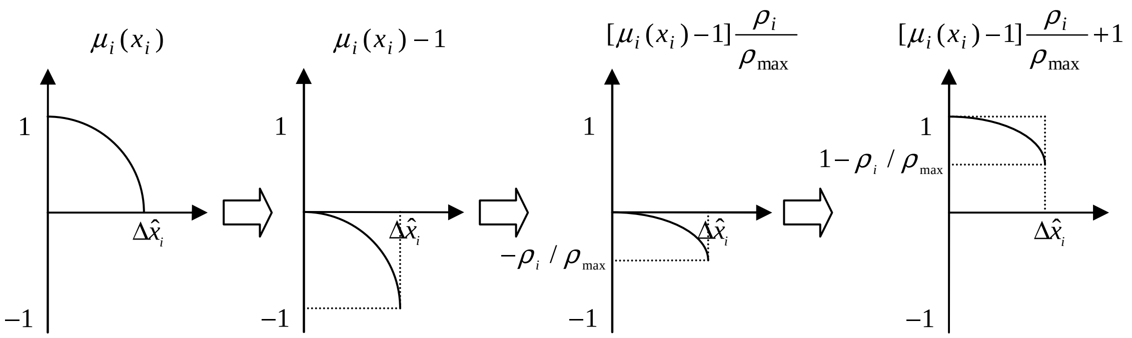


Fig. 5.8. Transformation from $\mu_i(x_i)$ to $\alpha_i(x_i)$

satisfaction functions will be different according to the different limits, step sizes and priorities employed. An example of $n=4$ is shown in Fig. 5.9. If proposition P1 is false, then Fig. 5.9 provides a counter-example.

At the beginning of the negotiation process, the agent will set the attributes to their most preferable values, leading to an initial satisfaction of 1.0. Since the agent is

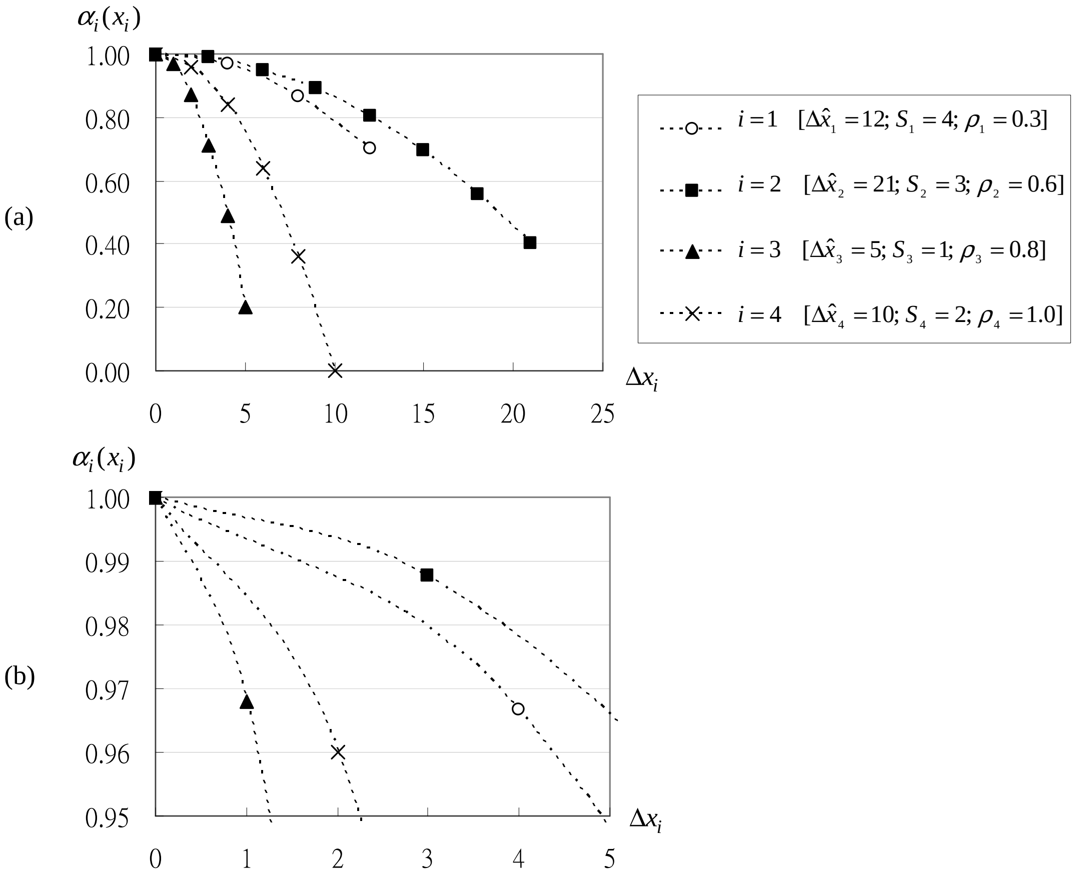


Fig. 5.9. Counter-example to prove P1 (a) complete view (b) enlarged view

minimising the loss in overall satisfaction, the agent will select the attribute which contributes to the smallest decrease in satisfaction. As a result, the attribute order for the initial relaxation will follow the decreasing order of satisfaction at $k=1$, hence the order in Fig. 5.9b is $2 \rightarrow 3 \rightarrow 1 \rightarrow 4$. So, even with the lowest value of ρ_i, attribute x_1 is not the first constraint to be chosen. This completes the proof.

Remark: Although ρ_i is referred as the priority of x_i and it has impact on the agent behaviour, its value does not explicitly indicate the true (effective) priority during negotiation. As indicated in proposition P2, ρ_i is an indirect control variable for the sequence of relaxation.

Proposition P2: The sequence of initial relaxation follows the ascending order of $z_i = \dfrac{\rho_i S_i^{\ 2}}{\Delta \hat{x}_i^{\ 2}}$.

Proof: Suppose $z_i > z_j$ and x_i is relaxed prior to x_j. At their initial relaxations, $k=1$. Since attribute x_i is selected prior to x_j, it follows:

$$\alpha_i(1) \geq \alpha_j(1)$$
$$1 - \frac{\rho_i S_i^{\ 2}}{\rho_{\max} \Delta \hat{x}_i^{\ 2}} \geq 1 - \frac{\rho_j S_j^{\ 2}}{\rho_{\max} \Delta \hat{x}_j^{\ 2}}$$
$$\frac{\rho_i S_i^{\ 2}}{\Delta \hat{x}_i^{\ 2}} \leq \frac{\rho_j S_j^{\ 2}}{\Delta \hat{x}_j^{\ 2}}$$
$$z_i \leq z_j$$

This contradicts $z_i > z_j$, which completes the proof.

Remark: If $z_1 \leq z_2 \leq ... \leq z_n$, the sequence of initial relaxation will be $1 \rightarrow 2 \rightarrow ... \rightarrow n$. In other words, the TSP operator may use z_i as the effective priority of the attributes. A higher z_i indicates an attribute of higher importance.

ρ_i may then be computed accordingly by $\rho_i = \frac{\Delta\hat{x}_i^{\ 2}}{S_i^{\ 2}} z_i$.

Proposition P3: Attribute x_i will be relaxed by m times before the initial relaxation of x_j is relaxed if $m^2 z_i \le z_j \le (m+1)^2 z_i$.

Proof: Suppose $z_j < m^2 z_i$, and x_j has already been relaxed for its first time when x_i is relaxed for the m-th times. The latter assumption implies:

$$\begin{gathered}
\alpha_i(m) \ge \alpha_j(1) \\
1 - \frac{\rho_i S_i^{\ 2}}{\rho_{\max}\Delta\hat{x}_i^{\ 2}} m^2 \ge 1 - \frac{\rho_j S_j^{\ 2}}{\rho_{\max}\Delta\hat{x}_j^{\ 2}} \\
\frac{\rho_j S_j^{\ 2}}{\Delta\hat{x}_j^{\ 2}} \ge \frac{\rho_i S_i^{\ 2}}{\Delta\hat{x}_i^{\ 2}} m^2 \\
z_j \ge m^2 z_i
\end{gathered}$$

This contradicts $z_j < m^2 z_i$. Therefore, when $m^2 z_i \le z_j$, x_i must have relaxed by m or more times prior to the initial relaxation of attribute x_j.

Moreover, suppose $z_j \le (m+1)^2 z_i$, and attribute x_j has already been relaxed for its first time when x_i is relaxed for the $(m+n)$-th times, where $n \ge 1$, this means:

$$\begin{gathered}
\alpha_i(m+n) \ge \alpha_j(1) \\
1 - \frac{\rho_i S_i^{\ 2}}{\rho_{\max}\Delta\hat{x}_i^{\ 2}} (m+n)^2 \ge 1 - \frac{\rho_j S_j^{\ 2}}{\rho_{\max}\Delta\hat{x}_j^{\ 2}} \\
\frac{\rho_j S_j^{\ 2}}{\Delta\hat{x}_j^{\ 2}} \ge \frac{\rho_i S_i^{\ 2}}{\Delta\hat{x}_i^{\ 2}} (m+n)^2 \\
z_j \ge (m+n)^2 z_i
\end{gathered}$$

However, since $z_j \le (m+1)^2 z_i$, so $n < 1$. This completes the proof.

Remark: This proposition may be used as a conditional constraint between several related attributes. For example, for an expenditure-reducing agent, the track access charge should not begin to increase unless the train service has been greatly distorted. In such case, the effective priority of TAC should be set as, for instance, 9 times larger than that of the station dwell times. This means the agent is willing to pay a higher TAC before one of the dwell times is altered for its fourth times.

Preliminaries: Propositions P4 and P5 (defined below) assume that the membership function of x_i is defined by two quadratic functions as shown in (5.29), where Δx_{Li} and Δx_{Ri} are the maximum left and right deviations from the most preferable value respectively.

$$\mu_i(\Delta x_i) = \begin{cases} 1 - \left(\dfrac{\Delta x_i}{\Delta \hat{x}_{Li}} \right)^2 & \Delta \hat{x}_{Li} \le \Delta x_i < 0 \\ 1 - \left(\dfrac{\Delta x_i}{\Delta \hat{x}_{Ri}} \right)^2 & 0 \le \Delta x_i \le \Delta \hat{x}_{Ri} \end{cases} \tag{5.29}$$

The corresponding satisfaction functions for continuous variable Δx_i and discrete variables k are given in (5.30) and (5.31) respectively.

$$\alpha_i(\Delta x_i) = \begin{cases} 1 - \dfrac{\rho_i}{\rho_{\max}} \left(\dfrac{\Delta x_i}{\Delta \hat{x}_{Li}} \right)^2 & \Delta \hat{x}_{Li} \le \Delta x_i < 0 \\ 1 - \dfrac{\rho_i}{\rho_{\max}} \left(\dfrac{\Delta x_i}{\Delta \hat{x}_{Ri}} \right)^2 & 0 \le \Delta x_i \le \Delta \hat{x}_{Ri} \end{cases} \tag{5.30}$$

$$\alpha_i(k) = \begin{cases} 1 - \dfrac{\rho_i S_i^{\ 2}}{\rho_{\max} \Delta \hat{x}_{Li}^{\ 2}} k^2 & \Delta \hat{x}_{Li} \le \Delta x_i < 0 \\ 1 - \dfrac{\rho_i S_i^{\ 2}}{\rho_{\max} \Delta \hat{x}_{Ri}^{\ 2}} k^2 & 0 \le \Delta x_i \le \Delta \hat{x}_{Ri} \end{cases} \tag{5.31}$$

Proposition P4: The sequence of initial relaxation follows the ascending order of

$z_i = \dfrac{\rho_i S_i^{\ 2}}{\max(\Delta x_{Li}^{\ 2}, \Delta x_{Ri}^{\ 2})}$.

Proof: Let $z_{Li} = \frac{\rho_i S_i^{\ 2}}{\Delta\hat{x}_{Li}^{\ 2}}$ and $z_{Ri} = \frac{\rho_i S_i^{\ 2}}{\Delta\hat{x}_{Ri}^{\ 2}}$. Consider the following three cases:

Case 1: When $\Delta\hat{x}_{Li}^{\ 2} > \Delta\hat{x}_{Ri}^{\ 2}$, $z_{Li} < z_{Ri}$. By P2, the left-membership function relaxes prior to the right-membership function.

Case 2: When $\Delta\hat{x}_{Ri}^{\ 2} > \Delta\hat{x}_{Li}^{\ 2}$, $z_{Ri} < z_{Li}$. By P2, the right-membership function relaxes prior to the left-membership function.

Case 3: When $\Delta\hat{x}_{Li}^{\ 2} = \Delta\hat{x}_{Ri}^{\ 2}$, $z_{Li} = z_{Ri}$. In this case, either membership function may be chosen for relaxation.

In other words, $\max(\Delta x_{Li}^{\ 2}, \Delta x_{Ri}^{\ 2})$ indicates whether the initial relaxation of x_i is governed by the left or the right-membership function. If the effective priority is defined as above, the ascending order of z_i indicates the order of initial relaxation. This completes the proof.

Remark: If $z_1 \le z_2 \le ... \le z_n$, the sequence of initial relaxation will be $1 \to 2 \to ... \to n$. In other words, the TSP operator may use z_i as the effective priority of the attributes. A higher z_i indicates an attribute of greater importance. ρ_i may then be computed accordingly by $\rho_i = \frac{\max(\Delta x_{Li}^{\ 2}, \Delta x_{Ri}^{\ 2})}{S_i^{\ 2}} z_i$.

Proposition P5: Attribute x_i will be relaxed by at least m times before the initial relaxation of x_j if $z_j \ge m^2 z_i$.

Proof: Let $a^2 z_{Li} < z_{Ri} < (a+1)^2 z_{Li}$, where $a > 1$. By P3, the left-membership function of x_i will be relaxed by a times prior to the initial relaxation from the right-membership function. Also, let,

$$m^2 z_i < z_j < (m+1)^2 z_i$$

$$m^2 \frac{\rho_i S_i^{\,2}}{\max(\Delta x_{Li}^{\,2}, \Delta x_{Ri}^{\,2})} < z_j < (m+1)^2 \frac{\rho_i S_i^{\,2}}{\max(\Delta x_{Li}^{\,2}, \Delta x_{Ri}^{\,2})}$$

$$m^2 \min(z_{Li}, z_{Ri}) < z_j < (m+1)^2 \min(z_{Li}, z_{Ri})$$

$$m^2 z_{Li} < z_j < (m+1)^2 z_{Li}$$

By P3, the left-membership function of x_i will be relaxed by m times prior to the initial relaxation of x_j. Now, consider the following two cases:

Case 1: $a > m$. By the time of the initial relaxation of x_j, the left-membership function of x_i has relaxed m times. However, the right-membership function of x_i has not been relaxed. So P5 holds.

Case 2: $a \le m$. By the time of the initial relaxation of x_j, the left-membership function of x_i has relaxed m times. In addition, the right-membership function of x_i has already relaxed for its first time. In other words, the initial relaxation of x_j occurs at least after $m+1$ relaxations of x_i. Hence, P5 also holds.

The proof of the argument for $a^2 z_{Ri} < z_{Li} < (a+1)^2 z_{Ri}$ can be constructed similarly. This completes the proof.

5.2.2. IP-Model

5.2.2.1. Combinatorial Optimisation

The maximisation of the utility function in (5.15) is combinatorial because the independent variables are all discrete as restricted by the constraint set Ξ. The common deterministic techniques in solving this kind of optimisation problems are integer linear programming (integer-LP), dynamic programming (DP), and branch-and-bound (BNB) algorithm (Papadimitriou & Steiglitz, 1998). However, integer-LP formulation is not suitable because (5.15) is nonlinear. While DP may handle nonlinearity, it has the limitation that a choice (state) selected for a decision (stage) must be independent to the choices made for subsequent decisions. Unfortunately, the underlying variables (i.e. $\phi, \omega, \zeta, t_{Dj}, t_{Rk}$) in (5.15) are strongly dependent as observed from the definitions in (5.17), (5.18), (5.20), (5.23) and (5.24). Consequently, DP is also not applicable to this IP optimisation problem.

BNB algorithm is based on the notion of intelligently enumerating all the feasible points of a combinatorial optimisation problem (Papadimitriou & Steiglitz, 1998). The solution space of the problem is partitioned into non-overlapping discrete subsets by branching. A subset generated by branching is represented as a node, which defines a relaxed problem to the original optimisation one. Within a node, a bound (a numerical value) is calculated to indicate the best possible solution for its leaf nodes. By appropriately selecting the nodes for expansion, the optimal solution is constructed without exhaustively evaluating all instances.

There are three rules for constructing a solution with BNB. Firstly, if there is no solution to the relaxed problem, there is no solution to the original problem. Secondly, if the solution to the relaxed problem is feasible, it is optimal for the original problem. Finally, if the solution to the relaxed problem is infeasible, the cost at that node provides a bound for its leaf nodes. Therefore, the requirements for resolution are to partition the solution space and to define the relaxed problem. The following subsection specifies a feasible BNB algorithm for the IP optimisation problem.

5.2.2.2. The Basic Branch-and-Bound Algorithm

The partitioning of the solution space is defined by the sequence of branching, which follows the order of variables $\phi \rightarrow \omega \rightarrow \zeta \rightarrow t_{D1} \rightarrow t_{R1} \rightarrow t_{D2} \rightarrow \ldots \rightarrow t_{Rn_s-1} \rightarrow$

t_{Dn_s} in this study. An example of the branching tree is illustrated in Fig. 5.10. In fact, other sequences with different ordering of variables are also possible because the given sequence is only one of the feasible instances. However, the proposed sequence has the advantage of chorological arrangement of the schedule times so that the arrival and departure times at stations can be computed at a node throughout the algorithm. Moreover, the restrictions on flex and rolling stock are considered at the beginning of the sequence to facilitate the reduction of computation demand (see Section 5.2.2.3).

The definition of the relaxed problem is defined as the optimisation of (5.15) when a partial constraint set $\Xi' \subseteq \Xi$ is considered. For example, when the tree is expanded to node M in Fig. 5.10, the constraint set becomes Ξ' = { $\phi = \phi_1$, $\omega = \omega_2$, ζ =07:45, $3 \le t_{D1} \le 5$, $t_{R1} = 15$, $3 \le t_{D2} \le 4$, $h_{\min} = 2$ }. By progressively adding a constraint at each level of the search tree, the global optimal solution is guaranteed. The bound at a node is computed by the sum of maximising the individual sub-charges and minimising the capacity utilisation subject to the associated constraints summarised in Table 5.1. The maximum TUC is identified by comparing the products of $c_1^{\omega} n_v^{\omega}$. For TEC, since maximum energy consumption is achieved by the operation at the minimum

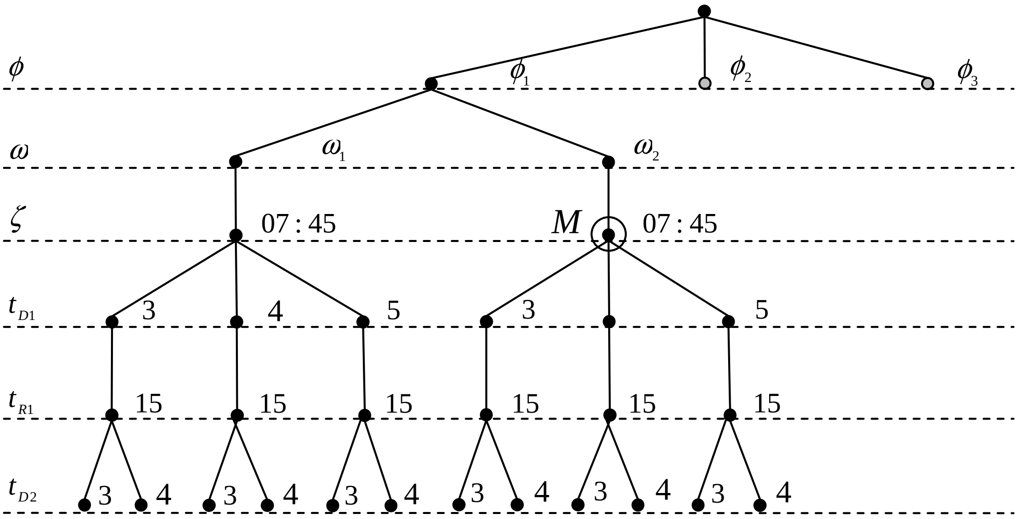


Fig. 5.10. Illustration of BNB search tree

Table 5.1. Objectives Functions and Constraints for Relaxed Problem

Terms	Objective function	Constraints in Ξ'
TUC	$\max\{c_1^{\omega} n_v^{\omega} \sum L_i\}$	Headway; rolling stock
TEC	$\max\{ c_2 \sum E(\omega, t_{Ri})\}$	Headway; rolling stock; inter-station runtimes
PDC	$\max\{ c_3 \Delta P(\omega, \Psi) \}$	Headway; rolling stock; commencing time, dwell times, inter-station runtimes
CGC	$\max\{ c_4 d_{\phi} \sum A_i \exp(\eta_i) \}$	Headway; flex; rolling stock; inter-station runtimes
TAC	$\max(c)$	Cost
CPU	$\min(\Delta\eta)$	Headway; rolling stock; inter-station runtimes

inter-station runtimes, it corresponds to comparing the maximum energy consumption of the available rolling stock when employing the set of shortest runtimes. The maximum PDC is evaluated by exhaustively enumerating the total power demand of all feasible schedules. CGC is maximal when the lowest available discount rate and the rolling stock with the longest cumulative inter-station runtimes (i.e. when congestion is most severe) are employed. The minimisation of capacity utilisation is achieved with the rolling stock travelling with the shortest cumulative inter-station runtimes.

Fig. 5.11 shows the flowchart for the BNB algorithm. k represents the current evaluating node. Initially, k is set to 0, which is the root node of the search tree. This node is inserted in $LIST$ which maintains the potential nodes generated in the algorithm. k^* and U^* record the best node and the corresponding utility value found during the algorithm, which are set to null and zero respectively at the beginning. The algorithm then adopts a depth-first search. If a node have a utility value smaller than the current best value, the node is declared 'fathomed' and the algorithm continues with the next node in $LIST$. Otherwise, the node will be evaluated for its feasibility. If the node is feasible, it is labelled 'lived'. Since its utility value is greater than that of the current best node, k^* and U^* are updated. However, in case of identifying an infeasible solution (e.g. the root node), the node is declared 'expand'. Since its leaf nodes may contain the optimal solution, they are generated and inserted in $LIST$. When all nodes have been evaluated, the best node is returned. If the best node exists,

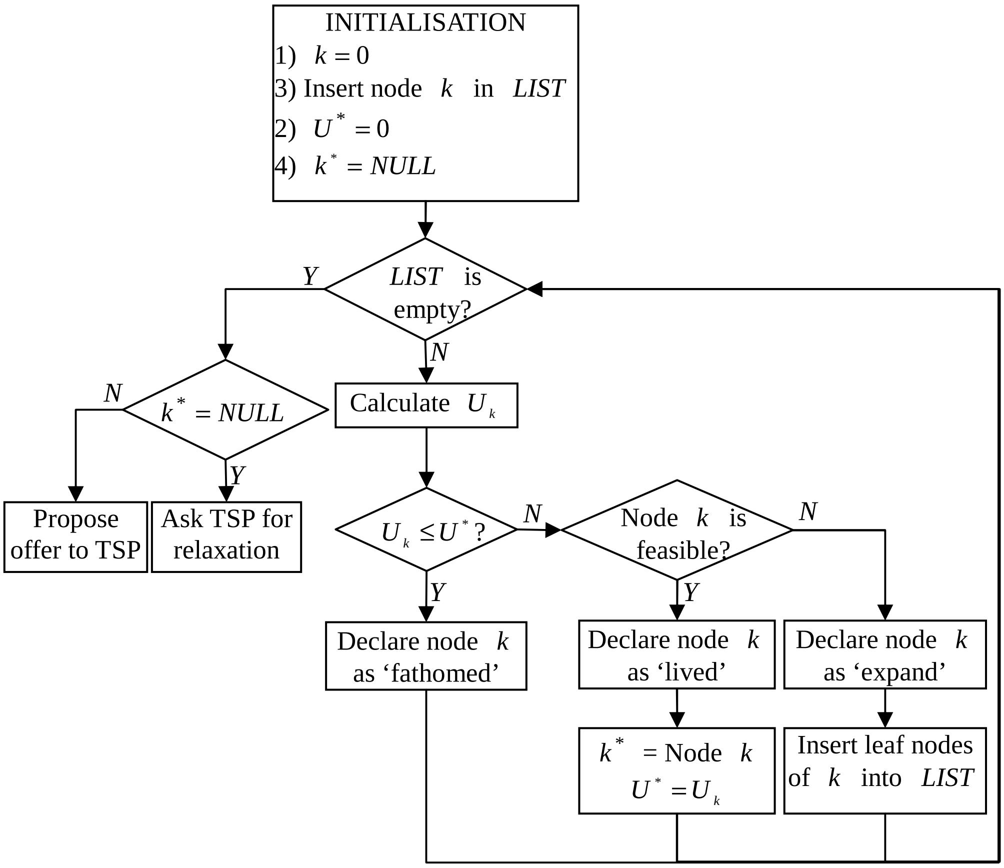


Fig. 5.11. A flowchart of the BNB algorithm for IP agent

the track access rights is proposed to the TSP via a CHECK message. Otherwise, a RELAX message is issued to the TSP.

5.2.2.3. Computation Demand Reduction

In the worst scenario, the computation complexity of a BNB algorithm is no better than an exhaustive search when all nodes are expanded. For the proposed algorithm, the complexity can be shown to be NP with $O(n_\phi n_\omega{}^2 n_\zeta{}^2 n_D{}^{2n_s} n_R{}^{2n_s})$, where $n_D = \max_{1\le i\le n_s}(n_{Di})$ and $n_R = \max_{1\le j\le n_s-1}(n_{Rj})$. The details of proof are given in Appendix A. In other words, the applicability of the algorithm is limited by n_s. In order to generate results within a reasonable time-span, three procedures are incorporated into the basic algorithm to reduce the number of node evaluations and hence the computation demand.

i) *Facilitation of the Most Preferable Schedule*: To minimise the information revealed to the seller agent, the original BSBP only allows the buyer agent to submit one crisp constraint within a FIND message in each negotiation. If the same restriction is imposed to the IP optimisation problem, the schedule times (ζ, t_{Di} and t_{Ri}) will often be unbounded by the TSP, and the problem space is then limited solely by the headway constraints. This sometimes leads to an overwhelming size of domains (Fig. 5.12) which significantly increases the number of node evaluations in the algorithm.

In practice, it is natural for the TSP to express the most preferable schedule at the beginning of negotiation so that the IP may provide a feasible schedule in the proximity of its requirements. With this consideration, the efficiency of the algorithm may be improved by allowing the TSP agent to submit the most preferable schedule during the first round of negotiation (the submission of the TAC constraint is however, not compulsory). Not only does this reduce the number of node evaluations in the algorithm, but the transaction also requires fewer negotiation rounds since those used in submitting the individual constraints are now condensed to a single one.

ii) *Pruning by Headway Constraints*: Despite the facilitation of the most preferable schedule, when the TSP agent progressively relaxes the constraints during the

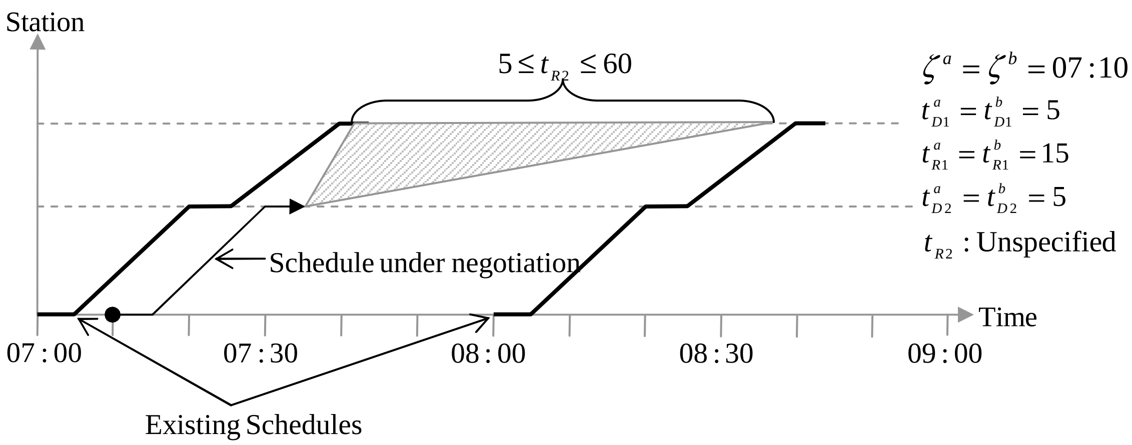


Fig. 5.12. Possible size of domain without specifying the most preferred schedule

negotiation, the problem space for the IP inflates significantly. This often gives rise to substantial computation demand.

Fig. 5.13 shows a special case, when the minimum inter-station runtime ($MIRT_k$) at an inter-station run k is greater than the maximum allowable runtime ($MART_k$) governed by the headway constraints. In such case, the leaf nodes corresponding to the situation are all infeasible. If this condition can be detected prior to the expansion at the node, all leaf nodes can be pruned.

Let EDT_k and LDT_k denotes the earliest and latest departure times at station k respectively. These are computed by (5.32) and (5.33) using the lower and upper limits of the TSP constraints.

$$EDT_k = \zeta^a + \sum_{i=1}^{k} (t_{Di}^a + t_{Ri}^a) \tag{5.32}$$

$$LDT_k = \zeta^b + \sum_{i=1}^{k} (t_{Di}^b + t_{Ri}^b) \tag{5.33}$$

Let EDT_k^j be the earliest departure time and LAT_k^j be the latest arrival time at station k due to the j-th service that arrives at AT_k^j and departs at DT_k^j, for

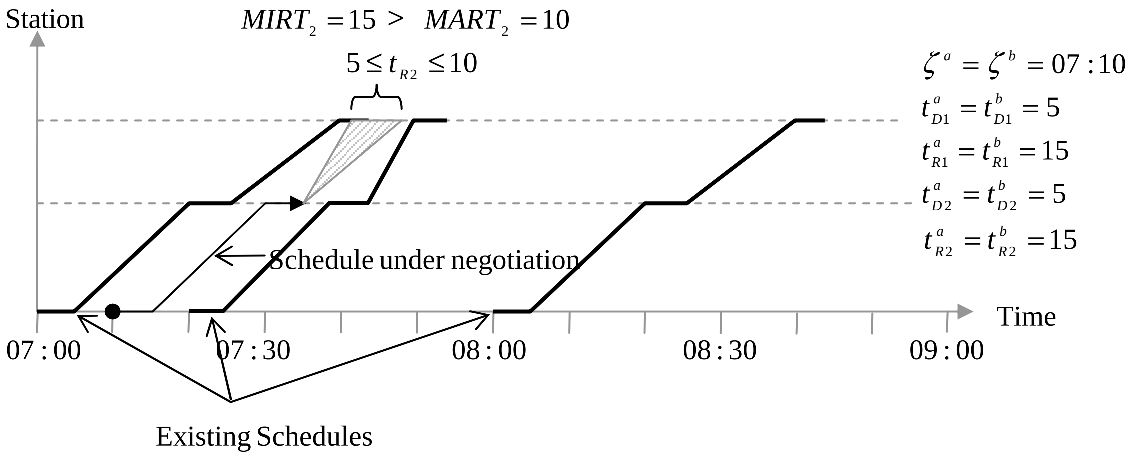


Fig. 5.13. Condition for pruning using capacity constraints

$EDT_k \leq DT_k^j \leq LDT_k$. EDT_k^j and LAT_k^j are computed by (5.34) and (5.35), and $MART_k$ and $MIRT_k$ can then be expressed in terms of (5.36) and (5.37). A node is pruned if $\exists k$, $MIRT_k > MART_k$.

$$EDT_k^j = DT_k^j + h_{\min} \tag{5.34}$$

$$LAT_k^j = AT_k^j - h_{\min} \tag{5.35}$$

$$MART_k = \max_{\forall j}\{LAT_{k+1}^{j+1} - EDT_k^j\} \tag{5.36}$$

$$MIRT_k = t_{Rk}^a \tag{5.37}$$

iii) *Pruning by REFIND Message*: Pruning is also possible when the IP agent receives a REFIND message in the previous round of negotiation. When this occurs, the TSP agent is requesting the IP to generate a new offer based on the previous set of constraints. As the constraint set remains unchanged, the TSP is in fact unsatisfied with the restrictions imposed by the IP, that is ω and/or ϕ. In other words, all nodes that employ the same set of rolling stock and flex level can be eliminated from evaluation. Hence, if ω and ϕ are used as the first two branching parameters, the entire branch beneath the combination is not required for evaluation.

5.3. Simulation Setup

There are three objectives for conducting the case studies described below. Firstly, the studies aim to examine whether the proposed TSP-model is capable of representing TSPs possessing different operation objectives. Secondly, the simulation also investigates whether the IP-model is acting rationally in an IP-TSP transaction in response to requests from different TSP agents. Thirdly, a preliminary study is constructed to examine whether the overall capacity utilisation is affected in a series of these IP-TSP transactions if the IP agent is to negotiate with the TSP agents in a different order.

In all studies, the number of stations is set to four. There are three types of rolling stock and five flex levels available for negotiation. Table 5.2 shows the vehicle numbers and track usage charge rates of the rolling stock, in addition to their relative traction requirements (i.e. energy and power consumptions). The charge rates reflect the degree of track damage incurred by the rolling stock. In Table 5.3, the lowest level ϕ_1 represents no flexibility and each incremental level allows an addition of 2-minute flex time. The flex discount factors reduce CGC by 5% in each successive level.

Twelve case studies have been performed (Table 5.4). These simulations are conducted under the same track configuration consisting of three track sections that connect stations A, B, C and D (Table 5.5). The track length for the middle section is comparatively long and the track specific constants for the first and third sections are higher in order to emphasise the long-distance service provisions between two cities.

In cases 1 to 9, only one IP-TSP transaction is conducted in each case. These transactions serve the purposes of examining the ability of reaching rational agreements.

Table 5.2. Definition of Rolling Stock

Type	Vehicles	Track usage rate c_1^{ω} (\$/veh·km)	Traction level
ω_1	10	0.04	Medium
ω_2	8	0.06	Low
ω_3	9	0.16	High

Table 5.3. Definition of Flex Levels

Level	Flex time (min)	Discount factor
ϕ_1	0	1.00
ϕ_2	2	0.95
ϕ_3	4	0.90
ϕ_4	6	0.85
ϕ_5	8	0.80

Table 5.4. Simulation Cases

Case	IP agent	TSP agent
1	IP-1	TSP-A1
2	IP-1	TSP-A2
3	IP-2	TSP-A3
4	IP-2	TSP-A4
5	IP-3	TSP- A4
6	IP-4	TSP- A4
7	IP-5	TSP-A3
8	IP-6	TSP-A3
9	IP-2	TSP-A5
10	IP-2	TSP-B {8, 2, 9, 1, 5, 3, 6, 10, 4, 7}
11	IP-2	TSP-B {6, 9, 1, 5, 8, 4, 2, 7, 10, 3}
12	IP-2	TSP-B {9, 8, 4, 10, 6, 2, 1, 7, 3, 5}

Table 5.5. Track and Station Data

Track	Origin station	Destination station	Length (km)	Track specific constant (min)
1	A	B	20	1.2
2	B	C	200	1.0
3	C	D	15	1.1

The remaining 3 cases form a preliminary study on an IP agent handling multiple negotiations in a sequential manner. Each of these cases involves 10 IP-TSP transactions, whose order of negotiations are randomly generated.

The definitions of the TSP agents are shown in Tables 5.6 and 5.7. A TSP agent is denoted by the prefix ‘TSP’, followed by a unique suffix representing the name of its train service (e.g. A1, A2, B3, etc.). Each agent is therefore responsible for conducting the negotiation for a single train service which is assumed to be operated by different train service providers. This also means that a negotiation will lead to a schedule for one train if an agreement is reached. The agents given in Table 5.6 are used in cases 1 to 9. Apart from TSP-A1 (which is a passenger-oriented TSP), all the other agents are carrying the objective of reducing expenditure, as reflected by their relatively high effective priorities on track access charge. Agents in Table 5.7 are employed in cases 10 to 12 and they possess a mixture of expenditure-reducing and passenger-oriented objectives. Owing to the limitation in space, the detailed settings are not shown but

Table 5.6. TSP-A Definitions

Attribute	TSP-A1	TSP-A2	TSP-A3	TSP-A4	TSP-A5
$\hat{\zeta}$ (hh:mm)	07:50	07:50	07:50	07:05	07:50
$\hat{T}_D$ (min)	{5, 5, 3, 3}	{5, 5, 3, 3}	{5, 5, 3, 3}	{5, 5, 3, 3}	{5, 5, 3, 3}
$\hat{T}_R$ (min)	{8, 70, 7}	{8, 70, 7}	{10, 75, 9}	{10, 75, 9}	{10, 75, 9}
$\hat{c}$ ($)	1600	1600	1600	1600	1600
z_δ	16	2	2	2	2
z_{T_D}	{36, 36, 36, 36}	{1, 1, 1, 1}	{1, 1, 1, 1}	{1, 1, 1, 1}	{1, 1, 1, 1}
z_{T_R}	{25, 25, 25}	{8, 1, 5}	{8, 1, 5}	{8, 1, 5}	{8, 1, 5}
z_c	1	5	5	5	5
f_{ω_1}	1.0	0.0	0.0	0.0	0.0
f_{ω_2}	0.6	0.6	0.6	0.6	0.0
f_{ω_3}	0.0	1.0	1.0	1.0	1.0
f_{ϕ_1}	1.0	1.0	1.0	1.0	1.0
f_{ϕ_2}	0.8	0.9	0.9	0.9	0.9
f_{ϕ_3}	0.0	0.8	0.8	0.8	0.8
f_{ϕ_4}	0.0	0.0	0.6	0.6	0.6
f_{ϕ_5}	0.0	0.0	0.0	0.0	0.0
τ	0.1	0.1	0.1	0.1	0.1

Table 5.7. TSP-B Definitions

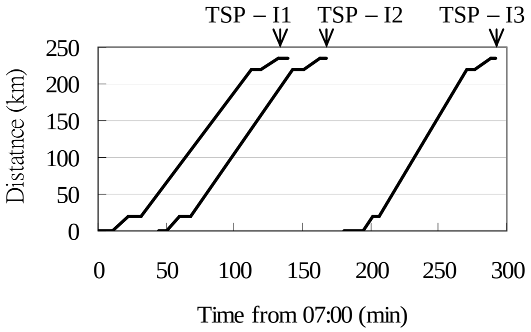
Name	Commencement time limits (min)	Cost limits ($)	Attribute(s) of top priority	Runtime requirements
TSP-B1	[07:00 | 07:10]	[1650 | 2300]	Dwell and run times between A and B	Moderate
TSP-B2	[07:05 | 07:15]	[1900 | 2750]	All schedule times	Short
TSP-B3	[07:20 | 07:30]	[1550 | 2500]	Cost	Moderate
TSP-B4	[07:30 | 07:40]	[1600 | 2850]	All schedule times	Long
TSP-B5	[07:35 | 07:50]	[1800 | 2600]	All schedule times	Moderate
TSP-B6	[07:45 | 07:50]	[1500 | 2300]	Cost	Short (between B and C)
TSP-B7	[07:50 | 08:00]	[1700 | 2500]	Dwell and run times between B and C	Moderate
TSP-B8	[08:00 | 08:20]	[2000 | 2550]	All schedule times	Moderate
TSP-B9	[08:10 | 08:20]	[1750 | 3100]	All schedule times	Long
TSP-B10	[08:15 | 08:30]	[1850 | 2950]	Dwell and run times between C and D	Long

the vital information and objectives are described in Table 5.7.

Table 5.8 summarises the definitions of six IP agents. The initial traffic condition and power distributions associated with these IP agents are shown in Figs. 5.14 and 5.15 respectively. I1 to I5 are used to represent the committed train services prior to the commencement of the negotiations.

Table 5.8. IP Definitions

Attribute	IP-1	IP-2	IP-3	IP-4	IP-5	IP-6
w_η (\$)	8000	5000	10,000	5000	5000	5000
c_2 (\$/kWh)	0.05	0.05	0.05	0.05	0.05	0.05
c_3 (\$/MW)	10.0	10.0	10.0	10.0	10.0	10.0
c_4 (\$/min)	250	250	250	350	250	250
Traffic model	TF-1	TF-2	TF-2	TF-2	TF-2	TF-2
Power model	PD-1	PD-2	PD-2	PD-2	PD-3	PD-4

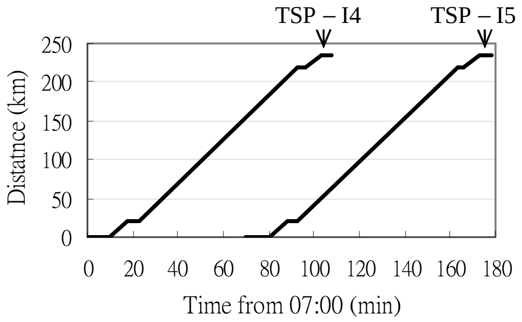


$$\Psi_{I1} = \{ \{A, B, C, D\},\ 07:00,\ \{10, 10, 8, 8\},\ \{12, 80, 8\} \}$$

$$\Psi_{I2} = \{ \{A, B, C, D\},\ 07:45,\ \{10, 10, 8, 8\},\ \{12, 80, 8\} \}$$

$$\Psi_{I3} = \{ \{A, B, C, D\},\ 10:00,\ \{10, 10, 8, 8\},\ \{12, 80, 8\} \}$$

(a) TF-1



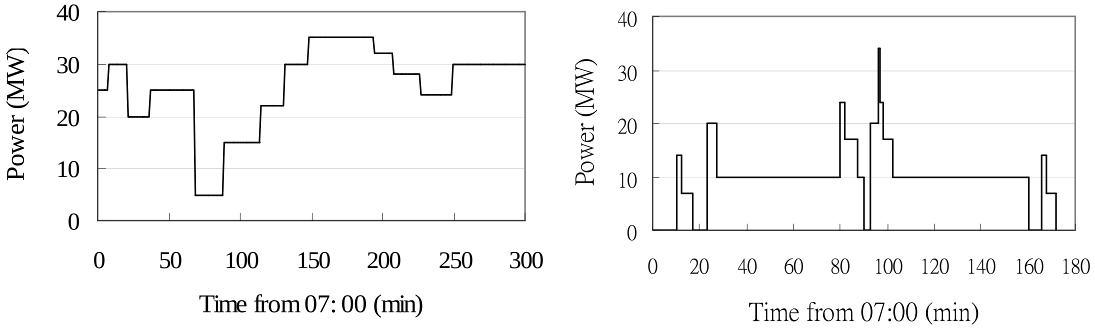
$$\Psi_{I4} = \{ \{A, B, C, D\},\ 07:00,\ \{10, 10, 5, 3\},\ \{8, 70, 7\} \}$$

$$\Psi_{I5} = \{ \{A, B, C, D\},\ 08:10,\ \{10, 10, 5, 3\},\ \{8, 70, 7\} \}$$

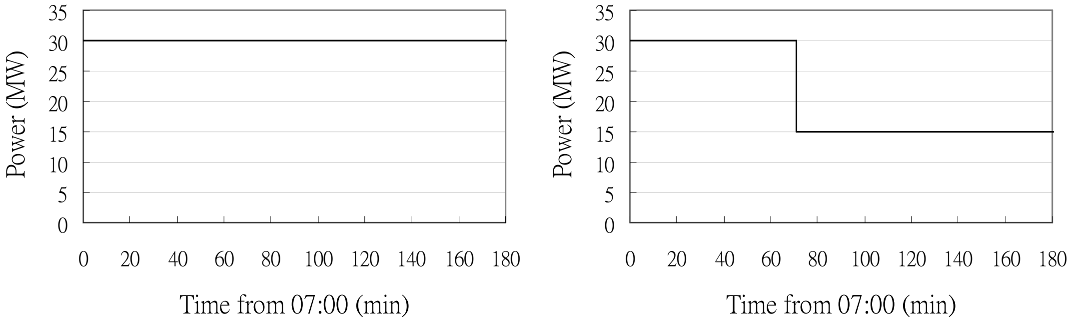
(b) TF-2

Fig. 5.14. Committed train schedule prior to negotiation

All simulations are conducted on a P4 1.6GHz PC and the simulation time is summarised in Table 5.9. Simulation results of the track access agreements in cases 1 to 9 are depicted in Table 5.10 and the resultant timetables of cases 10 to 12 are shown in Table 5.11.



(a) PD-1 (b) PD - 2



(c) PD - 3 (d) PD - 4

Fig. 5.15. Power distribution prior to negotiation

Table 5.9. Simulation Time per Transaction

Time range (min)	Frequency
1 – 10	32
11 – 20	2
21 – 30	0
31 – 60	2
60 – 120	2
120+	1

5.4. Results and Findings

5.4.1. Rational Responses of TSP Agents

According to Table 5.6, the two TSP agents in cases 1 and 2 request the same set of most preferable schedule times and access charge. TSP-A1 denotes a service provider with passenger-oriented operational objective having a strong commitment to punctual station dwell times and inter-station runtimes, while TSP-A2 aims to reduce expenditure and puts a relatively high effective priority on the access charge. Both agents negotiate

Table 5.10. Simulation Results (Cases 1-9): Final Agreements between IP and TSP Agents

	Attribute	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7	Case 8	Case 9
Track access rights	ζ (hh:mm)	07:53	07:50	07:50	07:05	07:05	07:05	07:50	07:50	07:51
	T_D (min)	{5, 7, 3, 3}	{4, 4, 3, 3}	{5, 5, 3, 3}	{7, 6, 3, 3}	{7, 6, 3, 3}	{7, 7, 3, 3}	{5, 5, 3, 3}	{5, 6, 3, 3}	{9, 9, 3, 3}
	T_R (min)	{8, 72, 9}	{8, 72, 8}	{10, 74, 9}	{9, 72, 8}	{9, 72, 8}	{8, 72, 7}	{9, 72, 8}	{10, 73, 9}	{8, 69, 7}
	c ($)	1781	1696	1650	1554	1554	1900	1744	1686	1999
	ω	ω_2	ω_2	ω_2	ω_2	ω_2	ω_2	ω_2	ω_2	ω_3
	ϕ	ϕ_2	ϕ_3	ϕ_4	ϕ_4	ϕ_4	ϕ_4	ϕ_4	ϕ_4	ϕ_4
Breakdown of IP utility value	U ($)	1662	1524	1521	1463	1371	1827	1653	1567	1935
	TUC ($)	113	113	113	113	113	113	113	113	338
	TEC ($)	567	567	561	567	567	567	567	564	671
	PDC ($)	130	90	95	0	0	0	190	130	120
	CGC ($)	972	926	881	875	875	1220	875	879	870
	$\Delta\eta$	0.0149	0.0215	0.0256	0.0183	0.0183	0.0147	0.0183	0.0238	0.0128

Table 5.11. Simulation Results (Cases 10 To 12): Committed Timetables

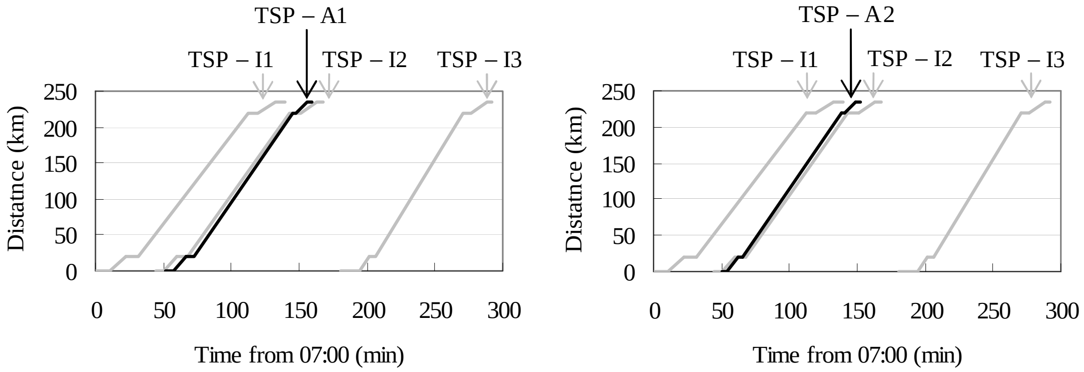
	TSP-B1			TSP-B2			TSP-B3			TSP-B4			TSP-B5		
Case	10	11	12	10	11	12	10	11	12	10	11	12	10	11	12
Arr. at A	07:05	07:05	07:05	07:05	07:05	07:05	07:20	07:20	07:20	07:30	07:30	07:30	07:35	07:35	07:44
Dep. at A	07:12	07:12	07:12	07:08	07:08	07:08	07:25	07:25	07:25	07:38	07:38	07:34	07:40	07:40	07:51
Arr. at B	07:20	07:20	07:20	07:14	07:14	07:14	07:35	07:33	07:34	07:46	07:46	07:45	07:48	07:48	07:59
Dep. at B	07:25	07:26	07:25	07:17	07:17	07:17	07:40	07:38	07:39	07:55	07:55	07:50	07:53	07:53	08:06
Arr. at C	08:35	08:35	08:36	08:22	08:22	08:22	08:51	08:49	08:50	09:05	09:06	09:06	09:03	09:03	09:18
Dep. at C	08:38	08:38	08:39	08:25	08:25	08:25	08:54	08:52	08:53	09:08	09:09	09:10	09:06	09:06	09:21
Arr. at D	08:46	08:46	08:49	08:31	08:31	08:31	09:02	09:01	09:02	09:16	09:17	09:21	09:13	09:13	09:28
Dep. at D	08:49	08:49	08:52	08:34	08:34	08:34	09:05	09:04	09:05	09:19	09:20	09:25	09:16	09:16	09:31

	TSP-B6			TSP-B7			TSP-B8			TSP-B9			TSP-B10		
Case	10	11	12	10	11	12	10	11	12	10	11	12	10	11	12
Arr. at A	07:45	07:45	07:45	07:50	07:50	07:50	08:00	08:08	08:08	08:12	08:10	08:10	08:15	08:29	08:26
Dep. at A	07:49	07:49	07:49	07:55	07:55	07:55	08:05	08:15	08:15	08:15	08:13	08:13	08:18	08:32	08:29
Arr. at B	07:57	07:58	07:58	08:03	08:03	08:03	08:13	08:23	08:22	08:21	08:19	08:19	08:25	08:38	08:35
Dep. at B	08:01	08:03	08:05	08:08	08:08	08:08	08:18	08:28	08:27	08:24	08:22	08:22	08:28	08:42	08:39
Arr. at C	09:10	09:11	09:10	09:24	09:24	09:24	09:28	09:37	09:36	09:30	09:27	09:27	09:33	09:48	09:45
Dep. at C	09:13	09:14	09:15	09:27	09:27	09:27	09:31	09:40	09:39	09:33	09:30	09:30	09:36	09:52	09:49
Arr. at D	09:22	09:23	09:23	09:34	09:34	09:34	09:38	09:47	09:47	09:40	09:36	09:36	09:43	09:58	09:55
Dep. at D	09:24	09:25	09:25	09:37	09:37	09:37	09:41	09:50	09:50	09:43	09:39	09:39	09:46	10:01	09:58

with the same infrastructure provider agent (IP-1) but their most preferable schedules are in conflict with another train service I2 whose travelling profile is given in Fig. 5.14a.

The simulation results of the negotiations are depicted in Table 5.10. TSP-A2 is able to acquire a schedule that overtakes I2 at station B (Fig. 5.16b) at a lower tariff, fulfilling its objective in cost reduction. On the other hand, given the unavailability of track capacity, TSP-A1 is unable to obtain the most preferable dwell times and runtimes, even though it is willing to pay a higher fee. Nevertheless, TSP-A1 has avoided shortening the passenger alighting time at station B by extending the dwell time there, which has resulted in a later overtaking of I2 at station C (Fig. 5.16a).

The variations in these schedules are related to the effective priority assignments. TSP-A1, the passenger-oriented agent, relaxes its constraints on cost ahead of the other attributes. On the other hand, TSP-A2 tends to relax its constraints on schedule times first. When the TSP agents encounter a RELAX message, TSP-A2 will first broaden the feasible range on dwell times, but TSP-A1 will maintain the preferred times and compromise with a higher access charge. Eventually, the difference in behaviour on making concession for the dwell time at station B has resulted in a tighter acceptable range for A1 (5-7 mins) and a broader one for A2 (4-7 mins). It is by employing the



(a) Case 1

(b) Case 2

Fig. 5.16. Timing diagrams for schedules in cases 1 and 2

shorter dwell times of 4 minutes that A2 is able to overtake the conflicting service of I2 at station B.

The schedule secured by TSP-A2 also reduces the TAC mainly through the avoidance of train operation close to 08:30 when the peak demand is highest. This is reflected by the lower PDC of TSP-A2 shown in Table 5.10. In addition, although A2 consumes more capacity, the congestion charge (CGC) is lower when a higher flex level is accepted.

The agents are therefore able to resolve conflicts rationally according to their operating objectives. It is also important to point out that the TSP agents are unaware of either the rights-of-way conflicts with I2 or the existence of the peak demand during the negotiation. Besides, the TSP agents have no cooperative intention to compromise with the IP on such issues. However, by offering different schedules at different prices, the TSP agents indirectly respond to the availability of the market supply.

5.4.2. Pareto-optimal Solutions

In case 3, TSP-A3 is set up to negotiate with IP-2. According to Fig. 5.14b, the most preferable schedule requested by TSP-A3 (Table 5.4) is not occupied by other train services. Despite the availability of capacity, the request is not granted to the TSP in the final agreement (Table 5.10).

To explain the observation, Fig. 5.17 is given to illustrate a simplified search tree at the final round of negotiation. The accepted offer is located at node 146 while the most preferable schedule is located at node N'. In this search problem, any solution employing ϕ_1 to ϕ_3 results in the violation of the cost constraint (i.e. $c \leq 1650$) imposed by the TSP agent. Therefore, all the schedules under nodes 1 to 3 are all infeasible. Similarly, the solution at node 134 (which differs from the final offer by the

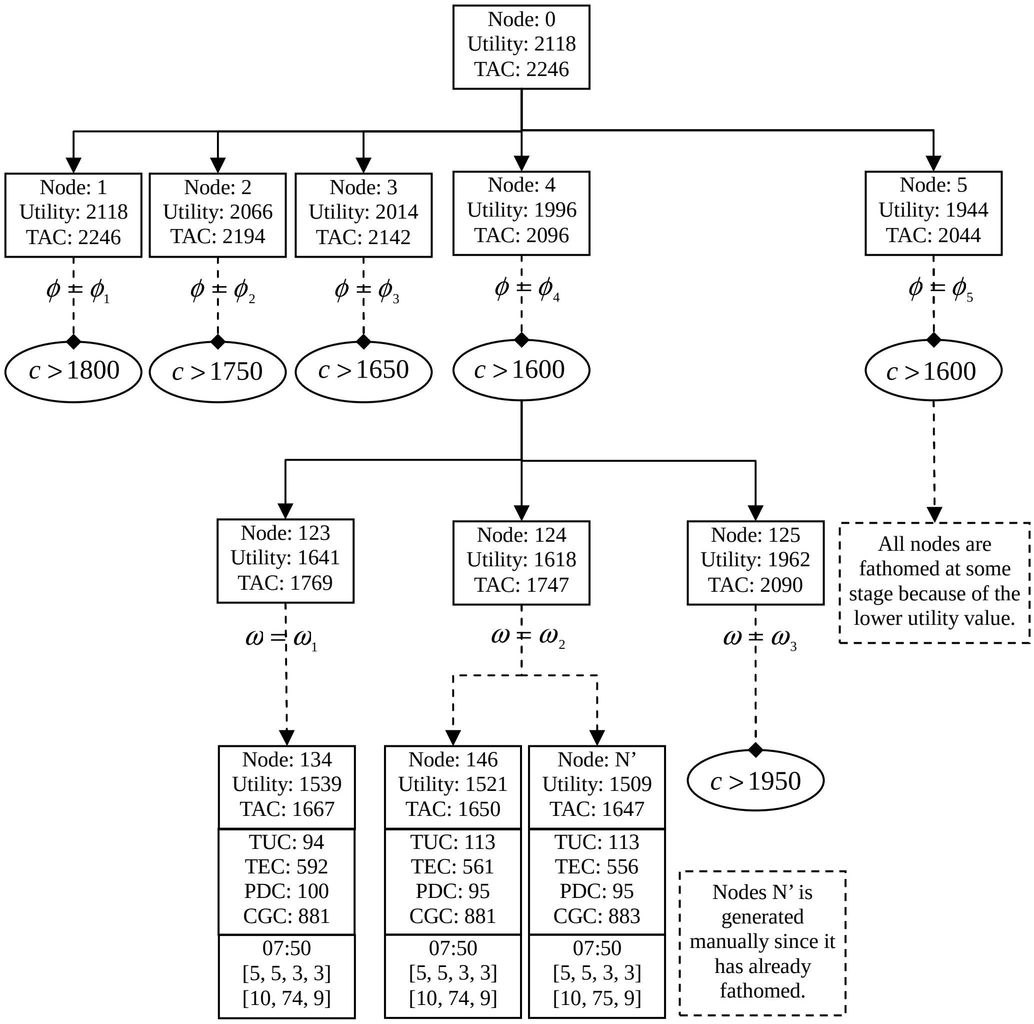


Fig. 5.17. Simplified tree for final round for case 3

type of rolling stock) also exceeds the upper cost limit. The first feasible solution is in fact the optimal solution at node 146. With the adoption of ω_2, the TAC is reduced to \$1650 by the lower energy and power consumption.

The most preferable schedule contained in node N' is also a feasible solution. Since the schedule has a slightly longer inter-station runtime between stations B and C, the TEC is reduced while the CGC is increased. As the change in TEC was greater than that in CGC, the overall TAC is settled at \$1647. Although the cost constraint is satisfied, the lower utility value leads to the rejection of the proposal.

This study therefore demonstrates the ability in reaching a Pareto-optimal (compromised) agreement. In a negotiation among several parties, a solution is Pareto-optimal if any deviations from this solution results in worse payoffs for at least one party (Ehtamo *et al.*, 1996). From the IP's perspective, node 134 is preferred due to its higher utility value, but it is excluded by the TSP's cost constraint. On the other hand, node N' is more favourable to the TSP in terms of the lower TACs, but it is not in the interest of the IP. The Pareto-optimal solution at node 146 is obtained through the use of BSBP (for submitting constraints) and the identification of optimal offer by the BNB algorithm.

5.4.3. Capacity Management

In cases 4 to 6, TSP-A4 is set up to perform an individual transaction with each of the three IP agents, IP-2, IP-3 and IP-4. Case 4 is the reference study, involving the negotiation with IP-2. The capacity weighting used by IP-3 in case 5 is doubled, while a higher congestion charge rate is employed by IP-4 in case 6. These simulations are constructed to examine the effects of raising these settings on capacity utilisation.

According to Table 5.10, apart from the difference in utility value, the track access agreements in cases 4 and 5 are identical. Apparently, the adoption of a higher capacity weighting carries no impact on the resultant schedule and capacity utilisation. Moreover, when the negotiation processes are inspected in detail, the sets of offers proposed during the negotiation in cases 4 and 5 are completely identical and the TSP agent's behaviour (i.e. the sequence of constraint relaxation) is unaffected by the choice of capacity weighting.

In fact, to influence the TSP's response, the protocol allows the IP to propose a different offer during the negotiation process. This may be achieved by any modification in values of TAC, schedule times, rolling stock or flex. However, since both cases employ the same set of charge rates, the TAC of a given set of schedule times and restrictions remains unchanged. According to (5.15), increasing the capacity weighting in case 5 only reduces the corresponding utility value U of the schedules in case 4, thus the rankings of satisfaction of the solutions are preserved. In other words, the IP will generate the same set of offers to the TSP in these two cases. Consequently, it is not surprising to find that the behaviour of the TSPs is identical.

However, raising the congestion rate in case 6 does improve the capacity utilisation. The use of a higher rate causes a more severe penalty on schedules having higher capacity consumption. As the expenditure-reducing TSP is unwilling to pay for an excessive increase in TAC, it settles for shorter inter-station runtimes, resulting in better capacity utilisation.

Therefore, the simulation results suggest that better capacity management can be achieved by increasing the congestion rate. In any open market, the price of a product is often used to manipulate the level of demand. By the same principle, when the intention of better capacity utilisation is reflected on the TAC, the demand on capacity usage may be altered. On the other hand, adjusting the capacity weighting fails to convey the same intention to the TSP agent. Although this may suggest the elimination of the term $-w_{\eta}\Delta\eta$ in (5.15), the term is still required when multiple schedules of equal TAC but different capacity utilisation are present. In these situations, the schedule that consumes the least capacity will be selected in negotiation.

5.4.4. PDC Recovery

Cases 7 and 8 are so constructed that the IP agents in the negotiation differ only by the initial power distributions. In case 7, when IP-5 has a constant power distribution, the TSP agent obtains the track access rights at $1653, of which $190 is the PDC. This is derived from the 49MW of peak demand (Fig. 5.18) when the service departs from station B at 08:09. As a consequence, a step decrease in peak demand is deliberately inserted slightly after 08:09 (at 08:11) in case 8. In this case, the first inter-station runtime and dwell time at station B have been extended, leading to a cumulative delay of 2 minutes. This postpones the departure time at the station B to 08:11, where the decline in peak demand was located. The peak power is reduced to 43MW when the service departs from station A, which lowers the PDC to $130.

Since the IP is negotiating with an expenditure-reducing TSP, the schedule time constraints will usually be relaxed prior to the cost constraint. When the IP encounters this type of negotiating partner, it responds by identifying any schedule with a better premium of TAC. In case 8, a lower TAC is possible through the slight adjustment of the timetable, which reduces the peak demand. By satisfying the buyer’s demand, the likelihood of securing a transaction is increased. On the other hand, in case of negotiating with a TSP who does not permit deviations on schedule times, the IP will offer the original schedule in case 7. The higher burden on the cost of peak demand will then be transferred to the TSP.

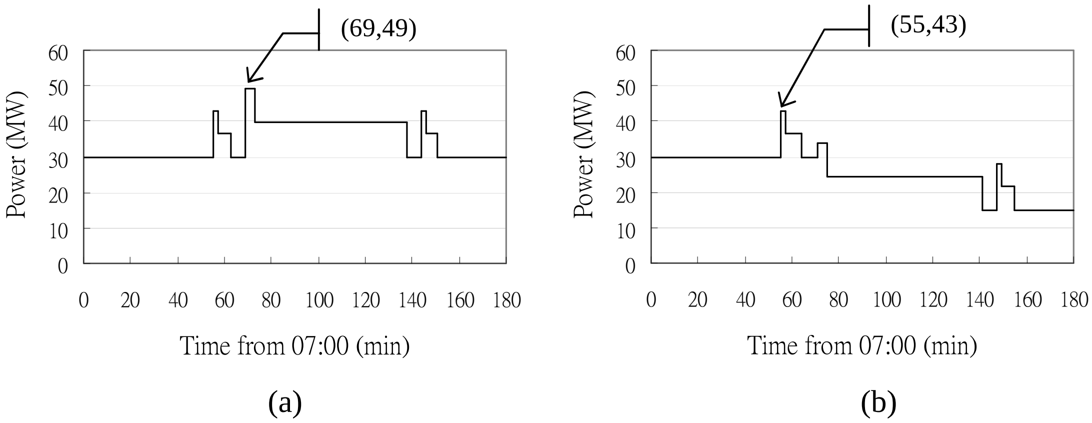


Fig. 5.18. Power distribution after negotiation (a) case 7; (b) case 8

5.4.5. TUC Recovery

Cases 3 and 9 employ the same IP agent but different TSPs. In case 3, TSP-A3 is willing to accept ω_2 and ω_3, but TSP-A5 in case 9 has a more restrictive demand on operating with ω_3 only. Despite the slight modification, there are significant variations in the resulting track access agreements.

As ω_3 causes the most serious tear-and-wear to the rails, it has the highest track usage charge rate of $0.16/veh·km (Table 5.2). This causes an increase in TUC from $113 to $338. Moreover, as ω_3 demands more energy and power, the TEC and PDC also become higher. To reduce the burden of the overall rise in TAC, TSP-A3 accepts shorter runtimes to reduce the CGC. Nonetheless, there is still an overall increase in TAC to $1999 (compared to $1650 in case 3).

Similar to PDC recovery, the IP is acting rationally by transferring the maintenance cost to the TSP. When the negotiating opponent is determined to employ a poor quality rolling stock, the IP increases the TAC so that the cost incurred on track maintenance is recovered or the TSP is put off.

5.4.6. Multiple Bilateral Negotiations

Ten TSPs with different cost and schedule time requirements are set up to compete for capacity over an interval of 3 hours in cases 10 to 12. According to Table 5.11, apart from two train services, B2 and B7, the track access agreements vary when different negotiation sequences are employed.

Several train services are worth for inspection. The schedule times for service B4, B5 and B6 are nearly identical in cases 10 and 11. The timing diagrams for these services in the two cases are shown in Fig. 5.19a. B4 departs from station A at 07:38

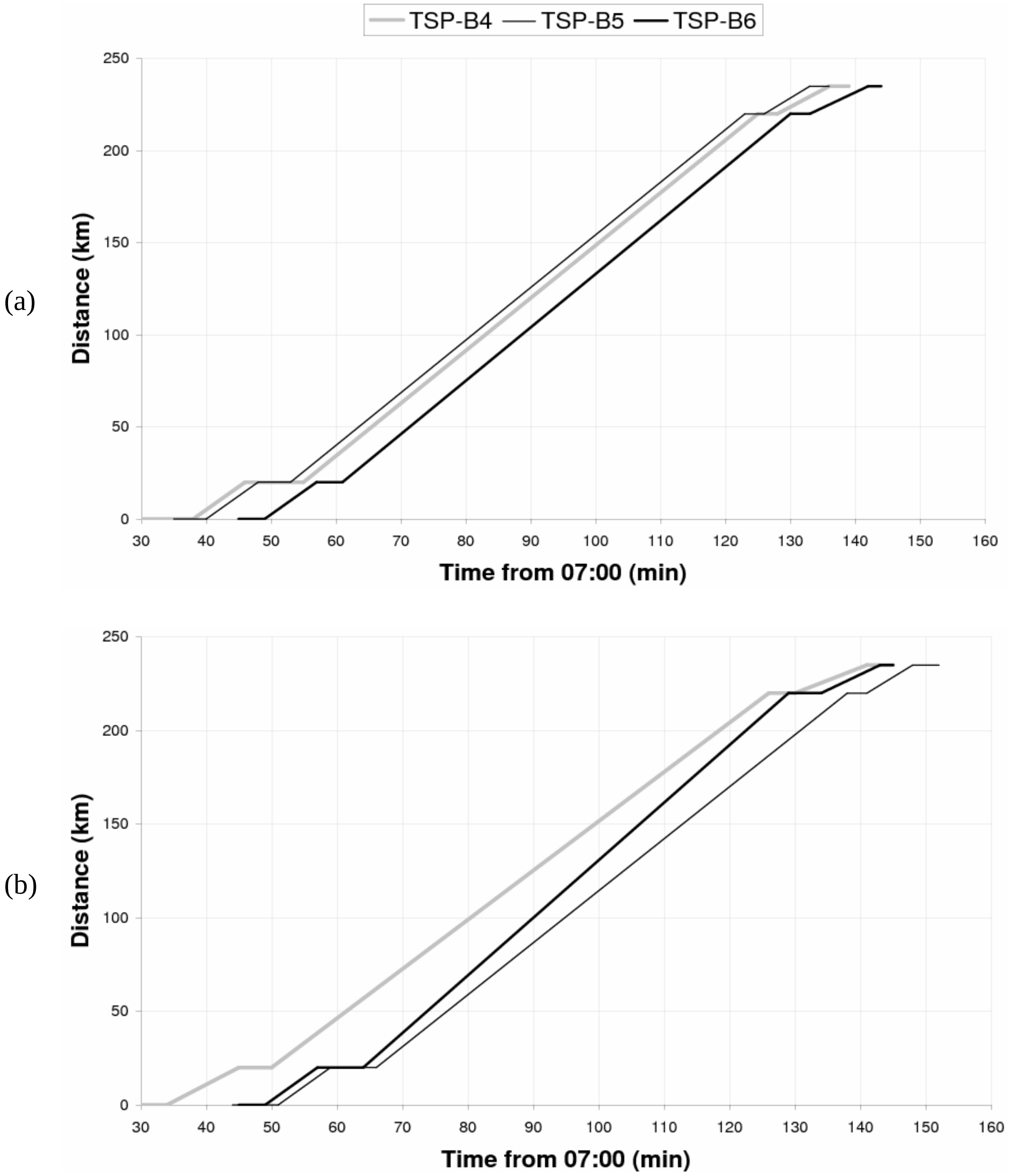


Fig. 5.19. Timing diagrams for B4, B5 and B6 in (a) cases 10 and 11; (b) cases 12

and it is overtaken by B5 at station B at 07:53. B6 leaves station A approximately 10 minutes after B4 and it travels behind B5 throughout the journey. However, a marked difference occurs in case 12 (Fig. 5.19b), B4 departs from station A at 07:34 and the inter-station runtimes are longer. There is no overtaking of B4 by B5, which now operates behind B6. Without the leading effect from B5, B6 is able to operate with faster inter-station runtimes.

The above result is a direct consequence of the IP's negotiation order of TSP agents. The sequences in cases 10 to 12 are TSP-B{5→6→4}, TSP-B{6→5→4} and TSP-B{4→6→5} respectively. In the first two cases, the negotiation with TSP-B4 is conducted last. By the time TSP-B4 has been served, the requested train capacity has already allocated to TSP-B5 and TSP-B6. TSP-B4 therefore needs to accept shorter inter-station runtimes and gives way to the faster service of TSP-B5 when it arrives at station B. When TSP-B4 is served first in case 12, the IP agent is able to satisfy its requirements on longer runtimes. The next service of TSP-B6 can also be scheduled with its preferred (short) runtimes because the two services are separated by sufficient distance. Nevertheless, as B6 gradually reduces the separation from B4 at the approach of station C, the remaining capacity is inadequate for B5 to operate between the two services. As a result, B5 is scheduled to run behind B6.

Similarly, the negotiation order for TSP-B8, TSP-B9 and TSP-B10 in case 10 is TSP-B{8→9→10}, and TSP-B{9→8→10} is the order used in cases 11 and 12. In case 10, TSP-B8 is able to obtain an early commencement time when capacity is available. The allocation of capacity causes more restrictions to TSP-B9, which needs to settle for small deviations in commencement time and runtimes. Although this only leaves a limited amount of capacity for TSP-B10 to operate its service between B9 and I4 (one of the initial services), it is still possible to operate the service tightly behind B9 owing to their similar runtime characteristics. In cases 11 and 12, as the negotiation with TSP-B9 is conducted before TSP-B8, TSP-B9 can now obtain its required capacity, but the service of TSP-B8 has to be scheduled behind it. In addition, since both B8 and I4 are running with moderate runtimes, B10 cannot utilise the remaining capacity between the two services. Eventually, B10 is delayed so that it is operated after I4.

The above results are in fact consistent with the timetables achieved by the scheduling principles adopted in practice. Experience suggests that if there are conflicts in the rights-of-way between train services, the service considered first usually has an advantage. Train planners often exploit this by scheduling according to the priority of services. In this application, as trains are progressively scheduled, there are more constraints to be considered. The first TSP in the sequence is therefore more likely to obtain its preferred requirements. Conversely, when several trains have already been allocated on the track, a competing TSP will probably need to compromise with less favorable schedules. In addition, when a TSP has its service postponed, there may be a knock-on effect to the subsequent transactions.

Another observation from the result is on scheduling non-homogenous traffic. From Fig. 5.20, sequencing the negotiation of TSP agents as in case 10 consumes the least track capacity, whereas the configuration in case 12 requires the highest capacity. In case 10, the better capacity utilisation is achieved by first scheduling the moderate-speed train (B5), and then the faster (B6) and slower (B4) trains. By

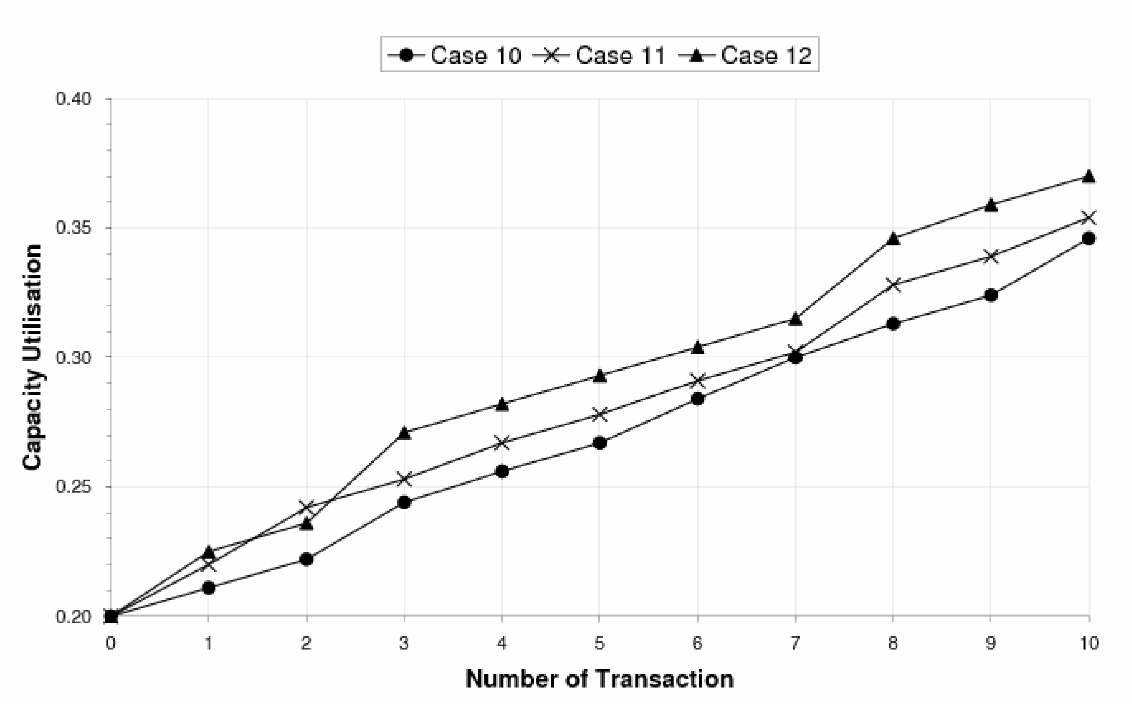


Fig. 5.20. Evolution of capacity utilisation in cases 10 to 12

selecting the moderate case as a reference service, the compromise on homogeneity may be shared by the two extreme services. Otherwise, one particular TSP could have been overburdened, in which case the service might not be scheduled at all. Furthermore, capacity is also improved by sequencing TSPs with similar servicing characteristics together (e.g. TSP-B9 and TSP-B10 in case 10). Conducting a transaction with considerably different train speeds (e.g. TSP-B8) between these TSPs will consume more capacity than required.

5.4.7. Simulation Time

Table 5.9 summarises the time required by the simulation cases. The length of simulation depends on the computational complexity in generating the optimal solutions with each negotiation round and the number of rounds required in each transaction. It should be noted that the majority of cases requires less than 10 minutes to accomplish a transaction and only three cases require more than an hour to reach an agreement.

While the above simulation time is reasonable for the chosen small-scale scenarios, the length of simulation will undoubtedly increase geometrically since the BNB algorithm is inherently NP as discussed in Section 5.2.2.3. Even with the adoption of three heuristic procedures, the overall algorithm is unlikely to sustain when the problem size increases with more stations. However, it is usually at localised track sections that fierce competitions for track access are observed. The BNB algorithm is therefore suitable to conduct critical analysis in these areas prior to the physical implementation of a regulatory or operational adjustment. In addition, since the simulation tool is not intended to be used for real-time scheduling, the order of simulation time (even in hours) for small scale studies is still reasonable.

To evaluate a system involving a large number of stations, parallel processing techniques (e.g. employing dual-core machines) may shorten the simulation time. Nevertheless, since computing power only grows in polynomial order while the problem size increases exponentially, the use of parallel processing techniques still poses limitations on the level of improvement on computation time. An alternative approach is to incorporating more efficient algorithms (e.g. genetic algorithm and simulated annealing) which can generate near-optimal solutions within a specified timeframe. Although these algorithms cannot guarantee the optimal solution, in competitive markets, a speedy negotiation with more flexible (i.e. sub-optimal) solutions is preferred over the absolute optimal solution that requires extensive time of generation.

5.5. Remarks

This chapter has presented an agent model for an IP-TSP negotiation in open railway access markets. With the aid of BSBP (the negotiation protocol), the rule-based TSP reasoning model (for solving the PFCS problem), and the BNB algorithm (for the combinatorial optimisation of the IP) that incorporates rules to reduce computation demand, simulation of the negotiation activities has been made possible. In addition, results have shown that the behaviour of the agents is rational, and the agents are competent to achieve their desired objectives.

In particular, when the model is incorporated with the BSBP, simulation results have demonstrated the ability of the agents to arrive at Pareto-optimal solutions that are beneficial to both parties. By modelling the TSP’s objectives as a PFCS problem, the TSP agent is able to determine the sequence of constraint relaxation that minimises the loss in making concessions. In addition, TSPs of different operational objectives (e.g. passenger-oriented and expenditure-reducing) can be represented by appropriate settings on the effective priority values. Using the BNB algorithm, the IP is also able to reflect the costs of track maintenance, peak power, and traffic congestion on the track access charge, so that the resultant schedules may recover the actual cost imposed by the train services. The results on handling multiple bilateral negotiations by the IP agent also confirm the competitive advantage on the first-served TSP and the difficulties in scheduling non-homogenous traffic demand. As these findings are supported by the in-depth examination on the flow and proof of algorithms employed by the IP and TSP agents, they can be generalised to other railway systems with different parameter settings.

A practical railway network often has a complex track layout consisting of multiple tracks and junctions to support both unidirectional and bidirectional traffic. Even though the approach presented here may not be applicable to solve the entire scheduling problem in such a network owing to the assumption of the simple track configuration (single-track for unidirectional traffic) and the restriction by the exponential growth of computation demand, it is certainly useful for operations at localised track sections where fierce competitions for track access among a number of TSPs are quite common. The models and algorithms devised here are therefore suitable to conduct critical analysis in these areas prior to the physical implementation of a regulatory or operational adjustment. In addition, since the simulation tool is not intended to be used for real-time scheduling, the order of simulation time (even in hours) for small scale studies is still reasonable.

The study also provides a foundation for further research in modelling railway open access markets by multi-agent systems. With the implementation of this core IP-TSP transaction in an open market, further research opportunities are in twofold. Firstly, studies may investigate the possibility of devising more sophisticated and efficient scheduling algorithms. The adoption of a heuristic algorithm (such as genetic algorithm and simulated annealing) is certainly a potential means to reduce the computation time (yet with no guarantee of optimality). The second direction of research is the continual modelling of negotiations occurred in open access markets. For instance, more structural research may be undertaken to investigate different strategies (e.g. first-come-first serve, highest potential TAC first) to sequence the bilateral negotiations so that objectives, such as capacity utilisation and cost recovery, may be optimised.

The simulation cases conducted in this chapter therefore highlight the ability of the IP agent to respond rationally to a single TSP agent at a time and confirm that the IP agent is able to obtain different train schedules when it is negotiating with various TSP agents under the same set of traffic conditions. The next chapter will investigate how the IP agent will allocate capacity to two or more competing TSP agents using different scheduling strategies and how the different settings of TSP agents may influence the decision making of the IP agent using statistical analysis.

Chapter 6

Multilateral Negotiation for Track Access Rights

A bilateral negotiation between an IP and a TSP on track access rights allocation has been modelled in Chapter 5. While it is important to study the basic interaction between the two parties, it is also vital to examine the effects on the IP and the quality of train services when competition between TSPs arises (i.e. the IP-TSP^n transaction). As a result, a preliminary study on sequencing a set of IP-TSP negotiations was conducted in the previous chapter. Simulation results have suggested that the ordering of these bilateral negotiations does indeed have an impact on the overall capacity utilisation. Therefore, this chapter continues to explore suitable but simple rules on ordering the negotiations (i.e. sequencing policies) for the IP.

A transaction-based generation approach is employed by the IP agent to manage the order of the bilateral negotiations in the IP-TSP^n transaction. Apart from this generation approach, other types of management techniques are also discussed. Nevertheless, the rest of this chapter will concentrate on devising a mathematical model for the transaction-based IP-TSP^n negotiation. In addition, a statistical analysis for evaluating two sequencing policies against four performance indices is proposed. A set of randomly generated case studies is then performed and the simulation results are discussed.

6.1. Problem Description

The process of network timetable development in the UK was briefly reviewed in Chapter 2. To generate track access rights offers for the TSPs, the IP may employ either combinatorial or sequence generation approach during the draft-timetable modification process. In the discussions, it was pointed out that sequence generation is preferred over the combinatorial approach if the timetable generation is too complex for the IP to meet the scheduling deadline.

Nevertheless, when considering the detail implementation of sequence generation in a negotiation process, there are in fact several interleaving mechanisms, such as transaction-based and round-based generations, together with pre-reception and post-reception generations (Fig. 6.1). Thus, before a mathematical model for IP-TSP^n transaction is presented, the definitions and merits of these mechanisms are discussed in greater depth.

6.1.1. Combinatorial and Sequence Generations

In combinatorial generation, the IP collects bids from all TSPs and then collectively derives a set of feasible train schedules, taking consideration of all constraints imposed

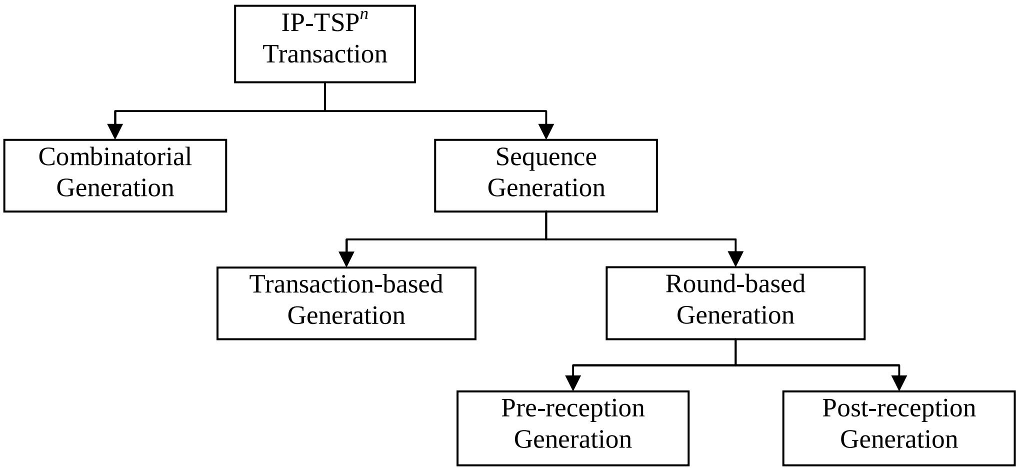


Fig. 6.1. Interleaving mechanisms in the IP-TSP^n transaction

by the TSPs. These offers are then proposed to the TSPs and the IP waits for their replies. Afterwards, the process is repeated with a new set of constraints.

In sequence generation, the IP also collects bids from the TSPs. Instead of devising the set of schedules collectively, the IP first derives a sequence whose order represents the chronological arrangement in which the TSPs are to be served. Using this approach, the IP is not attempting to identify the best way of allocating the capacity, but to negotiate with a proper TSP that is likely to result in better capacity management.

The optimisation problem in combinatorial generation clearly requires high computation demand since the problem space is substantial when all train services are considered at once. As demonstrated in the IP-TSP transaction, there are many feasible schedules even for a negotiation involving only a single train service and the proposed BNB algorithm also demands high computation effort when the service covers a number of stations. In other words, when there are multiple train services, the total permutation of feasible schedules are undoubtedly excessive so that identifying the optimal set of schedules becomes impractical even if there are only few trains and stations.

Therefore, sequence generation is a more practical approach since it only considers one IP-TSP transaction at a time. The requirement of generating a sequence for the TSPs greatly reduces the search space when compared to combinatorial generation. However, regarding to the optimality attained, it is anticipated that the combinatorial approach should obtain better performance because it always considers the entire set of solutions and selects the optimal one.

6.1.2. Transaction-based and Round-based Generations

When employing the sequence generation approach, the resulting sequence may lead to two different interpretations. Firstly, it may indicate the order of securing the transactions with the TSPs (transaction-based). In such case, the IP will complete a transaction successfully, or otherwise, before commencing another transaction with another TSP. Secondly, the sequence may indicate the order in which the IP should propose offers to the TSPs within a negotiation round (round-based). The IP will therefore propose an offer to another TSP immediately after the completion of the current negotiation round with a TSP (but not necessarily the completion of a transaction). After all TSPs have been served in a given round, the IP generates another sequence again and continues to negotiate with the TSPs according to this new sequence. The IP will therefore interleave between the TSPs during the bargaining process.

The computation demand involved in transaction-based generation is obviously lower than that in round-based generation. Hence, the choice between transaction-based or round-based generation depends mainly on the complexity in generating the sequence of TSPs. If the computation demand in deriving the sequence is not substantial, then the latter approach is expected to yield better solutions since it allows the flexibility to update the sequence in successive rounds of transaction. Such activity therefore allows the IP to consider the updated information on the TSPs' requirements and hence the IP will be more readily responded to the changes of TSPs, leading to more favourable solutions.

6.1.3. Pre-reception and Post-reception Generations

Under round-based generation, there are still two alternatives. The IP may either start to propose offers to the next TSP as soon as the proposal to the current agent is sent, or begin the next negotiation after a reply from the current agent has been received. In the pre-reception case, the IP will not receive the updated information from the current TSP. Thus, the IP may assume the current TSP either accepts the offer (capacity is secured) or rejects the offer (capacity is released). A higher degree of uncertainty therefore exists. In the post-reception case, the IP will be certain whether the proposed capacity has been secured or released so that when the IP approaches the next TSP, the offer generated will be optimal at that instant.

6.2. Mathematical Modelling

Among the various mechanisms of generating a negotiation sequence, combinatorial generation and round-based generation are likely to involve excessive simulation time, especially when the IP-TSP transaction has been shown to be NP. As a result, the transaction-based generation is employed in this study because of its lower computation demand and the possibility to reuse the IP-TSP transaction model developed in the last chapter. In this section, the agent interaction protocol is defined, and two sequencing policies are proposed for incorporating in transaction-based generation so that the IP agent can manage the negotiations with a group of TSP agents. In addition, four measuring indices are defined for performance evaluation of the sequencing policies.

6.2.1. Multiple Buyer and Seller Behaviour Protocol

Fig. 6.2 shows the Multiple Buyers and Seller Behaviour Protocol (MBSBP). It is modified from the BSBP (Luo *et al.*, 2003) employed in the IP-TSP negotiation. Since there are multiple buyer agents (TSPs) in the simulated environment, the seller agent (IP) is required to inform the buyers when it is ready to receive their bids. The announcement is achieved by an RFB (Request-for-bid) message. Thus, from time $t = 0$, the IP agent periodically broadcasts an RFB message to the TSP agents registered

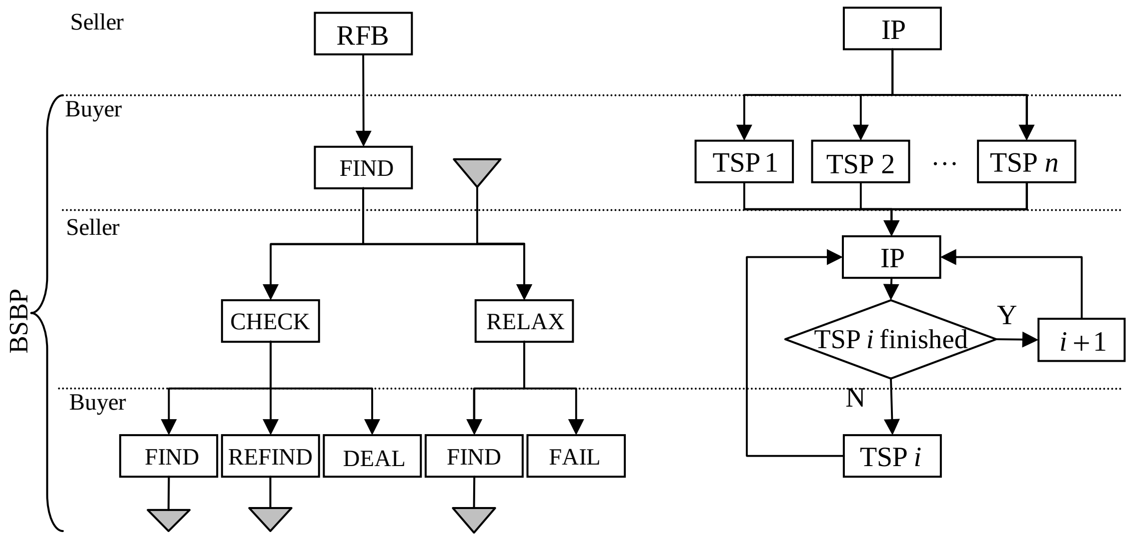


Fig. 6.2. Illustration of MBSBP in a transaction-based generation approach

on the directory facilitator (DF). The message contains a Commencement Time Period (CTP) and a submission deadline. CTP is simply a time window (e.g. from 07:00 to 07:59). Should a TSP wish to operate train services commencing in the stated time window, the TSP is asked to submit their bids before the submission deadline. Bids that are received after the deadline are not guaranteed to be processed.

When the submission deadline is reached, the negotiation activities between the IP and TSP follow the standard BSBP. However, the IP needs to evaluate the bids to determine an appropriate negotiation sequence for maximising its potential benefits.

6.2.2. Sequencing Policies for IP Agent

6.2.2.1. First-Come-First-Serve

The first sequencing policy proposed in this study is First-Come-First-Serve (FCFS), which is adopted in many scheduling practices in railway traffic control (Ho *et al.*, 1997) and computer engineering (Winograd & Kumar, 1996). In the context of the IP-TSP^n transaction, the sequence produced by FCFS indicates a chronological order of TSP agents that corresponds to the time when the bids from the TSP agents are received by the IP agent. The procedure of deriving a FCFS sequence is depicted in Fig. 6.3.

Step 1: IP agent receives a message from a TSP agent. The message contains a performative, an action, a bid (expressed in terms of a set of constraints) and a conversation ID (a reference number for the ease of identifying the negotiation with a particular TSP). This message is stored in the message queue.

Step 2: If the current time exceeds the submission deadline T_{end}, the process stops, and the sequence $FCFSList$ is ready to be used. Otherwise, go to step 3.

Step 3: The message is retrieved from the message queue, and a timestamp is created.

Step 4: The message and the timestamp are inserted into the sequence $FCFSList$. Go to step 1.

FCFS is simple for implementation. Negotiation employing this sequencing policy can in fact begin as soon as the messages are received without the need to wait until T_{end} is elapsed.

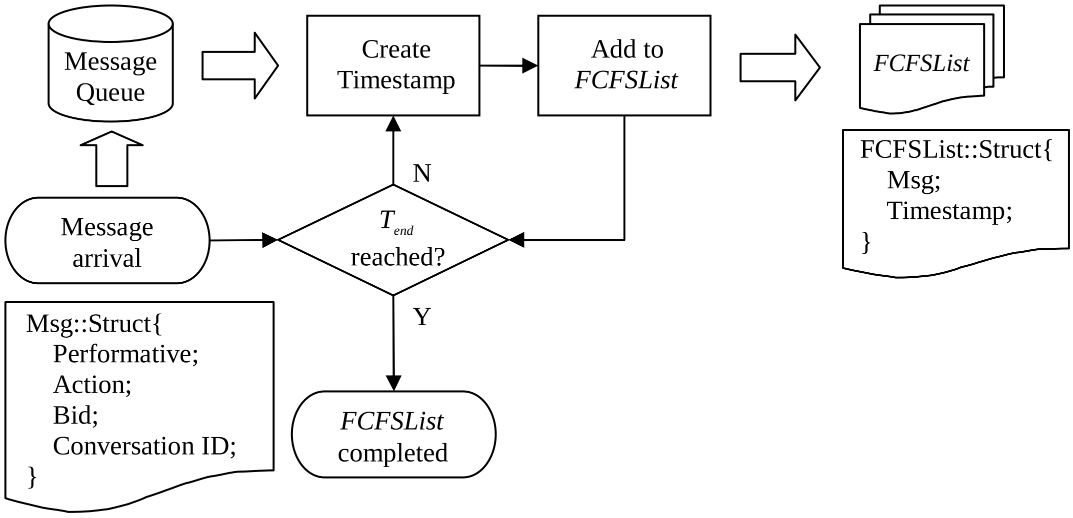


Fig. 6.3. Flowchart on FCFS sequencing policy

6.2.2.2. Highest-Willingness-to-Pay-First

The second sequencing policy is Highest-Willingness-to-Pay-First (HW2PF). In this policy, the TSP agents are ordered according to the potential track access charge recovered (Fig. 6.4).

Step 1: From the $FCFSList$ generated from the flowchart illustrated in Fig. 6.3, count the total number of messages in the list (i.e. N_{bid}) and set the counter $i = 1$.

Step 2: If $i = N_{bid}$, sort $HW2PFList$ according to the descending order WP_i, which is the highest expected willingness-to-pay by the i-th TSP agent. Stop the procedure and the sequence $HW2PFList$ is ready for use. Otherwise, go to step 3.

Step 3: Retrieve message i from $FCFSList$.

Step 4: Obtain c_i^p which is the track access charge requested by the i-th TSP agent specified in its constraint set. If it is unavailable, assume $c_i^p = 0$.

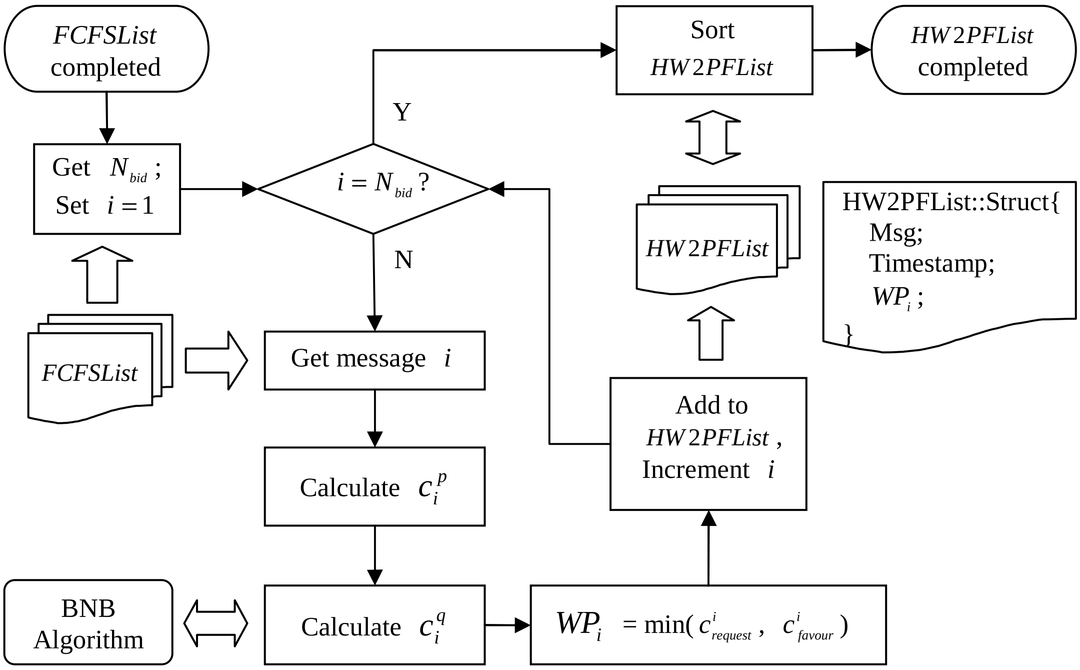


Fig. 6.4. Flowchart on HW2PF sequencing policy

Step 5: Calculate c_i^q using the BNB algorithm defined in Chapter 5. c_i^q is the track access charge of the optimal offer that can be proposed to the i-th TSP agent under the existing traffic condition.

Step 6: Compute $WP_i = \min(c_i^p, c_i^q)$.

Step 7: Add the message and the associated WP_i to the sequence $HW2PFList$. Increment i. Go to step 2.

The implementation of HW2PF is slightly more complicated than FCFS, and it requires the evaluation of c_i^q using the BNB algorithm. HW2PF therefore needs more computation time. However, due to the intention of assigning higher priority to the TSPs having higher willingness-to-pay, the IP agent is more likely to obtain a better cost recovery at the end.

6.2.3. Measuring Indices

Having devised the interaction protocol and the sequencing policies, the IP agent is able to handle the multilateral negotiation with a group of TSP agents. However, for the purpose of this study, in order to compare the performance of the sequencing policies quantitatively, it is necessary to obtain a set of measuring indices. As a result, four indices, total IP utility (IPU_T), average IP utility (IPU_A), extension in journey time (EJT_θ), and deviation from regularity (DFR_θ), are used to determine the effect of the operation of the sequencing policies on the IP and the overall quantity of services to the consumers.

6.2.3.1. Total and Average IP Utility

The first index is the total IP utility, IPU_T. It is the cumulative sum of the IP utility obtained from all transactions as shown in (6.1). U_i (in \$) is the utility value of the i-th successful negotiation, and n_k is the total number of successful negotiations in the N_{bid} transactions. A higher value of IPU_T indicates a higher revenue for the IP.

$$IPU_T = \sum_{i=1}^{n_k} U_i \tag{6.1}$$

The second index is IPU_A, which is the arithmetic mean of the IP utility among the successful transactions. It is computed by (6.2) and it yields a high value if IPU_T is large and/or the number of successful transactions is small.

$$IPU_A = \frac{1}{n_k} \sum_{i=1}^{n_k} U_i \tag{6.2}$$

6.2.3.2. Extension in Journey Time

In general, from the perspective of passengers and freight customers, a better quality of service is perceived when the journey time can be reduced. Thus, EJT_θ measures the average deviation in journey time of a train service operated by a TSP of type θ (i.e. freight, regional, intercity, etc.) from its desired schedule. For a TSP operating a set of n_s services, EJT_θ is defined by (6.3), where $\tilde{t}_i^{\,j}$ and $\hat{t}_i^{\,j}$ (in min) are the actual and expected inter-station runtime of train i between stopping station j and $j+1$ respectively.

$$EJT_\theta = \frac{1}{n_\theta} \sum_{i=1}^{n_\theta} \sum_{j=2}^{n_s} \max(\tilde{t}_i^{\,j} - \hat{t}_i^{\,j}, 0) \tag{6.3}$$

When $EJT_\theta = 0$, the schedule is described as 'without extension' because all trains arrive at the stations no later than the time requested in the bid. When EJT_θ takes a definite value other than zero, the schedule is said to be 'extended', in which one or more of the services suffers from extension in journey time. In addition, when $n_\theta = 0$ (i.e. no trains can be scheduled), EJT_θ is undefined, which is represented by $EJT_\theta = \infty$.

6.2.3.3. Deviation from Regularity

From the viewpoint of a commuter, trains are preferred to arrive at a station at equally spaced time intervals. Any deviations, either earlier or later, may lead to discontentment arising from overcrowding at platforms and trains. Let DFR_θ be the mean deviation from regularity of TSP θ at all stopping station j defined by (6.4). $\hat{n}_\theta$ is the expected number of trains in an one-hour operation, n_θ is the actual number of trains in service, t_i^j (in min) is the arrival time of the i-th train at station j, and $t_{n_t+1}^j = t_1^j + 60$, which assumes the timetable repeats in the subsequent hour.

$$DFR_\theta = \frac{1}{n_\theta} \sum_{j=1}^{n_s} \sum_{i=1}^{n_\theta} \left| (t_{i+1}^j - t_i^j) - 60 / \hat{n}_\theta \right| \tag{6.4}$$

When $DFR_\theta = 0$, the schedule is referred as 'periodical' because the TSP operates trains with equally spaced time intervals at all stations. When DFR_θ takes a definite value other than zero, the schedule is said to be 'non-periodical'. In addition, when $n_\theta = 0$ (i.e. no trains can be scheduled), DFR_θ is undefined, which is represented by $DFR_\theta = \infty$.

6.3. Statistical Analysis

Instead of employing a case-based analytical approach, a statistical approach is used to study the quality of schedules generated from different sequencing policies. In order to simulate a competitive scenario in open markets, the simulator needs a number of input variables. For instance, the number of TSP agents involved and their preferences on the train schedules are required. Since there is a rich combination of these input variables, a case-based approach is inappropriate because conclusions drawn from the results are only valid to the specific set of input variables, which may hardly be representative in practice. In order to obtain generalised findings associated with the agent activities, a statistical analysis is more appropriate.

In a statistical analysis based on simulation, the set of input variables $\Theta_Y = \{y_i \mid i = 1, 2, ..., v\}$ are modelled by a set of known probability functions $P_i : y_i \to [0, 1]$. A random instance $\hat{y}_i$ is generated for each variable and they are delivered to a simulator, which produces a set of output instances $\hat{x}_i$ for the variable set $\Theta_X = \{x_i \mid i = 1, 2, ..., u\}$. If the simulation is repeated for m times, it is possible to construct the sample distribution X_i and compute the sample mean $\bar{x}_i$ for each output variable. Although the population distributions are unknown, the distributions of their sample means $\overline{X}_i$ will be (approximately) normal if the sample size m is sufficiently large (Walpole *et al.*, 1998; Ayyub & McCuen, 2003). As a result, by selecting an appropriate test-statistics (e.g. z-test or t-test statistics) to analyse the output data, the population means can be estimated. This process is summarised in Fig. 6.5.

Throughout the following study, it is assumed that three types of TSP agents are in competition of track capacity. I-TSP, R-TSP and F-TSP stand for intercity-, regional-

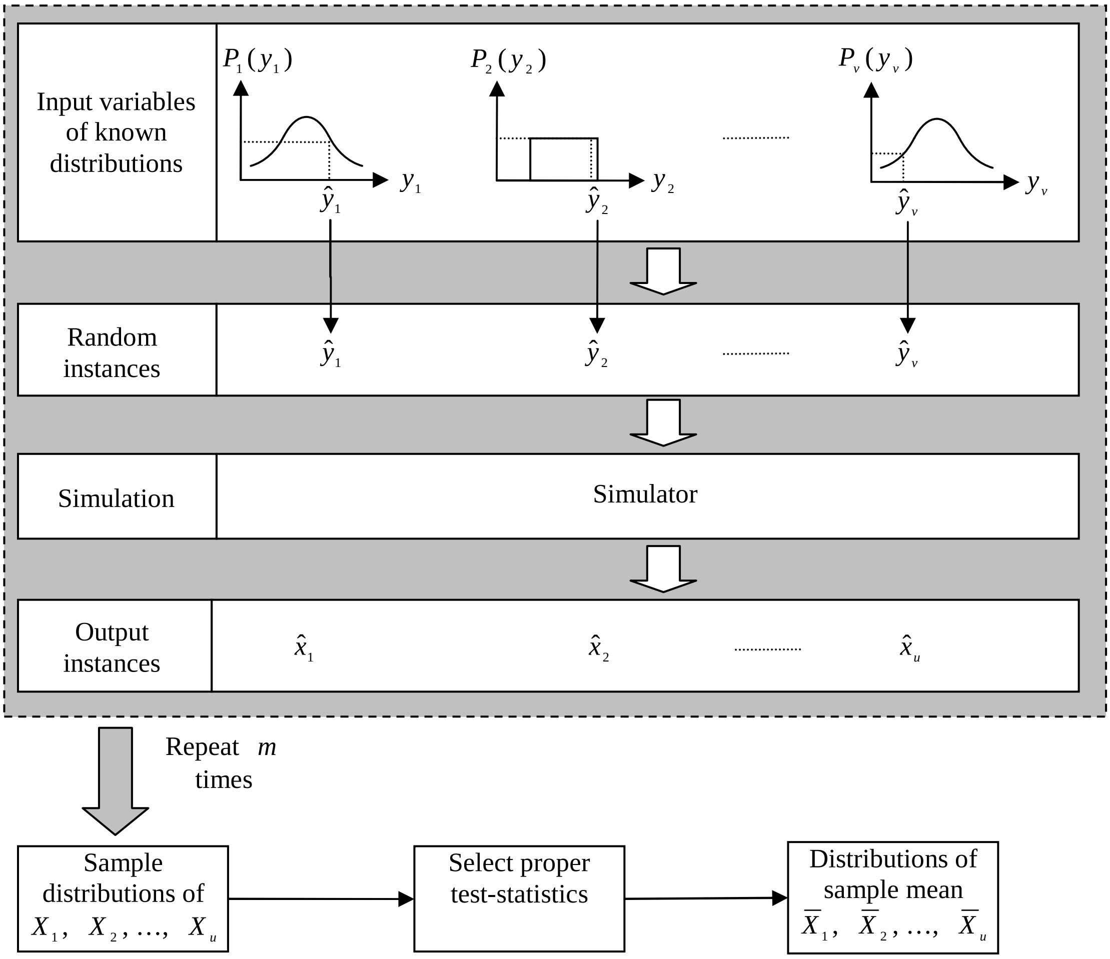


Fig. 6.5. A statistical analysis based on simulation

and freight- train service providers respectively. All these agents are included in the set $\Theta = \{\theta \mid \theta = I, R, or\ F\}$. In one simulation, there are n_I I-TSP agents, n_R R-TSP agents and n_F F-TSP agents, which gives rise to $N_{bid} = n_I + n_R + n_F$ number of IP-TSP transactions. In other words, a TSP agent is responsible for conducting one transaction, which involves the operation of one train service. n_I, n_R and n_F are random variables with probability density functions of $P_I(n_I)$, $P_R(n_R)$ and $P_F(n_F)$. In addition, for each type of TSP agent, the settings of track access charge c, service commencement time ζ, dwell time t_{Di} at station i, and inter-station runtime t_{Rj} between station j and $j+1$ are distributed by their respective probability density functions of $P_\theta(c)$, $P_\theta(\zeta)$, $P_\theta(t_{Di})$ and $P_\theta(t_{Rj})$.

Fig. 6.6 illustrates the statistical approach to derive the output estimates in the IP-TSP^n transaction. Given the probability density functions of the input variables, m random samples are generated. The samples are then fed into the open market simulator. The mean and standard deviations for IPU_T, IPU_A, DFR_I, DFR_R, EJT_I, EJT_R and EJT_F are then computed. To compare FCFS and HW2PF against these output variables, a set of two-sample hypothesis tests on the mean obtained from the policies are performed. The hypothesis tests are based on the t-test statistics because the population variances are unknown (Walpole *et al.*, 1998; Ayyub & McCuen, 2003).

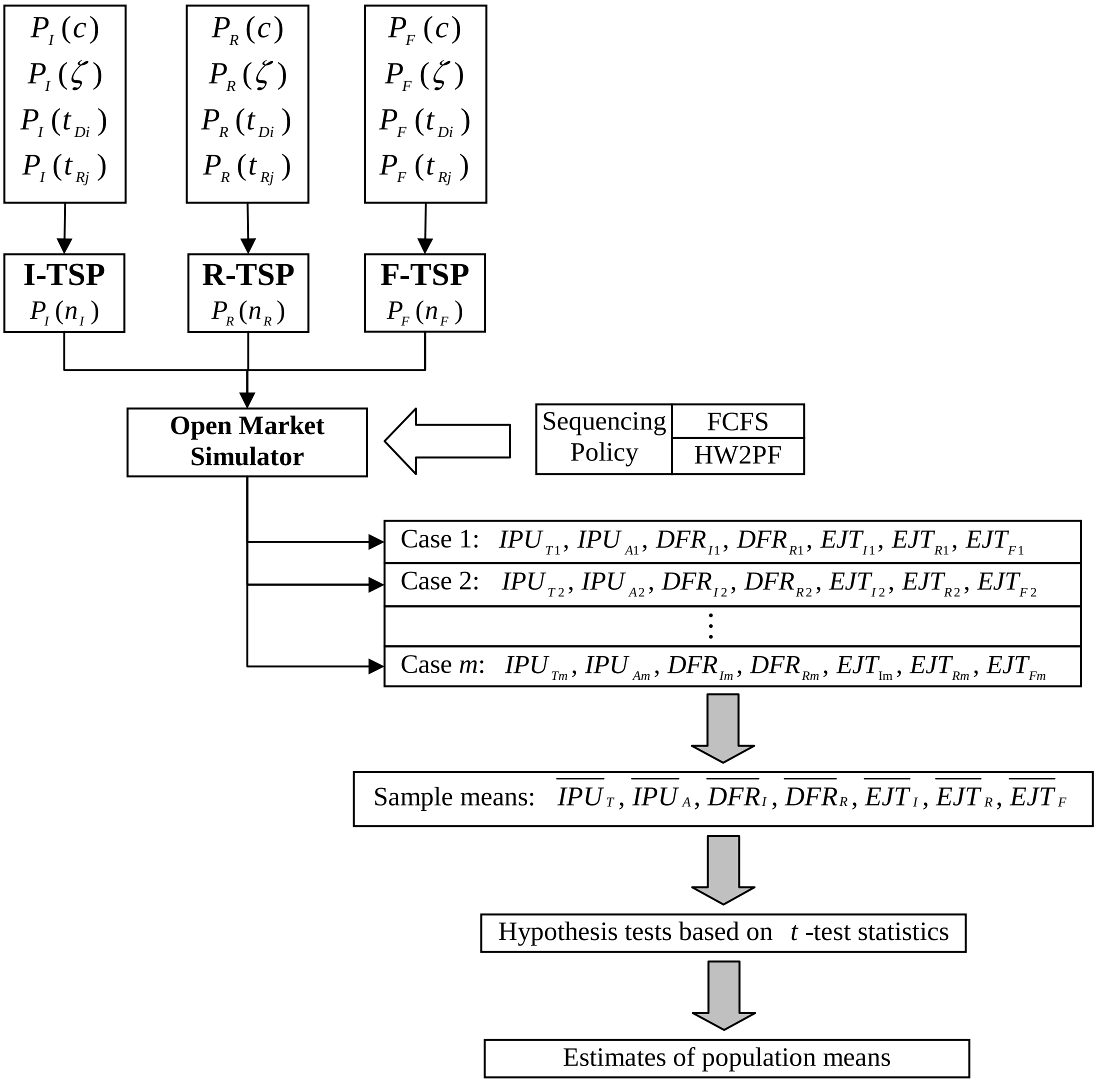


Fig. 6.6. Generation of statistical estimates using simulation

6.3.1. Sampling Distributions

When a random variable X is normally distributed with a population mean μ and variance σ^2, the sample mean $\overline{X}$ will be normally distributed with mean $\mu_{\overline{X}}$ and $\sigma^2_{\overline{X}}$. If m is the sample size, then $\mu_{\overline{X}} = \mu$ and $\sigma_{\overline{X}} = \sigma / \sqrt{m}$ (Fig. 6.7).

In practice, the sample mean $\overline{x}$ is used to estimate the population mean μ. However, $\overline{x}$ may have a deviation of δ due to the effect of random sampling (Fig. 6.7). It is often necessary to determine a confidence interval in which there is a probability of $1-\alpha$ that the population mean is resided in the interval ($\alpha \in [0,1]$ is known as the level of significance). With this definition, the interval is said to have a capture rate of $(1-\alpha)100\%$ of the population mean. If the population variance is known, the interval is computed using the z-statistics shown in (6.5), where $z_{\alpha/2}$ is the z-score corresponding to the level of significance α (Walpole *et al.*, 1998; Ayyub

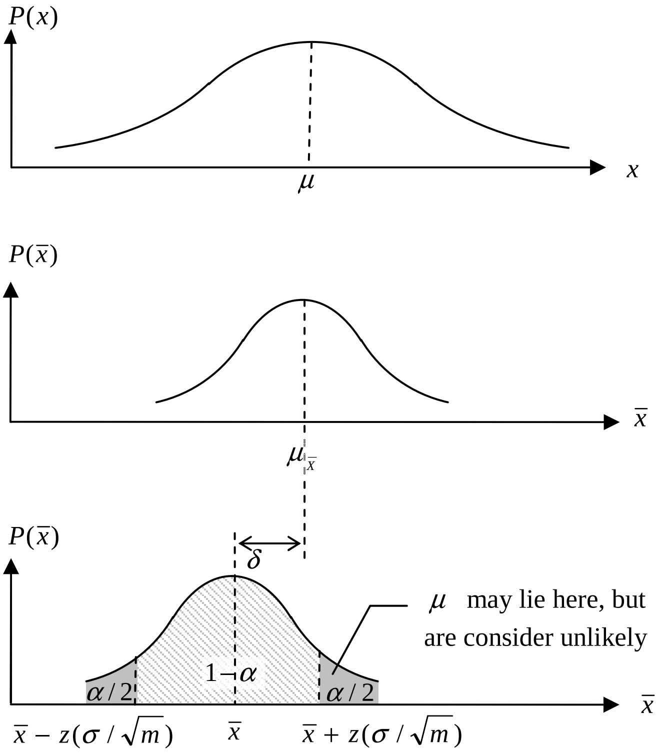


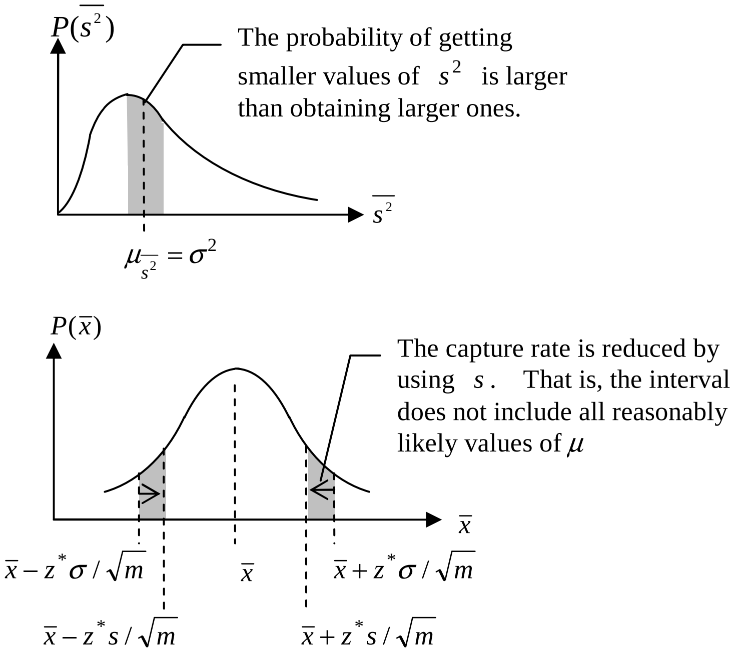
Fig. 6.7. Distributions of population and sample means

& McCuen, 2003).

$$\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{m}} \le \mu \le \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{m}} \tag{6.5}$$

Unfortunately, the population variance σ^2 is usually an unknown. In such case, the sample variance s^2 is used as an estimator. Since s^2 taken from a normal distribution follows a χ^2-distribution, it is more often to obtain s^2 whose value is smaller than σ^2 (Fig. 6.8). When the estimator is used directly, this causes $z_{\alpha/2}(s/\sqrt{m})$ to be smaller than $z_{\alpha/2}(\sigma/\sqrt{m})$ which lowers the capture rate. To resolve the problem, the use of the t-test statistic is needed to counter the effect. The resulting confidence interval is computed by (6.6), where $t_{\alpha/2,v}$ is the t-score corresponding to the level of significance α, and $v = m-1$ is the degree of freedom.

$$\bar{x} - t_{\alpha/2,v}\frac{\sigma}{\sqrt{m}} \le \mu \le \bar{x} + t_{\alpha/2,v}\frac{\sigma}{\sqrt{m}} \tag{6.6}$$



The probability of getting smaller values of s^2 is larger than obtaining larger ones.

The capture rate is reduced by using s. That is, the interval does not include all reasonably likely values of μ

Fig. 6.8. Distribution of sample variance and its effect on capture rate

For distributions that are not normally distributed, the t-test statistics may also be applied so long as the following assumption is satisfied. According to the Central Limit Theorem, if the sample size m is sufficiently large, then the distribution of the sample mean will be approximately normal, regardless of the distribution of the original population. By sufficiently large, it means the sample size approaches infinity. In practice, the condition is assumed to be satisfied when m is greater than 40 (Watkins *et al.*, 2004). Nevertheless, a more cautious approach to ensure the proper use of the statistics requires a transformation of the random variable beforehand. For instance, skewed distributions can be converted to approximate normal distributions by the natural logarithm or reciprocal transformations (Watkins *et al.*, 2004).

6.3.2. Two-sample Hypothesis Testing

When FCFS and HW2PF are applied in the simulation, it is essential to determine which one is (generally) better than the other. Although this may be examined by comparing the sample mean, the values of the mean are likely to be different due to the effect of random sampling. To examine whether the sample data suggests a significant difference between the two mean values, a set of two-sample hypothesis tests (Walpole *et al.*, 1998) is conducted.

In general, a null hypothesis H_0 is conjectured to describe a condition under test. The rejection of H_0 leads to the acceptance of the alternative hypothesis H_A. By convention, H_0 is specified by an equality while H_A is represented by an inequality.

To determine if the mean of sample S_1 of size m_1 is significantly smaller than the mean of sample S_2 of size m_2, the null and alternative hypotheses are set up as $H_0 : \mu_1 = \mu_2$ and $H_A : \mu_1 < \mu_2$ respectively. The t-score obtained from the two samples with unknown population variances (Watkins *et al.*, 2004) is computed by (6.7) and the approximated degree of freedom is calculated by (6.8).

$$t' = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{(s_1^2/m_1)^2 + (s_2^2/m_2)^2}} \tag{6.7}$$

$$v = \frac{(s_1^2/m_1 + s_2^2/m_2)^2}{\dfrac{(s_1^2/m_1)^2}{m_1 - 1} + \dfrac{(s_2^2/m_2)^2}{m_2 - 1}} \tag{6.8}$$

Suppose the level of significance required is α, the critical t-score can be looked up from the t-table as $t_{\alpha,v}$ (Table B.1 in Appendix B). If $t' > t_{\alpha,v}$, then the null hypothesis ($H_0 : \mu_1 = \mu_2$) is accepted, inferring that the data does not have adequate evidence to support a difference between the samples. Otherwise, the alternative hypothesis is accepted, indicating that the two mean values are indeed different.

6.3.3. Estimation of Sample Size

There are two types of errors in a hypothesis test. Type I error occurs when the alternative hypothesis is accepted when in fact the null hypothesis is true. On the other hand, type II error refers to the decision of accepting the null hypothesis when in fact the alternative hypothesis is true.

The probability of committing a type I error can be reduced by decreasing the level of significance α. However, this will also increase the probability of committing a type II error, denoted by β. In order to reduce α and β simultaneously, a larger sample size m is needed. The relationships of α, β and m are relatively simple for hypothesis tests based on z-statistics, but it becomes more complicated for tests based on t-statistics (Watkins *et al.*, 2004). For a two-sample hypothesis test of unknown variance but assumed equal, the sample size is determined by setting the desired values of α, β and a parameter $\Delta = |\mu_1 - \mu_2| / \sigma$, then m can be looked up from Table B.2 in Appendix B.

For the hypothesis tests conducted in this study, the variance is unknown and not likely to be equal. However, the sample size is still estimated by the above method for simplicity. α and β are set to 0.05 and $|\mu_1 - \mu_2| = 0.5\sigma$ or $\Delta = 0.5$. This yields a sample size of $m = 88$. Owing to the definitions of the output variables in (6.3) and (6.4), some of the simulations will produce infeasible values of DFR and EJT. In order to compensate for the possible reduction of sample size, the total number of simulation is set at 155 by assuming a rate of about 50% for generating infeasible values of DFR and EJT (Refer to Tables 6.9 – 6.13 for the actual sample sizes for different types of train services).

6.4. Simulation Setup

The case studies set up below aim to examine the performance of FCFS and HW2PF on the revenue collection of the IP agent and the quality of train services when three types of TSP agents are competing for track capacity. By adjusting the number of participating TSP agents, the case studies also simulate the scheduling of train services on a railway (single) track under light and heavy traffic conditions.

The track considered in the simulation consists of 5 stations (A to E). The entire track length is 85 km and the inter-station track lengths are shown in Table 6.1. In all simulation case studies, the IP agent issues a Request-For-Bid (RFB) message to the

Table 6.1. Track Configuration

Origin	Destination	Track Length (km)
A	B	20
B	C	30
C	D	15
D	E	20

TSP agents. The CTP spans from 07:00 to 07:59. Interested TSP agents are allowed to submit their bids before the submission deadline of 30s after the first issue of RFB.

Five cases have been constructed and the settings of their probability density functions are shown in Tables 6.2 to 6.6. $U(a_1, a_2, ..., a_n)$ denotes a uniform

Table 6.2. Case 1: Light Traffic Condition

	Intercity	Regional	Freight
Number of service	*U*(1)	*U*(3)	*U*(1)
TAC	*N* (1600, 625)	*N*(1500, 625)	*N*(1375, 100)
Commencement time	*U*(0:59)	*U*(0:19)	*U*(0:59)
Dwell Time at A	*N*(5, 0.25)	*P*(1, 0.2, 1)	*N*(15, 1)
Dwell Time at B	-	*P*(1, 0.2, 1)	*N*(15, 1)
Dwell Time at C	-	*P*(1, 0.2, 1)	-
Dwell Time at D	-	*P*(1, 0.2, 1)	*N*(15, 1)
Dwell Time at E	*N*(5, 0.25)	*P*(1, 0.2, 1)	*N*(15, 1)
Runtime at AB	*P*(11, 0.3, 1)	*P*(15, 0.5, 1)	*P*(24, 0.7, 1)
Runtime at BC	*P*(16, 0.3, 1)	*P*(24, 0.5, 1)	*P*(35, 0.7, 1)
Runtime at CD	*P*(9, 0.3, 1)	*P*(14, 0.5, 1)	*P*(23, 0.7, 1)
Runtime at DE	*P*(11, 0.3, 1)	*P*(15, 0.5, 1)	*P*(24, 0.7, 1)

Table 6.3. Case 2: Heavy Traffic Condition

	Intercity	Regional	Freight
Number of service	*U*(2)	*U*(6)	*U*(1)
TAC	*N* (1600, 625)	*N*(1500, 625)	*N*(1375, 100)
Commencement time	*U*(0:29)	*U*(0:9)	*U*(0:59)
Dwell Time at A	*N*(5, 0.25)	*P*(1, 0.2, 1)	*N*(15, 1)
Dwell Time at B	-	*P*(1, 0.2, 1)	*N*(15, 1)
Dwell Time at C	-	*P*(1, 0.2, 1)	-
Dwell Time at D	-	*P*(1, 0.2, 1)	*N*(15, 1)
Dwell Time at E	*N*(5, 0.25)	*P*(1, 0.2, 1)	*N*(15, 1)
Runtime at AB	*P*(11, 0.3, 1)	*P*(15, 0.5, 1)	*P*(24, 0.7, 1)
Runtime at BC	*P*(16, 0.3, 1)	*P*(24, 0.5, 1)	*P*(35, 0.7, 1)
Runtime at CD	*P*(9, 0.3, 1)	*P*(14, 0.5, 1)	*P*(23, 0.7, 1)
Runtime at DE	*P*(11, 0.3, 1)	*P*(15, 0.5, 1)	*P*(24, 0.7, 1)

Table 6.4. Case 3: Random Traffic Condition

	Intercity	Regional	Freight
Number of service	*U*(1, 2)	*U*(2, 3, 4, 6)	*U*(1)
TAC	*N* (1600, 625)	*N*(1500, 625)	*N*(1375, 100)
Commencement time	*U*(0:29), *U*(0:59)	*U*(0:9), *U*(0:14) *U*(0:19), *U*(0:29)	*U*(0:59)
Dwell Time at A	*N*(5, 0.25)	*P*(1, 0.2, 1)	*N*(15, 1)
Dwell Time at B	-	*P*(1, 0.2, 1)	*N*(15, 1)
Dwell Time at C	-	*P*(1, 0.2, 1)	-
Dwell Time at D	-	*P*(1, 0.2, 1)	*N*(15, 1)
Dwell Time at E	*N*(5, 0.25)	*P*(1, 0.2, 1)	*N*(15, 1)
Runtime at AB	*P*(11, 0.3, 1)	*P*(15, 0.5, 1)	*P*(24, 0.7, 1)
Runtime at BC	*P*(16, 0.3, 1)	*P*(24, 0.5, 1)	*P*(35, 0.7, 1)
Runtime at CD	*P*(9, 0.3, 1)	*P*(14, 0.5, 1)	*P*(23, 0.7, 1)
Runtime at DE	*P*(11, 0.3, 1)	*P*(15, 0.5, 1)	*P*(24, 0.7, 1)

Table 6.5. Case 4: Higher Willingness-to-Pay in Light Traffic Condition

	Intercity	Regional	Freight
Number of service	*U*(1)	*U*(3)	*U*(1)
TAC	*N* (2000, 625)	*N*(1500, 625)	*N*(1375, 100)
Commencement time	*U*(0:59)	*U*(0:19)	*U*(0:59)
Dwell Time at A	*N*(5, 0.25)	*P*(1, 0.2, 1)	*N*(15, 1)
Dwell Time at B	-	*P*(1, 0.2, 1)	*N*(15, 1)
Dwell Time at C	-	*P*(1, 0.2, 1)	-
Dwell Time at D	-	*P*(1, 0.2, 1)	*N*(15, 1)
Dwell Time at E	*N*(5, 0.25)	*P*(1, 0.2, 1)	*N*(15, 1)
Runtime at AB	*P*(11, 0.3, 1)	*P*(15, 0.5, 1)	*P*(24, 0.7, 1)
Runtime at BC	*P*(16, 0.3, 1)	*P*(24, 0.5, 1)	*P*(35, 0.7, 1)
Runtime at CD	*P*(9, 0.3, 1)	*P*(14, 0.5, 1)	*P*(23, 0.7, 1)
Runtime at DE	*P*(11, 0.3, 1)	*P*(15, 0.5, 1)	*P*(24, 0.7, 1)

Table 6.6. Case 5: Higher Willingness-to-Pay in Heavy Traffic Condition

	Intercity	Regional	Freight
Number of service	*U*(2)	*U*(6)	*U*(1)
TAC	*N* (2000, 625)	*N*(1500, 625)	*N*(1375, 100)
Commencement time	*U*(0:29)	*U*(0:9)	*U*(0:59)
Dwell Time at A	*N*(5, 0.25)	*P*(1, 0.2, 1)	*N*(15, 1)
Dwell Time at B	-	*P*(1, 0.2, 1)	*N*(15, 1)
Dwell Time at C	-	*P*(1, 0.2, 1)	-
Dwell Time at D	-	*P*(1, 0.2, 1)	*N*(15, 1)
Dwell Time at E	*N*(5, 0.25)	*P*(1, 0.2, 1)	*N*(15, 1)
Runtime at AB	*P*(11, 0.3, 1)	*P*(15, 0.5, 1)	*P*(24, 0.7, 1)
Runtime at BC	*P*(16, 0.3, 1)	*P*(24, 0.5, 1)	*P*(35, 0.7, 1)
Runtime at CD	*P*(9, 0.3, 1)	*P*(14, 0.5, 1)	*P*(23, 0.7, 1)
Runtime at DE	*P*(11, 0.3, 1)	*P*(15, 0.5, 1)	*P*(24, 0.7, 1)

distribution among feasible discrete values of $a_1, a_2, ..., a_n$. $U(a_1 : a_n)$ specifies a similar distribution with values $a_1, a_1 + 1, ..., a_n$. $N(\mu, \sigma^2)$ denotes a normal distribution with population mean μ and variance σ^2. $P(a, \lambda, t)$ represents a right-shifted Poisson distribution by a units with decay constant λ and time interval t.

Case 1 simulates the situation where a total of 5 train services are intended to operate on the track. The ratio of intercity, regional and freight services is 1:3:1. With a regional service headway time of 20 mins, the situation constitutes to a light traffic condition. More rail services are incorporated in case 2 to simulate a heavy traffic condition, where the ratio becomes 2:3:1, and the intercity and regional service headways are 30 mins and 10 mins respectively. In case 3, the number of train services is randomly generated according to a set of uniformly distributed probability functions. This corresponds to the case where the number of train services to be provided is unknown. Then, in cases 4 and 5, the (mean) track access charge of the intercity train services are raised from 16% to 45% higher than the regional ones compared to cases 1 and 2.

6.5. Results and Findings

The statistical results for the five cases are shown in Tables 6.7 to 6.14. Table 6.7 summarises the number of successful transactions found in each case. The statistics for IPU_A and IPU_T are given in Table 6.8. Tables 6.9 and 6.10 contain the EJT and DFR results for intercity services. The result of EJT for freight services is then included in Table 6.11. Tables 6.12 and 6.13 show the EJT and DFR statistics for regional services. Finally, Table 6.14 summarises the hypothesis tests comparing the mean obtained in case 1 and case 4, in addition to the mean obtained in case 2 and case 5.

In the statistical analysis, a natural logarithm-transformation (ln-transformation) is performed on the output random variables prior to the calculation of confidence interval and the hypothesis testing. This is because the distributions of the output variables are found to be skewed. The ln-transformation converts the distributions to approximately normal and it ensures the operation of the t-tests is meaningful (Watkins *et al.*, 2004). Illustrations of the distributions in case 3 (using FCFS) before and after transformation are depicted in Figs. 6.9 and 6.10. The normality of the distributions can be verified by chi-square tests.

Table 6.7. Transaction Successful and Failure Rates

Case		Case 1			Case 2			Case 3			Case 4			Case 5		
Sequencing policy		FCFS	HW2PF	(out of)	FCFS	HW2PF	(out of)	FCFS	HW2PF	(out of)	FCFS	HW2PF	(out of)	FCFS	HW2PF	(out of)
Successful transactions	Intercity	142	155	(155)	278	310	(310)	213	234	(234)	144	155	(155)	284	310	(310)
	Freight	150	151	(155)	131	149	(155)	145	148	(155)	149	151	(155)	134	146	(155)
	Regional	461	460	(465)	919	833	(930)	572	538	(581)	460	454	(465)	899	834	(930)
	Total	753	766	(775)	1328	1292	(1395)	930	920	(970)	753	760	(775)	1345	1290	(1395)
Successful rate (%)		97.2	98.8	-	95.2	92.6	-	95.9	94.8	-	97.2	98.1	-	96.4	92.6	-
Failure rate (%)		2.8	1.2	-	4.8	7.4	-	4.1	5.2	-	2.8	1.9	-	3.6	7.5	-

Table 6.8. IP Utility

Case		Case 1		Case 2		Case 3		Case 4		Case 5	
Sequencing policy		FCFS	HW2PF	FCFS	HW2PF	FCFS	HW2PF	FCFS	HW2PF	FCFS	HW2PF
Sample size		155	155	155	155	155	155	155	155	155	155
IPU_A	Mean ($/service)	1397	1408	1400	1416	1407	1418	1399	1409	1408	1415
	Standard deviation ($/service)	17.6	11.6	13.0	15.8	22.4	20.8	17.7	12.6	14.4	14.9
	H_0	$\mu_{FCFS} = \mu_{HW2PF}$		$\mu_{FCFS} = \mu_{HW2PF}$		$\mu_{FCFS} = \mu_{HW2PF}$		$\mu_{FCFS} = \mu_{HW2PF}$		$\mu_{FCFS} = \mu_{HW2PF}$	
	H_A	$\mu_{FCFS} < \mu_{HW2PF}$		$\mu_{FCFS} < \mu_{HW2PF}$		$\mu_{FCFS} < \mu_{HW2PF}$		$\mu_{FCFS} < \mu_{HW2PF}$		$\mu_{FCFS} < \mu_{HW2PF}$	
	Approx. degree of freedom v	266		298		306		276		308	
	Critical t-score $t_{\alpha,v}$, $\alpha = 0.05$	-1.640		-1.640		-1.640		-1.640		-1.640	
	t-score t'	-6.298		-9.553		-4.611		-5.522		-4.076	
	Conclusion	Reject H_0		Reject H_0		Reject H_0		Reject H_0		Reject H_0	
IPU_T	Mean ($)	6791	6958	11,999	11,795	8403	8415	6800	6908	11,946	11,741
	Standard deviation ($)	528	335	810	1164	2100	1958	531	421	883	1150
	H_0	$\mu_{FCFS} = \mu_{HW2PF}$		$\mu_{FCFS} = \mu_{HW2PF}$		$\mu_{FCFS} = \mu_{HW2PF}$		$\mu_{FCFS} = \mu_{HW2PF}$		$\mu_{FCFS} = \mu_{HW2PF}$	
	H_A	$\mu_{FCFS} < \mu_{HW2PF}$		$\mu_{FCFS} > \mu_{HW2PF}$		$\mu_{FCFS} < \mu_{HW2PF}$		$\mu_{FCFS} < \mu_{HW2PF}$		$\mu_{FCFS} > \mu_{HW2PF}$	
	Approx. degree of freedom v	259		264		307		291		282	
	Critical t-score $t_{\alpha,v}$, $\alpha = 0.05$	-1.640		1.640		-1.640		-1.640		1.640	
	t-score t'	-3.284		1.964		-0.171		-1.956		1.890	
	Conclusion	Reject H_0		Reject H_0		Accept H_0		Reject H_0		Reject H_0	

Table 6.9. Intercity Service EJT

Case	Case 1		Case 2		Case 3		Case 4		Case 5	
Sequencing policy	FCFS	HW2PF	FCFS	HW2PF	FCFS	HW2PF	FCFS	HW2PF	FCFS	HW2PF
Samples with no extension (%)	49.7	100	9.7	100	30.3	100	45.2	100	18.7	100
Samples with no service (%)	8.4	0	1.3	0	3.2	0	7.1	0	1.3	0
Remaining samples: ‘extended’ service (%)	41.9	0	89.0	0	66.5	0	47.7	0	80.0	0
Sample size of ‘extended’ service	65	-	138	-	103	-	74	-	124	-
Mean (min/service)	13.6	-	14.5	-	12.8	-	16.2	-	15.1	-
Standard deviation (min/service)	7.8	-	8.3	-	7.5	-	7.1	-	7.3	-

Table 6.10. Intercity Service DFR

Case	Case 1		Case 2		Case 3		Case 4		Case 5	
Sequencing policy	FCFS	HW2PF	FCFS	HW2PF	FCFS	HW2PF	FCFS	HW2PF	FCFS	HW2PF
Samples with periodical service (%)	-	-	7.7	64.5	48.4	84.5	-	-	13.5	100
Samples with no service (%)	-	-	1.3	0	3.2	0	-	-	1.3	0
Remaining samples: ‘non-periodical’ (%)	-	-	91.0	35.5	48.4	15.5	-	-	85.2	0
Sample size of ‘non-periodical’ services	-	-	141	55	75	24	-	-	132	-
Mean (min/service)	-	-	54.7	10.0	57.1	10.0	-	-	56.4	-
Standard deviation (min/service)	-	-	51.5	0	53.9	0	-	-	49.1	-

Table 6.11. Freight Service EJT

Case	Case 1		Case 2		Case 3		Case 4		Case 5	
Sequencing policy	FCFS	HW2PF	FCFS	HW2PF	FCFS	HW2PF	FCFS	HW2PF	FCFS	HW2PF
Samples with no extension (%)	78.1	74.8	71.6	65.8	80.0	75.5	81.9	83.9	75.6	63.2
Samples with no service (%)	3.2	2.6	15.5	3.9	6.5	4.5	3.9	2.6	13.5	5.8
Remaining samples: ‘extended’ service (%)	18.7	22.6	12.9	30.3	13.5	20.0	14.2	13.5	11.0	31.0
Sample size of ‘extended’ service	29	35	20	47	21	31	22	21	17	48
Mean (min/service)	4.9	4.9	5.2	4.9	4.0	5.5	5.5	4.2	3.7	4.8
Standard deviation (min/service)	4.2	4.2	4.3	3.7	3.0	4.5	4.2	4.0	2.0	3.8
H_0	$\mu_{FCFS} = \mu_{HW2PF}$		$\mu_{FCFS} = \mu_{HW2PF}$		$\mu_{FCFS} = \mu_{HW2PF}$		$\mu_{FCFS} = \mu_{HW2PF}$		$\mu_{FCFS} = \mu_{HW2PF}$	
H_A	$\mu_{FCFS} < \mu_{HW2PF}$		$\mu_{FCFS} < \mu_{HW2PF}$		$\mu_{FCFS} < \mu_{HW2PF}$		$\mu_{FCFS} > \mu_{HW2PF}$		$\mu_{FCFS} < \mu_{HW2PF}$	
Approx. degree of freedom v	59		30		46		41		32	
Critical t-score $t_{\alpha,v}$, $\alpha = 0.05$	-1.670		-1.697		-1.680		1.682		-1.695	
t-score t'	-0.058		-0.273		-1.193		1.37		-0.895	
Conclusion	Accept H_0		Accept H_0		Accept H_0		Accept H_0		Accept H_0	

Table 6.12. Regional Service EJT

Case	Case 1		Case 2		Case 3		Case 4		Case 5	
Sequencing policy	FCFS	HW2PF	FCFS	HW2PF	FCFS	HW2PF	FCFS	HW2PF	FCFS	HW2PF
Samples with no extension (%)	37.4	14.2	14.8	0	37.4	14.2	41.3	15.5	12.3	2.6
Samples with no service (%)	0	0	0	0	0	0	0	0.0	0	0
Remaining samples: 'extended' service (%)	62.6	85.8	85.2	100	62.6	85.8	58.7	84.5	87.7	97.4
Sample size of 'extended' service	97	133	132	155	97	133	91	131	136	151
Mean (min/service)	6.9	7.5	5.2	8.6	6.7	9.7	7.0	8.3	5.4	9.1
Standard deviation (min/service)	6.4	5.8	4.4	5.7	5.2	8.0	6.3	6.6	5.0	6.4
H_0	$\mu_{FCFS} = \mu_{HW2PF}$		$\mu_{FCFS} = \mu_{HW2PF}$		$\mu_{FCFS} = \mu_{HW2PF}$		$\mu_{FCFS} = \mu_{HW2PF}$		$\mu_{FCFS} = \mu_{HW2PF}$	
H_A	$\mu_{FCFS} < \mu_{HW2PF}$		$\mu_{FCFS} < \mu_{HW2PF}$		$\mu_{FCFS} < \mu_{HW2PF}$		$\mu_{FCFS} < \mu_{HW2PF}$		$\mu_{FCFS} < \mu_{HW2PF}$	
Approx. degree of freedom v	182		236		200		182		257	
Critical t-score $t_{\alpha,v}$, $\alpha = 0.05$	-1.650		-1.650		-1.650		-1.650		-1.650	
t-score t'	-1.595		-6.539		-2.810		-1.835		-5.960	
Conclusion	Accept H_0		Reject H_0		Reject H_0		Reject H_0		Reject H_0	

Table 6.13. Regional Service DFR

Case	Case 1		Case 2		Case 3		Case 4		Case 5	
Sequencing policy	FCFS	HW2PF	FCFS	HW2PF	FCFS	HW2PF	FCFS	HW2PF	FCFS	HW2PF
Samples with periodical service (%)	26.4	3.9	11.0	0	29.0	3.2	32.9	4.5	7.1	0
Samples with no service (%)	0	0	0	0	0	0	0	0.0	0	0
Remaining samples: 'non-periodical' (%)	73.6	96.1	89.0	100	71.0	96.8	67.1	95.5	92.9	100
Sample size of 'non-periodical' services	114	149	138	155	110	150	104	148	144	155
Mean (min/service)	18.6	20.0	15.1	25.9	22.2	31.2	18.3	23.1	15.9	24.9
Standard deviation (min/service)	15.0	15.2	9.5	16.4	22.4	35.5	16.3	19.6	10.8	13.3
H_0	$\mu_{FCFS} = \mu_{HW2PF}$		$\mu_{FCFS} = \mu_{HW2PF}$		$\mu_{FCFS} = \mu_{HW2PF}$		$\mu_{FCFS} = \mu_{HW2PF}$		$\mu_{FCFS} = \mu_{HW2PF}$	
H_A	$\mu_{FCFS} < \mu_{HW2PF}$		$\mu_{FCFS} < \mu_{HW2PF}$		$\mu_{FCFS} < \mu_{HW2PF}$		$\mu_{FCFS} < \mu_{HW2PF}$		$\mu_{FCFS} < \mu_{HW2PF}$	
Approx. degree of freedom v	223		255		227		216		236	
Critical t-score $t_{\alpha,v}$, $\alpha = 0.05$	-1.650		-1.650		-1.650		-1.650		-1.650	
t-score t'	-1.423		-8.384		-2.819		-2.437		-7.565	
Conclusion	Accept H_0		Reject H_0		Reject H_0		Reject H_0		Reject H_0	

Table 6.14. Hypothesis Tests against Different Willingness-to-pay for Total IP Utility

Sequencing policy	FCFS				HW2PF			
Case	Case 1	Case 4	Case 2	Case 5	Case 1	Case 4	Case 2	Case 5
Sample size	155	155	155	155	155	155	155	155
Mean ($)	6791	6801	11,999	11,946	6958	6908	11,795	11,741
Standard dev. ($)	528	531	810	883	335	420	1164	1150
H_0	$\mu_1 = \mu_4$		$\mu_2 = \mu_5$		$\mu_1 = \mu_4$		$\mu_2 = \mu_5$	
H_A	$\mu_1 < \mu_4$		$\mu_2 > \mu_5$		$\mu_1 > \mu_4$		$\mu_2 > \mu_5$	
Approx. degree of freedom v	308		305		293		308	
Critical t-score $t_{\alpha,v}$, $\alpha = 0.05$	-1.650		1.650		1.650		1.650	
t-score t'	-0.140		0.588		1.160		0.363	
Conclusion	Accept H_0		Accept H_0		Accept H_0		Accept H_0	

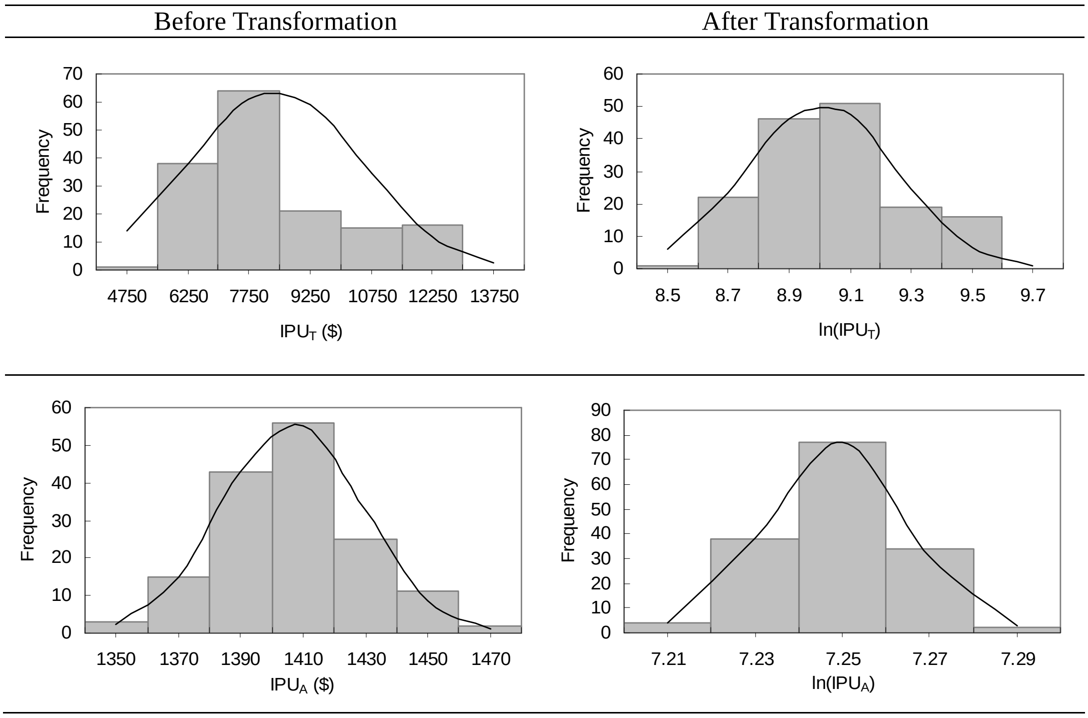


Fig. 6.9. Distributions of IPU before and after ln-transformation

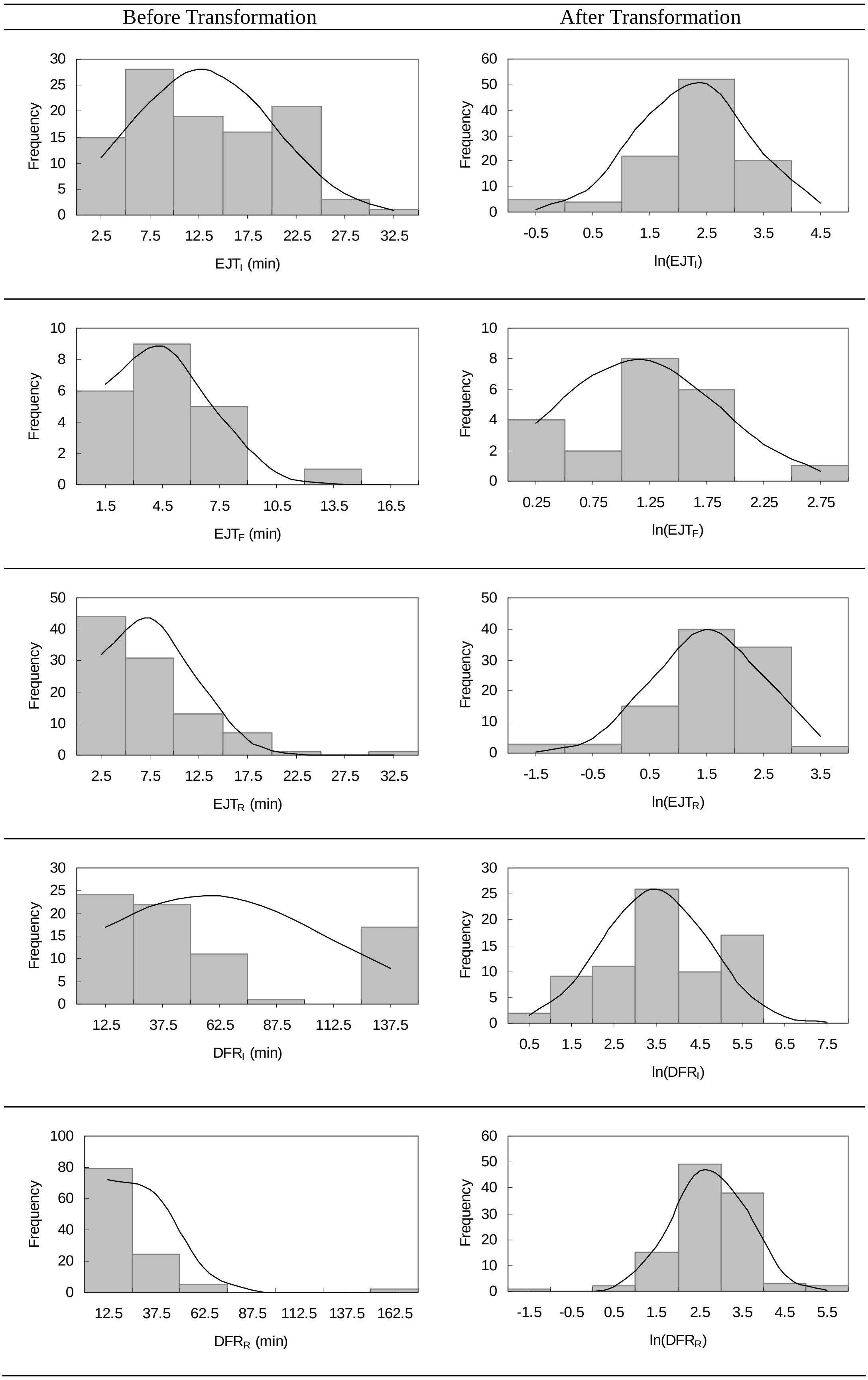


Fig. 6.10. Distributions of EJT and DFR before and after ln-transformation

6.5.1. Infrastructure Provider

6.5.1.1. Average IP Utility

In all cases, the null hypotheses regarding to IPU_A (i.e. whether the mean of the two sequencing policies are equal) are all rejected (Table 6.8). Thus, with the acceptance of the alternative hypotheses, the performance of HW2PF in terms of IPU_A is significantly better than that of FCFS.

In HW2PF, since track capacity is allocated to TSP agents with decreasing potential willingness-to-pay, capacity is first allocated to intercity services, followed by freight and then regional ones. Moreover, whenever a deal is secured, the scheduled service will impose more constraints on track capacity, so the subsequent transactions are more likely to lead to termination without a deal made. In other words, in this study, intercity and freight services stand more chance to obtain track access rights than their regional counterparts do. This is reflected in Table 6.7 by the increased number of successful I-TSP and F-TSP transactions, and the reduced number of successful R-TSP transactions when HW2PF is employed. In essence, the IP agent is attempting to replace the ‘low-valued’ regional transactions with the ‘high-valued’ intercity and freight transactions so as to increase IPU_A.

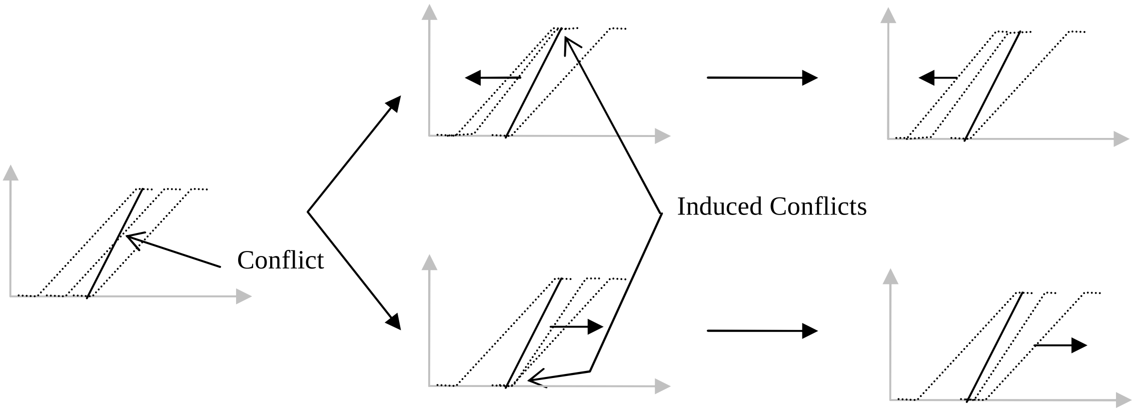
The fact that the IP agent favours the TSP agents based on solely the higher willingness-to-pay raises the concern on equity of track access. Since equity of infrastructure access is an important issue in transportation (BTRE, 2003; Zhang & Levinson, 2005) and it often raises social and political debates, most regulatory bodies impose restrictions on the IP to guarantee a minimal level of track access to the TSPs. Although the model assumes that the IP is not regulated on the equity of access, the inclusion of the parameter in further work will be beneficial to the regulatory bodies to investigate the trade-off between freedom and equity of access.

6.5.1.2. Total IP Utility

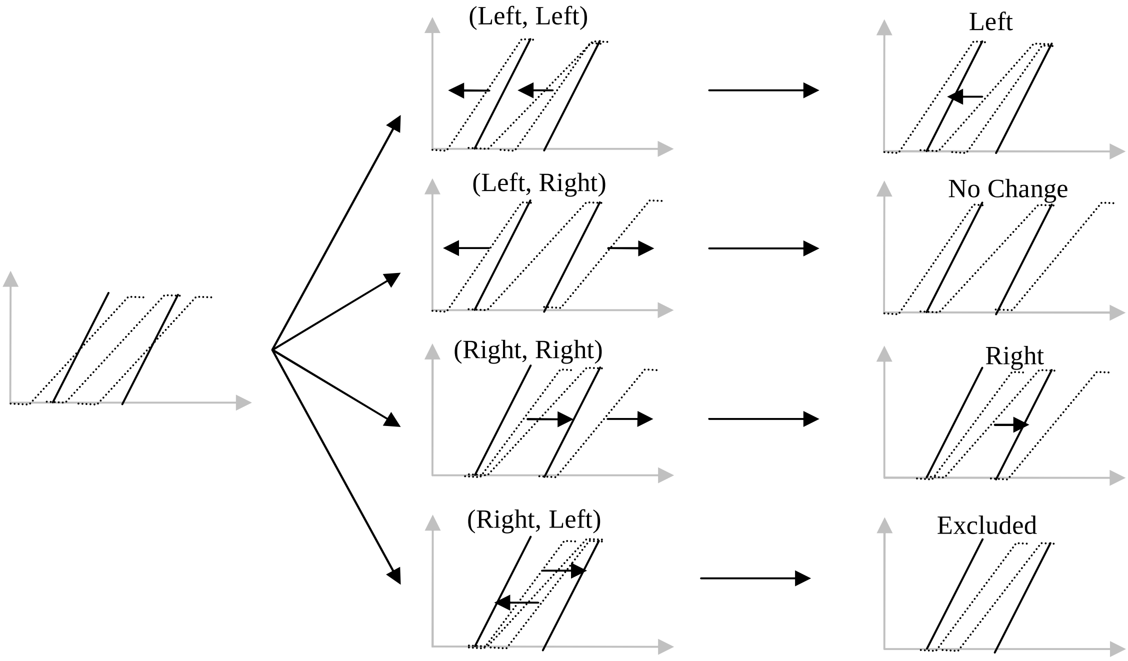
The result of IPU_A does not lead to the conclusion that HW2PF always improves the revenue collection of the IP. In fact, according to the results of IPU_T in Table 6.8, HW2PF is only better than FCFS under light traffic conditions (cases 1 and 4). When traffic becomes congested (cases 2 and 5), FCFS performs better than HW2PF.

The observation that HW2PF raises IPU_T under light traffic condition is the result of a larger improvement in the number of successful I-TSP transactions than the corresponding change for freight and regional services (Table 6.7). For instance, in case 1, there is an increase of 13 successful transactions for the intercity services when employing HW2PF, while the number of freight and regional services are changed only by a single transaction. In other words, there has been a net increase in the number of successful transactions conducted. As mentioned above, since capacity is first allocated to the intercity services in HW2PF, when competition for track capacity is not severe under light traffic condition, there is sufficient capacity to schedule the remaining rail services.

In contrast, even though there is still a reasonable improvement of intercity transactions under heavy traffic, there is a substantial reduction of regional transactions (Table 6.7). In fact, the total number of failure transactions in HW2PF outnumbers that of FCFS, leading to the weaker performance observed. The cause of the failure negotiations with R-TSP under heavy traffic condition can be realised in Fig. 6.11, which illustrates the consequence of ‘knock-on’ effect when intercity services (solid lines) are scheduled first. Fig. 6.11a shows a scenario in light traffic condition when a regional service (the dotted line in the middle) is in conflict of rights-of-way with an



(a) Possible consequences of induced conflict in light traffic



(b) Possible consequences of induced conflicts in heavy traffic

Fig. 6.11. Illustration of 'knock-on' effect

intercity train. From the figure, there are two possible methods to resolve the conflict – shifting the regional schedule to either left or right. In both cases, the shifting causes a new conflict with the adjacent services which leads to subsequent adjustment of schedules. Under this situation, all transactions are secured albeit the regularity of the regional services is affected. On the other hand, when there are two conflicts occurring in close proximity as shown in Fig. 6.11b, which is quite common under heavy traffic condition, there are four possible outcomes – the combinations of shifts are (left, left), (left, right), (right, right) and (right, left). In the first three situations, the sandwiched regional services may be adjusted correspondingly to resolve the induced conflicts. However, in the last scenario, when the middle service experiences two induced conflicts simultaneously, there is inadequate capacity available for adjustment. Such possibility highly increases the failure rate for regional trains under heavy traffic condition.

With a random sequence of incoming TSPs, FCFS may allow capacity allocation to a regional service first. This raises the transaction failure rate for I-TSP, but lowers the one for R-TSP. FCFS is therefore able to secure more ‘low-valued’ regional transactions than the ‘high-valued’ ones. Under heavy condition, this is an advantage since the increase in demand for regional services is greater than that for intercity services. The reduction of the number of intercity transactions is thus justified by the increase in the number of regional transactions.

6.5.2. Intercity Services

6.5.2.1. Extension in Journey Time

Simulation results for EJT_I are shown in Table 6.9. When employing HW2PF, all intercity services obtain the desired schedules, but when FCFS is used, there are usually over 50% of cases suffering from extension in journey time.

The exceptional performance by HW2PF is once again resulted from the transaction sequence of decreasing potential willingness-to-pay. The first allocation of intercity services on track implies the minimal capacity constraints when they are scheduled. Having more freedom to select their running profiles, these services are almost certain to obtain their desired schedules. On the other hand, in FCFS, it is possible that regional or freight services are scheduled prior to the intercity services. This may lead to severe extensions as illustrated in Fig 6.12. In this figure, the regional service has already been scheduled when the IP agent is attempting to allocate capacity to an intercity service with its desired profile shown. The crossing of the two profiles near 35 km gives rise to a conflict of rights-of-way. Although the conflict is finally resolved by negotiation, the journey of the intercity service has been extended substantially.

Moreover, it appears that FCFS performs better under light traffic condition. This is observed from the higher proportion of cases obtaining zero extension in cases 1 and 4 (49.7% and 45.2%) than the corresponding values in cases 2 and 5 (9.7% and 18.7%). Therefore, under light traffic condition, FCFS is more likely to obtain the desired intercity journey time, even though the overall performance in EJT_I is not comparable to that of HW2PF. The observation is in fact governed by the probability of selecting an intercity service for negotiation, which is determined from the ratio of the number of intercity, freight and regional services. Since there is only one intercity service competing with four other services under light traffic condition, the probability that the

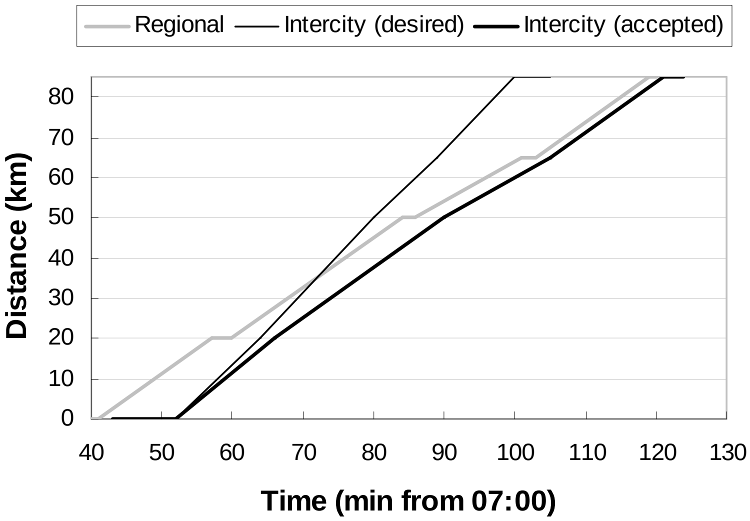


Fig. 6.12. Illustration of severe extension of an intercity service

intercity service is selected first for scheduling is 1/5 or 0.2. Under heavy traffic condition, there are 2 intercity services competing with 7 other types of services. So the probability that both services are selected first is substantially lowered to (2/9 × 1/8) = 0.028.

Among the extended services, EJT_I obtains extended journey times from about 13 to 15 mins (Table 6.9). Since the train only stops at station E, these figures also represent the average extension time experienced by the passengers on the train. As the desired journey time of an intercity service is approximately 60 mins, the travelling time has thus been increased by almost 25%. Such extension is generally not acceptable from the commuters' viewpoints.

6.5.2.2. Deviation from Regularity

HW2PF obtains a high percentage of periodic schedules. It achieves 64.5% in case 2 and even reaches 100% in case 5. However, the performance of FCFS is not comparable to HW2PF. It has only around 10% of cases obtaining periodic schedules. The majority is non-periodical with mean DFR_I of 55 mins. Such a large deviation is associated with the limited number of intercity services. As there are at most 2 intercity services under heavy traffic condition, if either one of them cannot be scheduled, the loss in regularity is given by (6.4):

$$\begin{aligned}
DFR_I &= \frac{1}{n_I}\sum_{j=1}^{n_s}\sum_{i=1}^{n_I}\left|(t_{i+1}^j - t_i^j) - 60/\hat{n}_I\right| \\
&= \frac{1}{1}\sum_{j=1}^{2}\sum_{i=1}^{1}\left|(t_{i+1}^j - t_i^j) - 60/2\right| \\
&= \left|(t_2^1 - t_1^1) - 30\right| + \left|(t_2^2 - t_1^2) - 30\right| \\
&= \left|(t_1^1 + 60 - t_1^1) - 30\right| + \left|(t_1^2 + 60 - t_1^2) - 30\right| \\
&= 60\,\text{min}
\end{aligned}$$

As the mean DFR_I obtained is close to 60 mins, it indicates that among the non-periodic cases in FCFS, a large proportion is caused by the loss of one intercity service. This is not surprising since capacity is not necessarily allocated to I-TSP first. If one of the intercity transactions is performed towards the end of the queue, then its service is unlikely to operate on the track, especially when the traffic is congested.

6.5.3. Freight Services

Results obtained for freight services using FCFS and HW2PF are similar (Table 6.11), regardless to the traffic conditions. In general, the majority (about 60-85%) obtains the desired capacity, and a small proportion of freight services (2.6-15.5%) are unable to be allocated on track. In addition, there are about 10-20% of cases suffering from extension in journey time. The mean of EJT_F ranges from 3.7 to 5.5 mins, and all hypothesis tests indicate there is no evidence for a significant different between the mean of the two sequencing policies.

It may be surprising to discover that F-TSP has such a high likelihood in obtaining their desired track capacity. In fact, the result is found to be biased because the freight trains under simulation rarely compete for capacity with other train services. A typical example is illustrated in Fig. 6.13. Initially, the IP agent requests bids from TSP agents operating trains that commence within the first 1-hour period (from 0-59 mins). One intercity, one freight and three regional services are entitled for submission. Since the journey time of the regional and intercity services only spans for about 60 to 70 mins, their trains will usually arrive at the destination station within 130 mins. However, the travelling time of the freight service is nearly three times longer (due to the lower speed and longer loading/unloading time at station). As a result, if the commencement time of the freight service approaches to 60 mins, the freight train will hardly experience any

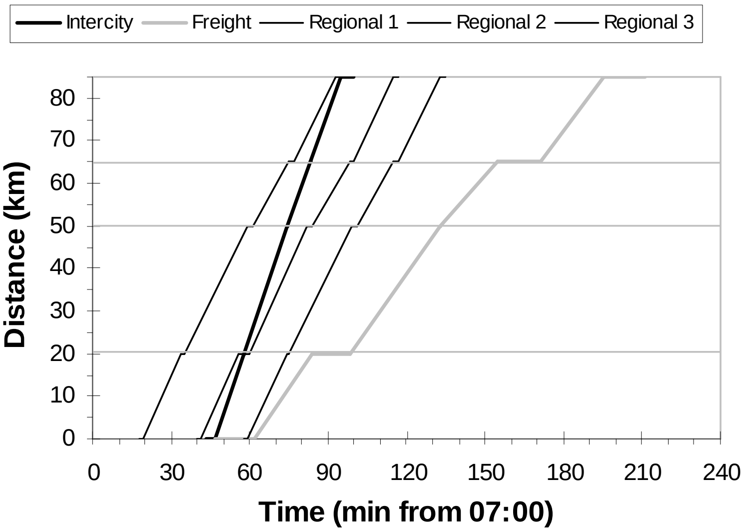


Fig. 6.13. Absence of competition for freight service

competition of track capacity with the intercity and regional services.

One implication from the above observation is that the CTP for the faster (intercity and regional) and the slower (freight) trains should be different in order to introduce the intended degree of competition in railway markets. The recommended timeframes are depicted in Fig. 6.14. If the total journey time for the faster and slower services are t_f and t_s, and the corresponding CTP are T_f and T_s, the approximation of $T_f + t_f \approx T_s + t_s$ should hold. As t_f and t_s are only determined during negotiations, the IP may assume their values from past experience. To avoid excessive simulation/processing time and to aid subsequent repetitive bid process, it is suggested

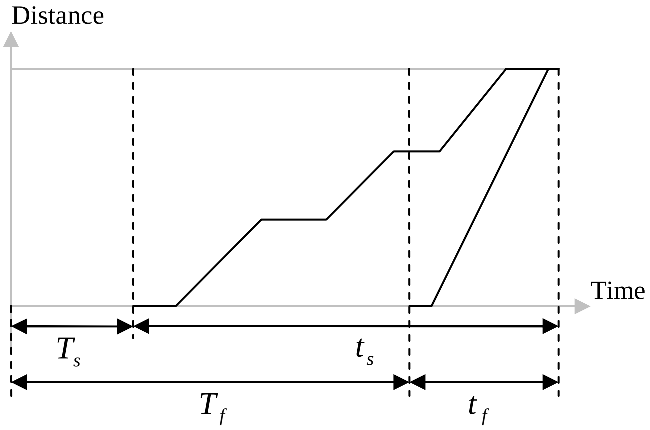


Fig. 6.14. Determination of commencement time periods for fast and slow trains

that $T_s = 30$ min and $T_f = T_s + t_s - t_f$ (rounding to the nearest 30 mins).

6.5.4. Regional Services

6.5.4.1. Extension in Journey Time

Similar to intercity services, there are more extended schedules under heavy traffic condition. Having more regional services, R-TSP is less likely to obtain the desired set of schedules, especially when their services are obscured by the competing intercity and freight services.

In addition to the effect of heavy traffic condition, the number of extended schedules is also increased by the use of HW2PF. From Table 6.12, FCFS obtains about 60% and 85% of extended schedules under light and heavy traffic conditions respectively, while the corresponding figures acquired by HW2PF increase to about 85% and 100%. Moreover, the hypothesis tests between FCFS and HW2PF for EJT_R indicate their differences are significant (except in case 1 where the conclusion is marginally accepted).

The increase in both the number and the mean of EJT_R in HW2PF is expected because of the lower willingness-to-pay by the R-TSP. Intercity services are therefore scheduled prior to the regional ones, which imposes more capacity constraints to the regional scheduling problems. Thus, the regional services are frequently subject to more adjustments in their journey time to avoid conflicts of rights-of-way with the intercity services.

The mean of EJT_R ranges from 5.2 to 9.7 mins. Since each regional service has four inter-station runs, the average extension time experienced per inter-station are about 1.3 to 2.5 mins. This seems reasonable to passengers who travels in short distances (e.g. 1 or 2 stations), but the cumulative extension for more stations may not be tolerable. In addition, the large standard deviations of 4 to 8 mins suggest that in some extreme cases, the extension can be excessively high.

6.5.4.2. Deviation from Regularity

According to the results in Table 6.13, the trends for regularity are identical to those found for EJT_R. Heavy traffic condition and the use of HW2PF yield a higher number of non-periodical services. Also, the hypothesis tests show that the mean DFR_R obtained by HW2PF is larger than that by FCFS. The explanations are the same as described for EJT_R.

Under light traffic condition, the mean DFR_R spans from 18 to 23 mins. Since it represents the cumulative sum of five stations, the value per station are thus between 3.6 to 4.6 mins. With the desired service headway of 20 mins, there is 20% deviation of regularity. In the case of heavy traffic condition, the range becomes wider, spanning from 15 to 26 mins per service, or about 3 to 5 mins per service per station. As the headway becomes 10 mins in tighter traffic condition, the deviations are increased to 30-50%.

6.5.5. Effect of Willingness-to-Pay

Table 6.14 summarises the results for the hypothesis tests comparing the mean of IPU_T when the willingness-to-pay of I-TSP has been increased from $1600 (in cases 1 and 2) to $2000 (in cases 4 and 5). In all cases, the null hypotheses are accepted which indicates that there is no evidence for significant difference between the mean.

According to the results, it appears that the IP agent is unable to take the advantage of the increased willingness-to-pay so that it may benefit from a higher revenue collection. The reason for such behaviour is explained by the maximum obtainable track access charge, or the ceiling price, which is derived from the sum of the maximum sub-charges for track usage (TUC), traction energy (TEC), peak demand (PDC) and congestion (CGC). These charges vary with the types of rolling stock in operation and the quality of the proposed train schedules. By inspecting the simulation data, these values are found to be \$536, \$416, \$184 and \$557 respectively. In other words, the ceiling price in this study is \$1693. Since the willingness-to-pay offered in cases 1 and 2 is already close to this value, it is reasonable that further increase of the willingness-to-pay in cases 4 and 5 should not lead to a significant increase in revenue collection.

In practice, ceiling price is usually regulated by the local juridical authority. It is used to prevent the IP from monopolising the infrastructure provision, especially in market sectors where railways are competitive to other modes of transportation (e.g. long-distance freight transportation) (BTRE, 2003). For example, the ceiling price set for the Australian Rail Track Corporation (an IP in Australia) is limited by the economic costs on infrastructure usage, return of asset and depreciation so that the access seekers are not exploited for overcharging by the IP (BTRE, 2003). In this study, although the ceiling price is not explicitly defined, because the four sub-charges are derived to recover the economic cost of infrastructure usage, the maximum sum of these charges will therefore yield the ceiling price.

6.6. Remarks

This chapter has investigated railway competition in open access markets on track capacity among several TSPs providing intercity, regional, and freight services. The study is enabled by the use of MBSBP as the agent interaction protocol, the development of FCFS and HW2PF as the sequencing policies, and the definitions of total IP utility, average IP utility, extension in journey time and deviation in regularity as the performance measuring indices.

In order to derive generalised conclusions on the performance of the sequencing policies, a statistical approach is employed. The adoption of the t-test statistics has successfully produced estimations of the mean of the four measuring indices. In addition, with the aid of the two-sample hypothesis tests, the performance between FCFS and HW2PF under light and heavy traffic has been compared.

It has been found that the use of HW2PF is more favourable to intercity services as capacity will be first allocated to these services because of their higher willingness-to-pay. This has resulted in almost no extensions of journey time and no deviations from regularity for intercity services where in the cases of using FCFS has resulted in severe deterioration from journey time and regularity. Unfortunately, when HW2PF is applied, the quality of regional services is slightly degraded for short-distance travelling and it becomes worse for longer distance travelling. In other words, there is a tug-of-war on the choice of sequencing policy concerning the quality of train services.

Nevertheless, if there are no regulations on equity of access and the IP only aims to increase its short-term revenue collection, the IP should employ HW2PF when intercity services are dominating the network traffic; and adopt FCFS when regional services are dominating. In the former, the IP may receive higher revenue by securing transactions of higher track access charge first so as to avoid large number of rejections of the favourable buyers. In the latter, FCFS is likely to increase the total number of train services (hence total revenue collection) albeit the possibility of lower premium in each transaction. However, as the issue on equity of access has become important in open access market, further works should investigate how the IP should be regulated so that a certain level of equity is introduced in the open market while impact on the commercial operation of the IP is minimised. One means to measure equity is the application of the Gini Coefficient (Zhang & Levinson, 2005), which can be incorporated in the objective function of the IP. The imposition of a minimum level of equity by the railway regulatory bodies may be represented by a constraint in the in the combinatorial optimisation problem.

It was also discovered that different Commencement Time Periods (CTPs) should be used for different types of train services, especially for freight and passenger services. When the same CTP is used by both categories of services, trains running at significantly lower speed will seldom be subject to competition from the faster trains. The CTP for the slower trains should then be reduced in order to ensure fair competition in open markets.

While this study has examined the performance of FCFS and HW2PF, other sequencing policies may be studied similarly by the statistical approach presented. For example, other possible sequencing policies using simple decision rules may involve ordering the IP-TSP transaction by the earliest-commencement-time-first, or the longest-dwell-time-first. On the other hand, more complex sequencing policies may be realised from the current practices adopted by experienced train planners. In such case, comprehensive interviews with the train planners are required before a realistic rule-based system can be devised.

In addition, the alternatives for implementing sequence generation (described at the beginning of in this chapter) can be developed and compared against the transaction-based generation approach. It is also interesting to set up an IP-TSP^n transaction based on the auctioning mechanism. As many countries have strong concern on developing an access charging regime that is suitable for their railway markets, the pursuit of the auction-based multi-agent system is definitely useful to evaluate the costs and benefits of implementing such charging regime. However, since it is difficult to explicitly define the ‘product’ in a railway auction (as different TSPs have different capacity requirements), considerable effort is needed to devise a proper definition of the auctioning product. With this change in definition and interaction between the parties, the agent interaction protocol and decision-making mechanisms have to be revised.

Chapter 7

Bilateral Negotiation for Schedule Coordination

In Chapters 5 and 6, the bilateral and multilateral negotiations between IP and TSP were modelled and studied. While these transactions represent the core interactions in open access markets, there are other types of negotiations in open markets that are also worth studying. This chapter focuses on the bilateral negotiation between two passenger-TSPs whose objective is to coordinate their train services at a common interchange station.

In modelling the track allocation processes in the IP-TSP^n transaction, train planning for different TSPs is assumed to be conducted independently. Although the assumption is valid for TSPs competing directly on a single track, when the TSPs are serving different routes, coordinated planning may be in fact beneficial to both the stakeholders and the passengers. As discussed in Chapter 2, the TSPs may increase revenue collection by boosting the passenger demand, while the travellers can also enjoy a shorter transfer time.

Although the idea of schedule coordination is consistent with the intention of the EU policies (or other national policies in various countries) that open markets should reduce barriers on providing seamless services, the TSPs, which belong to different parties, are still required to preserve a high degree of autonomy in their decision-making, and there are also regulations to prohibit collusions with monetary involvement between the stakeholders. The negotiation between two TSPs (the TSP-TSP transaction), yet without monetary involvement, is the objective of this study.

This chapter is organised as follows. Firstly, the schedule coordination problem is described. Secondly, a mathematical model is proposed for the decision-making process for a TSP agent capable of incorporating a negotiation strategy. Next, three negotiation strategies of different objectives are devised. Simulation case studies are then conducted to examine the actions of TSP agents employing different strategies. This is followed by the discussions on results and findings.

7.1. Problem Description

Despite the permission of competition among TSPs for track capacity and passengers in open railway access markets, direct competition through the provision of identical regional services seldom occurs in practice. In fact, the scope of operation of a regional TSP usually overlaps (or intersects) with another TSP, but they do not coincide completely. As a result, little or only a moderate level of competition is present in the overlapping scopes of service provision.

An example is shown in Fig. 7.1. TSP-1 is operating a railway line to and from stations A and F, stopping at the intermediate stations of B, C, D and E. On the other hand, TSP-2 is operating a line to and from stations G and J, with intermediate stops at H, C, D, E and I. Since these TSPs are not in direct competition with each other, there is a possibility of revenue improvements by coordinating their train schedules at a transfer node (e.g. station D) so as to attract an additional passenger demand travelling across the regions. Not only does such coordination create an inter-modal competition (e.g. with road traffic), but also introduces an intra-modal competition when there is an

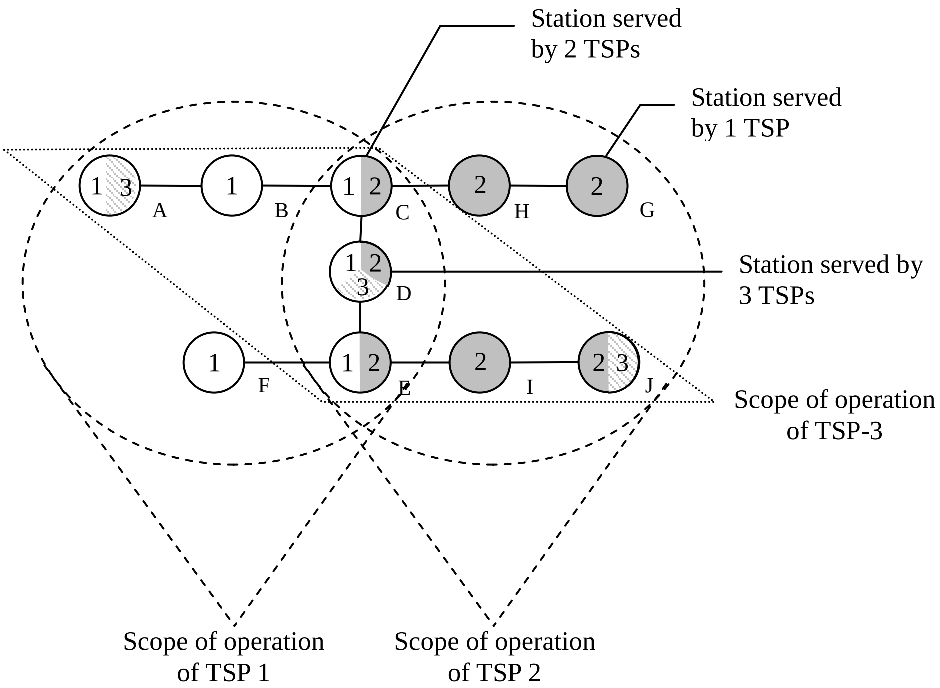


Fig. 7.1. Competition of rail services between three TSPs

intercity service (TSP-3) operating across the regions.

Passengers transferring between train services are often discouraged when the waiting time for the transit is substantial, especially when an alternative means of transportation is available. Therefore, the problem on schedule coordination mainly aims to reduce, and possibly minimise, the passenger waiting time at the interchange station. Such problem is not novel in railways, and it has been extensively modelled and examined under an integrated railway. Minimisation of waiting time is usually obtained by adjusting the commencement time of two services so that headways and travelling times are preserved to avoid disturbing the quality of service of the individual lines (Burkard, 1986; Brucker *et al.*, 1990; Nachtigall & Voget, 1996). In these studies, when coordinating schedules at a single station, the arrival times of a line at the station have been modelled by a set of vertices of a polygon within a unit circle (Burkard, 1986; Brucker *et al.*, 1990). The problem is then to minimise the total arc lengths between the vertices on the circumference of the circle. On the other hand, when coordinating a set of trains at multiple interchange stations, the problem has been shown to be NP-hard (Nachtigall, 1996) and it has been solved by a branch-and-bound algorithm for optimal solution (Nachtigall, 1996), and a genetic algorithm for near-optimal solutions (Nachtigall & Voget, 1996).

Despite the extensive effort in the schedule coordination problem in the integrated railways, the introduction of open access has altered the nature of the problem. Firstly, the lines are now managed by different TSPs instead of a single authority. As a result, the alignment of schedules requires a mutual agreement from more than one party, whose operating constraints and objectives may be in conflict with those of other operators. In particular, there may be constraints regarding to the earliest commencement time due to the availability of rolling stock, and it is also desirable to consider the cost of idle time for the rolling stock. Moreover, sensitive data such as cost rates are unlikely to be revealed to the other TSPs, which means decisions on the coordinated schedules are often made under incomplete information through negotiation activities. These changes prompt for the remodelling of the schedule coordination problem.

7.2. Mathematical Modelling

7.2.1. Cost Function

Let ζ_i be the commencement time of the train service of line L_i. Since the adjustment of ζ_i may result in an idle usage of rolling stock of L_i, let I_i represents the cost of idle time. In addition, let D_{ij} be the number of passengers transferring from L_i to L_j at interchange station X (where $i \neq j$). The coordination problem of two services therefore involves two demands, namely D_{12} and D_{21}. Suppose k_i

is the average charge for a transferring passenger travelling with L_i, then the improvement by coordinating the schedules (i.e. cost) can be expressed by the difference of the revenue gained and the idle cost of rolling stock in (7.1).

$$Y_i = k_i (D_{ij} + D_{ji}) - I_i \tag{7.1}$$

7.2.1.1. Idle Cost

Let $\hat{\zeta}_i$ be the release date of the rolling stock of L_i. If L_i commences at $\hat{\zeta}_i$, then the idle cost is zero. As the commencement time is postponed, the idle cost is increased proportionally. Let c_i be the unit cost of idle time for the rolling stock. The idle cost is then modelled by the function $F : Z \to R^+$ in (7.2). A commencement time earlier than the release date is not permissible.

$$I_i = F(\zeta_i) = c_i (\zeta_i - \hat{\zeta}_i) \quad \text{for} \quad \zeta_i \geq \hat{\zeta}_i \tag{7.2}$$

7.2.1.2. Passenger Demand

Let h_i be the total time required for L_i to arrive at X from the first station, d_i be the dwell time at X, and κ_{ij} be the minimum transfer time from L_i to L_j, then the arrival time A_i and departure time B_i for line L_i at station X are modelled by (7.3) and (7.4) respectively. The passenger waiting time w_{ij} can in turn be expressed by (7.5).

$$A_i = \zeta_i + h_i \tag{7.3}$$

$$B_i = \zeta_i + h_i + d_i \tag{7.4}$$

$$w_{ij} = B_j - A_i - \kappa_{ij} \tag{7.5}$$

D_{ij} is assumed to be affected by the waiting time at X. The longer is the waiting time, the lower is the demand. In this study, the demand from L_i to L_j in relation to w_{ij} is modelled by a quadratic function in (7.6).

$$D_{ij} = G_{ij}^*\left[1-\left(\frac{w_{ij}}{w_m}\right)^2\right] \quad \text{for} \quad 0 \le w_{ij} \le w_m \tag{7.6}$$

When the waiting time is zero, the function achieves maximum demand G_{ij}^*. As the waiting time increases, more passengers will opt for the alternative means of transportation. By substitution, D_{ij} can be expressed as a function of ζ_i and ζ_j ($G : \mathrm{Z}\times\mathrm{Z} \;\; \rightarrow [0, G_{ij}^*]$). This is given in (7.7), in which $z_{ij} = h_j + d_j - h_i - \kappa_{ij}$. z_{ij} can be regarded as the time lag that L_i should commence its service after L_j does in order to attain the maximum demand from L_i to L_j.

$$D_{ij} = G(\zeta_i, \zeta_j) = G_{ij}^*\left[1-\left(\frac{\zeta_j - \zeta_i + z_{ij}}{w_m}\right)^2\right] \quad \text{for} \quad 0 \le \zeta_j - \zeta_i + z_{ij} \le w_m \tag{7.7}$$

7.2.1.3. Analysis of Solution Space

The overall cost function is summarised by (7.8), where the definitions of $F(\zeta_i)$ and $G(\zeta_i, \zeta_j)$ are given in (7.2) and (7.7).

$$Y_i = k_i G(\zeta_i, \zeta_j) + k_i G(\zeta_j, \zeta_i) - F(\zeta_i) \tag{7.8}$$

According to the constraints associated with $F(\zeta_i)$ and $G(\zeta_i, \zeta_j)$, the solution space of the cost function is discrete. An example for two service providers is shown in Fig. 7.2 which has been constructed with $k_1 = 15$, $k_2 = 22$, $G_{12}^* = 100$, $G_{21}^* = 80$,

$h_1 = 20$, $h_2 = 30$, $d_1 = 5$, $d_2 = 7$, $\hat{\zeta}_1 = 7$, $\hat{\zeta}_2 = 5$, $c_1 = 50$, $c_2 = 60$ and $\kappa_{12} = \kappa_{21} = 2$. There are three discontinuities in the solution space. The first one occurs at $\zeta_i = \hat{\zeta}_i$ where for $\zeta_i < \hat{\zeta}_i$, there are no feasible solutions ($Y_i = 0$). The second and third discontinuities partition the solutions corresponding to unidirectional and bidirectional transfer of passengers. In this case, the global maximum resides in

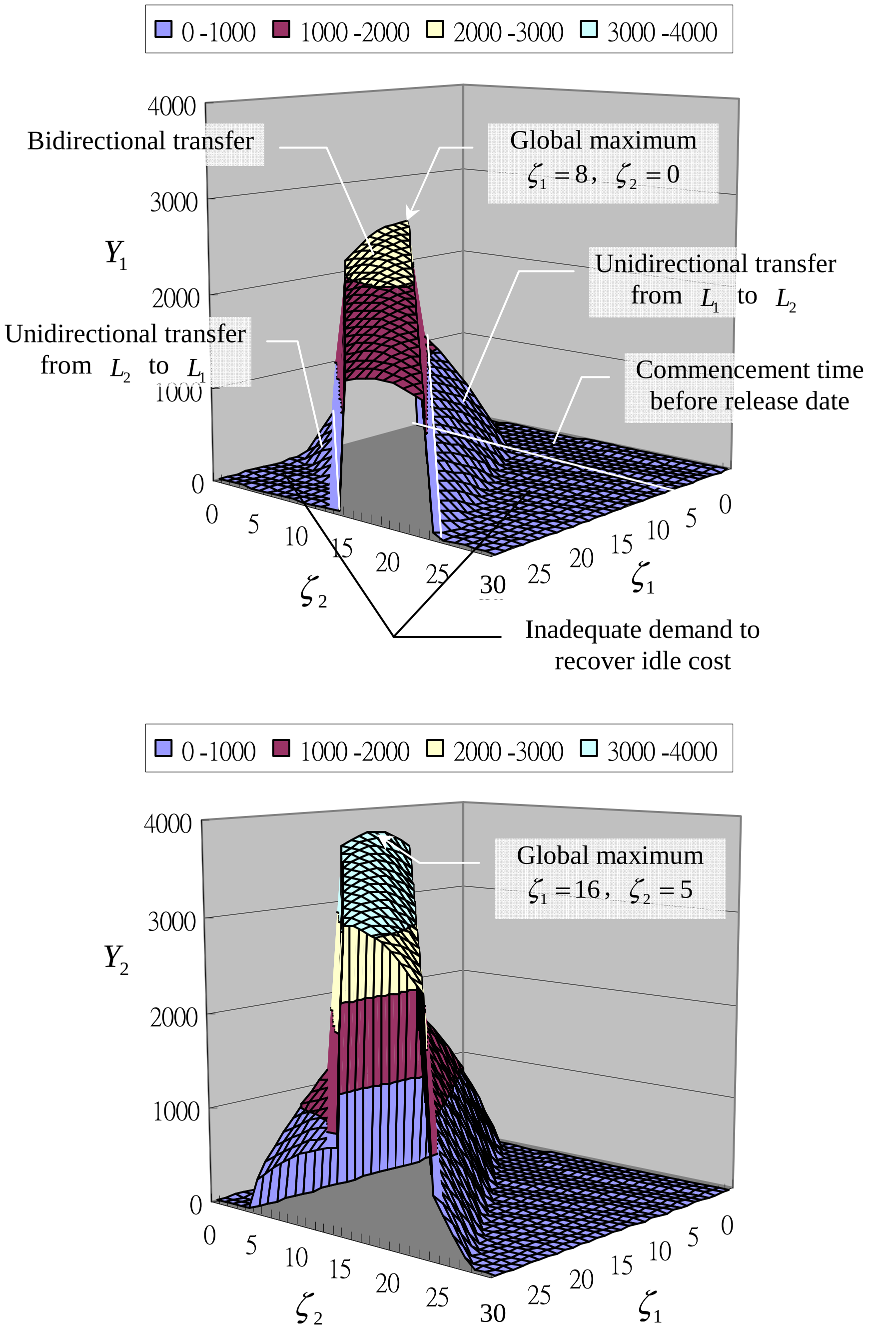


Fig. 7.2. Solution space of objective functions of TSP-1 and TSP-2

the bidirectional transfer region. There are also regions of zero costs owing to the inadequate demand to recover the idle cost.

7.2.2. Negotiation Protocol

In Fig. 7.2, the global maximum for TSP-1 occurs at $\zeta_1 = 8$ and $\zeta_2 = 0$, yielding $Y_1 = 2463$. However, from the perspective of TSP-2, this is an infeasible solution because $\zeta_2 < \hat{\zeta}_2 = 5$. On the other hand, the global optimum for TSP-2 is located at $\zeta_1 = 16$ and $\zeta_2 = 5$, giving $Y_2 = 3802$. Although this is a feasible solution for TSP-1, the solution only yields $Y_1 = 2142$. Since there are solutions with cost greater than 2142 for TSP-1, it is desirable to devise proper negotiation strategies for the TSPs with various expectations on the cost (i.e. Y_i) and the length of negotiation (i.e. number of negotiation rounds).

7.2.2.1. Simple Exchanges of Offers

Negotiation is conducted by the exchange of offers in a number of rounds. The TSP agent submitting the first offer is the initiator, while the negotiating partner (proponent) is the responder.

An offer at round k consists of the proposed commencement times of the initiator i and the responder j. An offer is therefore modelled by (7.9).

$$O^k = \{\zeta_i^k, \zeta_j^k\} \tag{7.9}$$

The cost associated with the offer O^k is assumed to be stored internally by the agent, represented by Y_i^k. Suppose TSP-1 is the initiator, then the offers in the odd rounds of negotiation (i.e. $k = 2m - 1$) are proposed by TSP-1, while offers in the even rounds of negotiation (i.e. $k = 2m$) are generated by TSP-2.

The negotiation procedure is shown in Fig. 7.3. The action set of an agent is given by {PROPOSE, ACCEPT, FAILURE}. At the beginning, the initiator generates the offer which maximises (7.8). If the offer exists, it is proposed to the proponent. Otherwise, no action is taken (no more negotiation activities are needed). Upon the arrival of the counteroffer from the proponent, the agent evaluates the associated cost of

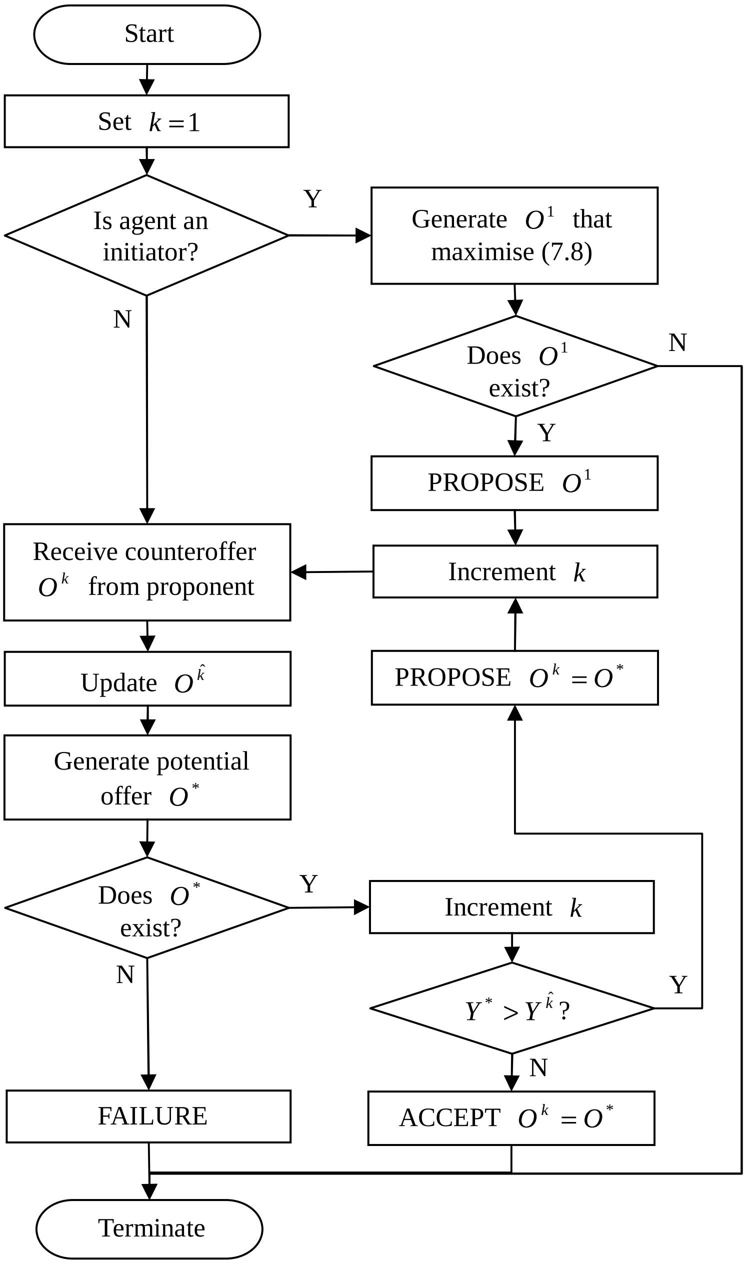


Fig. 7.3. Negotiation procedure for TSP-TSP transaction

the counteroffer and update $O^{\hat{k}}$, which is the first occurrence of counteroffer with the highest cost $Y_i^{\hat{k}}$ received at round $\hat{k}$. In addition, the agent also computes the next potential offer O^* using one of the strategies S_{po}, S_{min} or S_{max} (definitions are described in Section 7.2.2.2). If no potential offer can be found, the negotiation is terminated with the action FAILURE. If the offer exists, the agent proposes O^* if $Y_i^* > Y_i^{\hat{k}}$, and accepts $O^{\hat{k}}$ otherwise.

7.2.2.2. Strategies

i) *Strategy-PO (S_{po})*: This strategy aims to derive the Pareto-optimal solution and it requires both agents to employ this strategy to achieve the objective. According to the definition of Pareto-optimality (Ehtamo *et al.*, 1996), a solution s is Pareto-optimal if there does not exist any alternative solution s' which improves the costs of all negotiating parties.

Definitions: By definition, the initiator is proposing at rounds $k = 2m-1$ while the responder is proposing at rounds $k = 2m$. In other words, the sequence of offers generated by the initiator is $O^1, O^3, ..., O^{2m-1}$ and the sequence of offers of the responder is $O^2, O^4, ..., O^{2m}$. In this strategy, the feasible offers of an agent are arranged in descending order of their costs, that is, for the initiator, $Y_1^1 \geq Y_1^3 \geq ... \geq Y_1^{2m-1}$ and for the responder, $Y_2^2 \geq Y_2^4 \geq ... \geq Y_2^{2m}$.

Proof: The ability to arrive at Pareto-optimal solution can be proven by contradiction. Assume that the condition of acceptance is detected by the initiator after round k_D and $O^{\hat{k}}$ is accepted. If $O^{\hat{k}}$ is not Pareto-optimal, then there exists another offer O' that does not decrease the cost of either agent. To determine whether such offer does exist, the offers are divided into three partitions as shown in Fig. 7.4.

Partition A: This partition consists of the proposals prior to round $\hat{k}$. In the odd rounds within this set (i.e. $2m-1<\hat{k}$), although the costs of the initiator are higher (i.e. $Y_1^{2m-1} \geq Y_1^{\hat{k}}$), the costs of the responder are lower (i.e. $Y_2^{2m-1} < Y_2^{\hat{k}}$). Otherwise the condition of acceptance would have been detected by the responder (Fig. 7.4). Since these solutions cause a decrease in Y_2, they are not Pareto-optimal. On the other hand, in the even rounds (i.e. $2m<\hat{k}$), although the costs of the responder are higher (i.e.

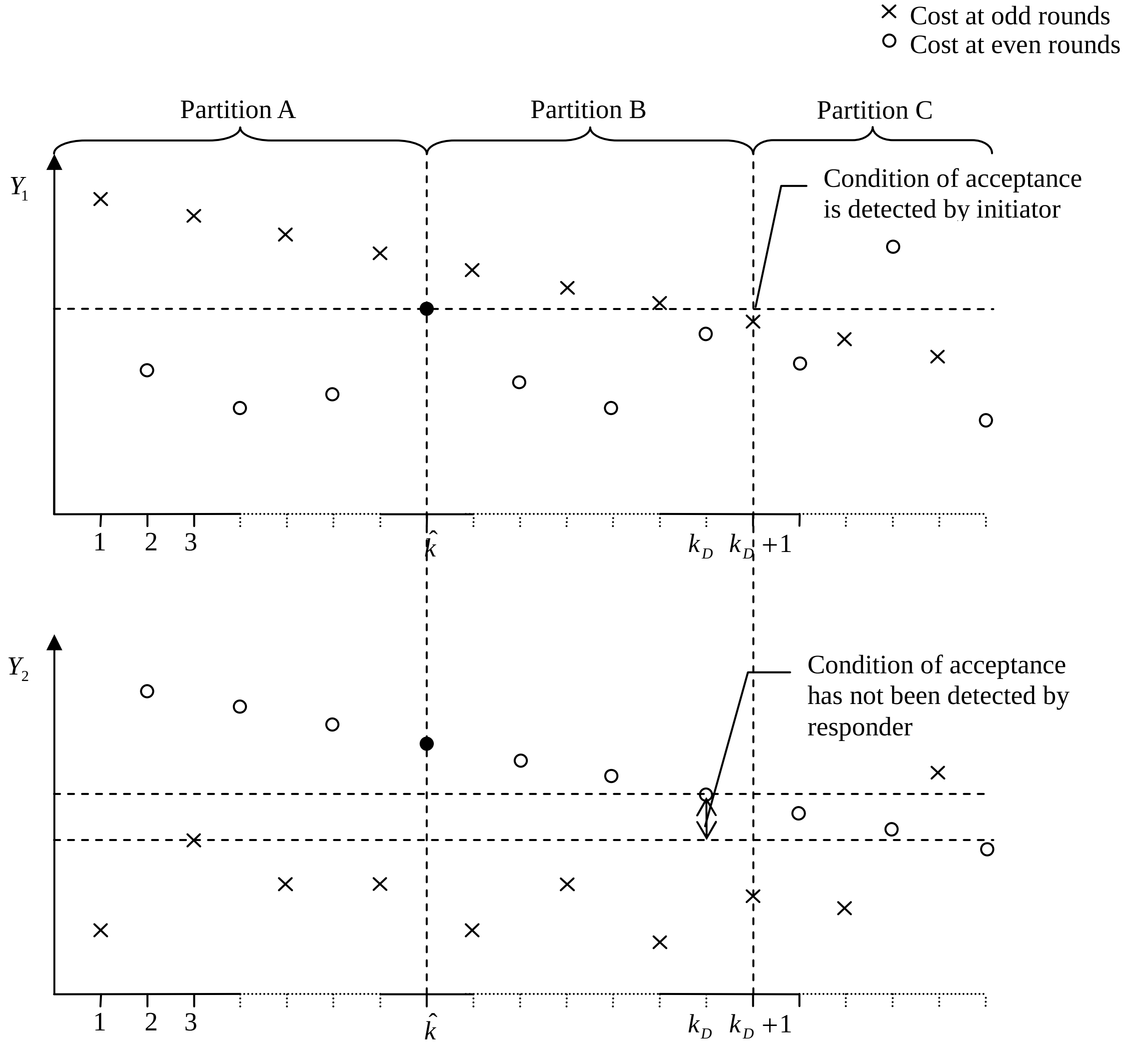


Fig. 7.4. Illustration of proof of Pareto-optimality

$Y_2^{2m} \geq Y_2^{\hat{k}}$), the costs of the initiator are smaller (i.e. $Y_1^{2m} < Y_1^{\hat{k}}$) because by definition, $Y_1^{\hat{k}}$ is the first highest cost of the counteroffers. Therefore, these solutions are also not Pareto-optimal.

Partition B: This partition consists of the proposals between round $\hat{k}$ and round $k_D + 1$ exclusively. For the costs in the odd rounds, the same argument holds as the odd rounds in partition A. In the even rounds ($\hat{k} < 2m < k_D + 1$), both costs are smaller by definition (i.e. $Y_2^{2m} < Y_2^{\hat{k}}$ and $Y_1^{2m} < Y_1^{\hat{k}}$). In other words, all the other offers that have been proposed (partition A and B) cannot improve Y_1 and Y_2 simultaneously.

Partition C: To examine the remaining offers that have not been proposed, the negotiation is assumed to continue. In the odd rounds of negotiation, Y_1 is decreasing, so even if $Y_2 > Y_2^{\hat{k}}$, the offer is not Pareto-optimal. Similarly, in the even rounds of negotiation, since Y_2 is decreasing, these proposals cannot be Pareto-optimal.

As a result, no offers can improve the costs of both parties simultaneously when the condition of acceptance is detected by the initiator. The proof for the responder can be constructed in a similar manner. This completes the proof.

To reach the Pareto-optimal solution, both parties must employ S_{po}. Despite the theoretical significance of such solution, stakeholders often aim to achieve a better cost in practice, even if the proponent suffers from a loss. As a consequence, it is also worth examining other negotiation strategies (or combination of strategies) and comparing their resulting offers from the Pareto-optimal solution obtained by S_{po}. Two additional strategies are proposed below. In these strategies, it is assumed that only one variable can be changed in O^{k+1} with respect to O^k or O^{k-1}. This is inserted to reduce the computational complexity.

ii) *Strategy-MIN (S_{min})*: This strategy attempts to reduce the concession made from the most recent offer proposed by the agent itself. Agents employing this strategy are expected to behave cautiously during the negotiation because the generated offers do not take the proponent's requirements into consideration. If the negotiating agents begin with extreme costs, the convergence can be slow. Even though the final costs obtained should be of good quality, they may not be necessarily optimal.

Definitions: Suppose an agent has just received the counteroffer O^k. In this strategy, the potential offer O^* is derived by (7.10), where O' and O'' are offers having costs Y' and Y'', which are found by (7.11) and (7.12) respectively. Y_i^{k-1} is the cost associated with $O^{k-1} = \{\zeta_i^{k-1}, \zeta_j^{k-1}\}$, which was the most recent offer proposed in round $k-1$. $Y_{\partial\zeta_i}$ and $Y_{\partial\zeta_j}$ are the cost of offers $O_{\partial\zeta_i} = \{\zeta_i, \zeta_j^{k-1}\}$ and $O_{\partial\zeta_j} = \{\zeta_i^{k-1}, \zeta_j\}$ respectively.

$$O^* = \begin{cases} O' & \text{for } Y' > Y'' \\ O'' & \text{otherwise} \end{cases} \tag{7.10}$$

$$Y' = \arg_{Y_{\partial\zeta_i}} \{\min(Y_i^{k-1} - Y_{\partial\zeta_i})\} \tag{7.11}$$

$$Y'' = \arg_{Y_{\partial\zeta_j}} \{\min(Y_i^{k-1} - Y_{\partial\zeta_j})\} \tag{7.12}$$

iii) *Strategy-MAX (S_{max})*: This strategy attempts to maximise the difference of cost from the most recent offer received from the proponent agent. Agents employing this strategy are expected to behave desperately during the negotiation because the generated offers are modified from the proponent's offers. The strategy is likely to obtain a fast convergence of solution, but the cost obtained may not be in good quality.

Definitions: Suppose an agent has just received the counteroffer O^k. In this strategy, the potential offer O^* is derived by (7.10), where O' and O'' are offers having costs Y' and Y'', which are found by (7.13) and (7.14) respectively. Y_i^k is the cost associated with $O^k = \{\zeta_i^k, \zeta_j^k\}$, which was the most recent counteroffer received. $Y_{\partial\zeta_i}$ and $Y_{\partial\zeta_j}$ are the cost of offers $O_{\partial\zeta_i} = \{\zeta_i, \zeta_j^k\}$ and $O_{\partial\zeta_j} = \{\zeta_i^k, \zeta_j\}$ respectively.

$$Y' = \arg_{Y_{\partial\zeta_i}} \{\max(Y_{\partial\zeta_i} - Y_i^k)\} \tag{7.13}$$

$$Y'' = \arg_{Y_{\partial\zeta_j}} \{\max(Y_{\partial\zeta_j} - Y_i^k)\} \tag{7.14}$$

7.3. Algorithms for Generation of Offers

In the definitions of the negotiation strategies, it has been assumed that there is an algorithm to generate the necessary offers. In all strategies, each TSP is required to find the global optimum with respect to its internal benefits defined by (7.8). An additional method is then needed to generate the remaining offers other than the global optimum. For S_{po}, a sequence of offers with decreasing cost is needed. For S_{min} and S_{max}, the generation of a subsequent offer is constrained by the optimisation of cost difference along one of the variables while holding the other one constant. This section proposes a feasible algorithm for both problems.

7.3.1. Optimisation Algorithm

7.3.1.1. Alternative Representation of Optimisation Problem

Prior to describing the algorithm to solve the optimisation problem in (7.8), the problem is expressed in a different manner so that the choice of algorithm is made more apparent. Owing to the natural separation of the solution space shown in Fig. 7.2, the problem of maximising (7.8) can be equivalently solved by minimising the following three transfer problems. The details of the conversion are described in Appendix C.

1) $\mathbf{P_{i\to j}}$: Unidirectional transfer from L_i to L_j

$$\begin{aligned}\min Z_i &= \frac{2k_i G_{ij}^* z_{ij}}{w_m{}^2}\begin{bmatrix} c_i w_m{}^2 / 2k_i G_{ij}^* z_{ij} - 1 \\ 1 \end{bmatrix}^T \begin{bmatrix} \zeta_i \\ \zeta_j \end{bmatrix} \\ &+ \frac{1}{2}\frac{2k_i G_{ij}^*}{w_m{}^2}\begin{bmatrix} \zeta_i \\ \zeta_j \end{bmatrix}^T \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} \zeta_i \\ \zeta_j \end{bmatrix}\end{aligned} \tag{7.15}$$

subject to:

$$\zeta_i, \zeta_j \geq 0, \in Z \tag{7.16}$$

$$\zeta_i \geq \hat{\zeta}_i \tag{7.17}$$

$$\zeta_j - \zeta_i + z_{ij} \geq 0 \tag{7.18}$$

$$\zeta_j - \zeta_i + z_{ij} \leq w_m \tag{7.19}$$

$$\zeta_i - \zeta_j + z_{ji} \leq -1 \tag{7.20}$$

Constraint (7.16) is the non-negativity constraint on the variables ζ_i and ζ_j, which are also integers for the purpose of timetabling. Constraint (7.17) is associated to the release date, and (7.18) and (7.19) ensure the validity of the passenger waiting time for the transfer from L_i to L_j. Constraint (7.20) is applied so that the simultaneous occurrence of transfer in the opposite direction (L_j to L_i) is excluded from consideration.

2) $\mathbf{P}_{\mathbf{j}\to\mathbf{i}}$: Unidirectional transfer from L_j to L_i

$$\begin{aligned} \min Z_i = & \frac{2k_i G_{ji}^* z_{ji}}{w_m{}^2} \begin{bmatrix} c_i w_m{}^2 / 2k_i G_{ji}^* z_{ji} + 1 \\ -1 \end{bmatrix}^T \begin{bmatrix} \zeta_i \\ \zeta_j \end{bmatrix} \\ & + \frac{1}{2} \frac{2k_i G_{ji}^*}{w_m{}^2} \begin{bmatrix} \zeta_i \\ \zeta_j \end{bmatrix}^T \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} \zeta_i \\ \zeta_j \end{bmatrix} \end{aligned} \tag{7.21}$$

subject to (7.16), (7.17) and

$$\zeta_i - \zeta_j + z_{ji} \geq 0 \tag{7.22}$$

$$\zeta_i - \zeta_j + z_{ji} \leq w_m \tag{7.23}$$

$$\zeta_j - \zeta_i + z_{ij} \leq -1 \tag{7.24}$$

The above constraints have similar interpretations as (7.18)-(7.20) described in the problem $\mathbf{P}_{\mathbf{i}\to\mathbf{j}}$.

3) $\mathbf{P}_{\mathbf{i}\leftrightarrow\mathbf{j}}$: Bidirectional transfer to and from L_i and L_j

$$\begin{aligned} \min Z_i = & \frac{2k_i}{w_m{}^2} \begin{bmatrix} G_{ji}^* z_{ji} - G_{ij}^* z_{ij} + \dfrac{c_i w_m{}^2}{2k_i} \\ G_{ij}^* z_{ij} - G_{ji}^* z_{ji} \end{bmatrix}^T \begin{bmatrix} \zeta_i \\ \zeta_j \end{bmatrix} \\ & + \frac{1}{2} \frac{2k_i}{w_m{}^2} \begin{bmatrix} \zeta_i \\ \zeta_j \end{bmatrix}^T \begin{bmatrix} G_{ij}^* + G_{ji}^* & -(G_{ij}^* + G_{ji}^*) \\ -(G_{ij}^* + G_{ji}^*) & G_{ij}^* + G_{ji}^* \end{bmatrix} \begin{bmatrix} \zeta_i \\ \zeta_j \end{bmatrix} \end{aligned} \tag{7.25}$$

subject to (7.16)-(7.18) and (7.22)-(7.23).

7.3.1.2. Quadratic Programming

The above problems can be expressed in the general form in (7.26), where $\mathbf{x} = [\zeta_i \quad \zeta_j]^T$ is the column vector of variables and the remaining matrices for the corresponding optimisation problems are summarised in Table 7.1.

Table 7.1. Matrices for Optimisation Problems

	c	**H**	**A**	**b**
$\mathbf{P}_{\mathbf{i\rightarrow j}}$	$\frac{2k_i G_{ij}^{*} z_{ij}}{w_m^{\ 2}} \begin{bmatrix} c_i w_m^{\ 2} / 2k_i G_{ij}^{*} z_{ij} - 1 \\ 1 \end{bmatrix}$	$\frac{2k_i G_{ij}^{*}}{w_m^{\ 2}} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$	$\begin{bmatrix} -1 & 0 \\ 1 & -1 \\ -1 & 1 \\ 1 & -1 \end{bmatrix}$	$\begin{bmatrix} -\hat{\zeta}_i \\ z_{ij} \\ w_m - z_{ij} \\ -1 - z_{ji} \end{bmatrix}$
$\mathbf{P}_{\mathbf{j\rightarrow i}}$	$\frac{2k_i G_{ji}^{*} z_{ji}}{w_m^{\ 2}} \begin{bmatrix} c_i w_m^{\ 2} / 2k_i G_{ji}^{*} z_{ji} + 1 \\ -1 \end{bmatrix}$	$\frac{2k_i G_{ji}^{*}}{w_m^{\ 2}} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$	$\begin{bmatrix} -1 & 0 \\ -1 & 1 \\ 1 & -1 \\ -1 & 1 \end{bmatrix}$	$\begin{bmatrix} -\hat{\zeta}_i \\ z_{ji} \\ w_m - z_{ji} \\ -1 - z_{ij} \end{bmatrix}$
$\mathbf{P}_{\mathbf{i\leftrightarrow j}}$	$\frac{2k_i}{w_m^{\ 2}} \begin{bmatrix} G_{ji}^{*} z_{ji} - G_{ij}^{*} z_{ij} + \frac{c_i w_m^{\ 2}}{2k_i} \\ G_{ij}^{*} z_{ij} - G_{ji}^{*} z_{ji} \end{bmatrix}$	$\frac{2k_i}{w_m^{\ 2}} \begin{bmatrix} G^{*} & -G^{*} \\ -G^{*} & G^{*} \end{bmatrix}$ where $G^{*} = G_{ij}^{*} + G_{ji}^{*}$	$\begin{bmatrix} -1 & 0 \\ 1 & -1 \\ -1 & 1 \\ -1 & 1 \\ 1 & -1 \end{bmatrix}$	$\begin{bmatrix} -\hat{\zeta}_i \\ z_{ij} \\ w_m - z_{ij} \\ z_{ji} \\ w_m - z_{ji} \end{bmatrix}$

$$\min \quad \{ f(\mathbf{x}) = \mathbf{c}^T \mathbf{x} + \frac{1}{2} \mathbf{x}^T \ \mathbf{H}\mathbf{x} : \mathbf{A}\mathbf{x} \leq \mathbf{b}, \mathbf{x} \geq \mathbf{0}, x_i \in Z \} \tag{7.26}$$

With the objective function being quadratic (non-linear) with a set of linear constraints, the problem in (7.26) is a standard quadratic programming problem (Bazaraa *et al.*, 1993; Hillier & Lieberman, 1995) if the integer constraint on the variables is eliminated from considerations.

Although not all quadratic programming problems can be solved analytically, it has been shown that if **H** is positive semi-definite, then the quadratic programming problem can be reduced to a linear programming problem with an additional complementary constraint (Appendix D), which can be solved efficiently by algorithms such as the Modified Simplex Algorithm (Hillier & Lieberman, 1995) or the Lemke's Complimentary Pivoting Algorithm (Bazaraa *et al.*, 1993). Verifications of **H** for the three transfer problems show that they are all positive semi-definite (Appendix E).

7.3.1.3. Lemke's Complementary Pivoting Algorithm

According to the discussion in Appendix D, the quadratic programming problem in

(7.26) excluding the integer constraint can be transformed to a linear programming problem in (7.27) subject to a complementary constraint (7.28).

$$\mathbf{w}-\mathbf{Mz}-\mathbf{1}z_0=\mathbf{q} \tag{7.27}$$

$$\mathbf{w}^{\mathbf{T}}\mathbf{z}=0 \tag{7.28}$$

$$\mathbf{w},\,\mathbf{z}\geq\mathbf{0} \tag{7.29}$$

where $\mathbf{M}=\begin{bmatrix}\mathbf{0} & -\mathbf{A}\\ \mathbf{A}^{\mathbf{T}} & \mathbf{H}\end{bmatrix}$, $\mathbf{q}=\begin{bmatrix}\mathbf{b}\\ \mathbf{c}\end{bmatrix}$, $\mathbf{w}=\begin{bmatrix}\mathbf{y}\\ \mathbf{v}\end{bmatrix}$, $\mathbf{z}=\begin{bmatrix}\mathbf{u}\\ \mathbf{x}\end{bmatrix}$, $\mathbf{u}$ and $\mathbf{v}$ are the Lagrangian multiplier vectors, and $\mathbf{y}$ is the vector for the slack variables. This problem can be solved by Lemke's Complementary Pivoting Algorithm (LCPA) (Bazaraa *et al.*, 1993). The algorithm consists of two stages: the initialisation stage and the main stage.

Initialisation stage: Display the system defined by (7.27)-(7.29) in a tableau format shown in Table 7.2. $\mathbf{w}=\mathbf{q}$ are the basic variables and $\mathbf{z}=\mathbf{0}$, $z_0=0$ are the non-basic variables. If $\mathbf{q}\geq 0$, terminate the algorithm; $(\mathbf{w},\mathbf{z})=(\mathbf{q},\mathbf{0})$ is the optimal solution.

Otherwise, let $-q_s=\max\{-q_i:1\leq i\leq p\}$ (the most negative element of $\mathbf{q}$). Update the tableau using Gaussian elimination by pivoting at row s and the z_0 column. This makes z_0 becomes basic and w_s becomes non-basic. Let $y_s=z_s$

Table 7.2. An Example of Tableau for Lemke's Complementary Pivoting Algorithm

	w_1	w_2	w_3	w_4	w_5	w_6	z_1	z_2	z_3	z_4	z_5	z_6	z_0	q
w_1	1	0	0	0	0	0	0	0	0	0	-1	0	-1	-7
w_2	0	1	0	0	0	0	0	0	0	0	1	-1	-1	15
w_3	0	0	1	0	0	0	0	0	0	0	-1	1	-1	5
w_4	0	0	0	1	0	0	0	0	0	0	1	-1	-1	6
w_5	0	0	0	0	1	0	1	-1	1	-1	-7.5	7.5	-1	-62.5
w_6	0	0	0	0	0	1	0	1	-1	1	7.5	-7.5	-1	112.5

and go to the main stage.

Main stage: The main stage consists of the following four steps.

Step 1: Let $\mathbf{d}_s$ be the updated column in the tableau under the variable y_s. If $\mathbf{d}_s \le 0$, go to step 4. Otherwise, determine the index r by the minimum ratio test below, where $\overline{\mathbf{q}}$ is the updated right-hand-side column. If the basic variable at row r is z_0, go to step 3. Otherwise, go to step 2.

$$\frac{\overline{q}_r}{d_{rs}} = \min_{1 \le i \le p} \left\{ \frac{\overline{q}_i}{d_{is}} : d_{is} > 0 \right\}$$

Step 2: The tableau is updated by pivoting at row r and the y_s column. The entering basic variable is y_s, and the leaving basic variable at row r is either w_l or z_l, for some $l \ne s$. In the former case set $y_s = z_l$, and in the latter, set $y_s = w_l$. Return to step 1.

Step 3: Here y_s enters the basis, and z_0 leaves the basis. Pivot at the y_s column and the z_0 row. The resulting solution $(\mathbf{w}, \mathbf{z})$ is optimal. Terminate the algorithm.

Step 4: Stop. The problem has no solution.

7.3.2. Algorithm for Strategy-PO

7.3.2.1. Pruning Tree Searching Algorithm

With the ability to find the optimal solution for the relaxed (non-integer) problems, the remaining problem is to find a sequence of solution of descending order of costs.

In theory, the definition of S_{po} requires an exhaustive list for offers. In practice, since the negotiation often terminates in a finite number of rounds, it is possible to generate a sub-list of size n_l that is greater than the expected number of rounds, but significantly less than the total number of feasible offers. In other words, it is not necessary to exhaustively search through the entire solution space. The following devises a pruning tree searching algorithm to produce a set of feasible offers of the highest cost.

Let $Y^* = \max(Y_{i\to j}, Y_{j\to i}, Y_{i\leftrightarrow j})$ be the optimal cost generated from solving the three transfer problems using LCPA. The objective of the search is to extract all the feasible offers with $Y \geq \alpha Y^*$, where $\alpha \in [0,1]$.

The structure of the tree consists of four levels (Fig. 7.5). Level 1 consists of the root node which corresponds to the best solution generated from the three transfer problems in level 2 using LCPA. These solutions are potentially infeasible because they may not be integers.

In level 3, ζ_i is assigned with integer values governed by the function $Q_\pi^i(u)$, where $\pi \in \{i \to j,\ j \to i, i \leftrightarrow j\}$ represents the type of transfer problem and $u = \{1, 2, ...\}$ is the number of leaf nodes generated from the parent node in level 2. The cost of a node at this level is evaluated by maximising Y_π^u at $\{\zeta_i^u, \zeta_j^*\}$. $Q_\pi^i(u)$

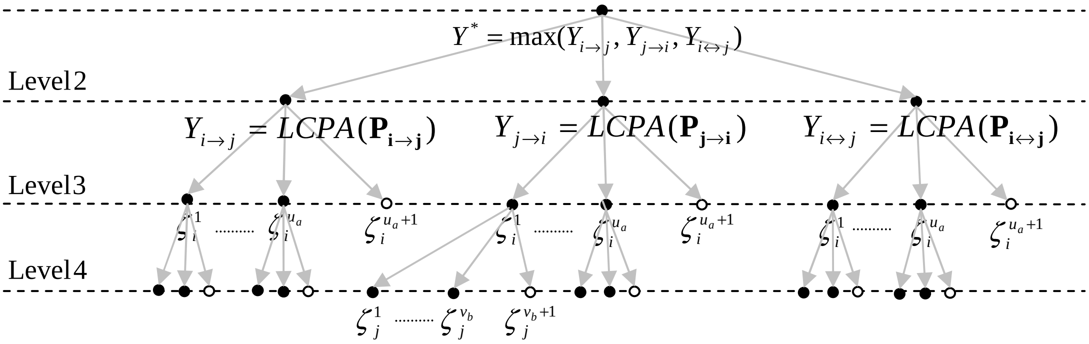


Fig. 7.5. Structure of the pruning tree searching algorithm

has the property such that $Y_{\pi}^{u} < \alpha Y^{*}$, when $u > u_a$. Hence, all nodes beyond u_a (and their leaf nodes) can be pruned.

Level 4 produces feasible solutions with integer values of ζ_i^u (inherited from the parent node) and ζ_j^v. ζ_j^v is similarly produced by the function $Q_{\pi}^{j}(v)$, where $v = \{1, 2, ...\}$, and has the property of $Y_{\pi}^{uv} \geq \alpha Y^{*}$, for $v = \{1, 2, ..., v_b\}$ and $Y_{\pi}^{uv} < \alpha Y^{*}$, for $v > v_b$. Again, all nodes can be pruned when $v > v_b$.

7.3.2.2. Optimisation at Level 3

Optimisation at this level can be summarised with respect to the three transfer problems. Since ζ_i at this level is a constant, the problems are reduced to single variable optimisation problems (of degree 2) coupled with linear constraints. These are easily solved by comparing the global optimal solutions and the boundary solutions.

1) $\mathbf{P_{i \to j}}$: Unidirectional transfer from L_i to L_j

$$\max \quad Y_u^{i \to j} = k_i G_{ij}^{*} \left[1 - \left(\frac{\zeta_j - \zeta_i + z_{ij}}{w_m} \right)^2 \right] - c_i(\zeta_i - \hat{\zeta}_i) \tag{7.30}$$

subject to (7.18)-(7.20) and

$$\zeta_j \geq 0 \tag{7.31}$$

The global optimum can be found from setting the first partial derivative of (7.20) to zero, that is, at $\dot{\zeta}_j = \zeta_i - z_{ij}$. On the other hand, the boundary solutions are provided by the constraints (7.18)-(7.20) and (7.31) which yields the lower bound of $\zeta'_j = \max\{(\zeta_i - z_{ij}), (\zeta_i + z_{ji} + 1), 0\}$ and upper bound of $\zeta''_j = w_m + \zeta_i - z_{ij}$ (Fig. 7.6).

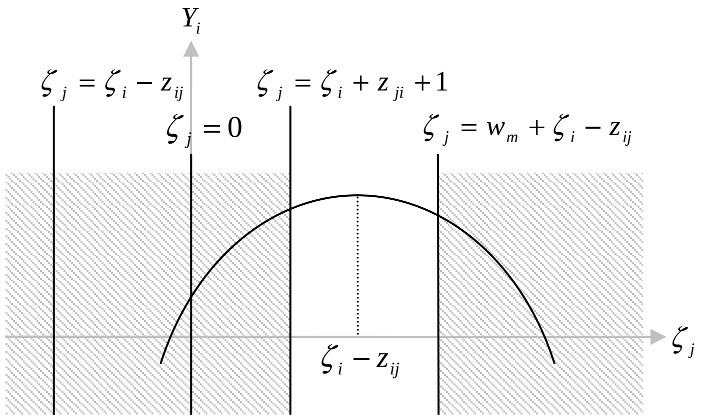


Fig. 7.6. Optimal and boundary solutions for transfer $L_i \to L_j$ at constant ζ_i

2) $\mathbf{P}_{\mathbf{j\to i}}$: Unidirectional transfer from L_j to L_i

$$\max \quad Y_u^{j\to i} = k_i G_{ji}^{*}\left[1-\left(\frac{\zeta_i - \zeta_j + z_{ji}}{w_m}\right)^2\right] - c_i(\zeta_i - \hat{\zeta}_i) \tag{7.32}$$

subject to (7.22)-(7.24) and (7.31).

The global optimum is situated at $\dot{\zeta}_j = \zeta_i + z_{ji}$ and the lower and upper boundary solutions are located at $\zeta_j' = \max\{(\zeta_i + z_{ji} - w_m), 0\}$ and $\zeta_j'' = \min\{(\zeta_i + z_{ji}), (\zeta_i - z_{ji} - 1)\}$ respectively (Fig. 7.7).

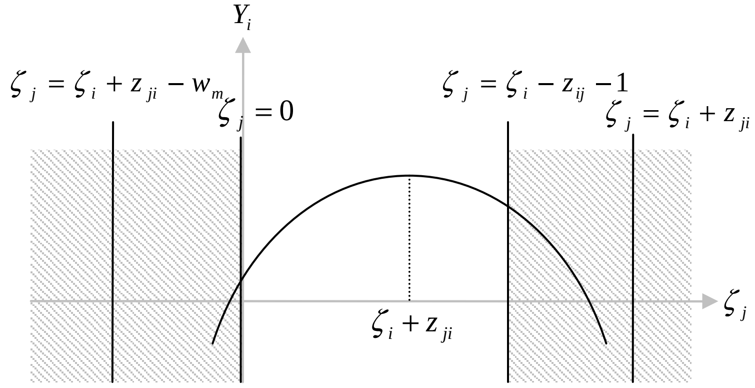


Fig. 7.7. Optimal and boundary solutions for transfer $L_j \to L_i$ at constant ζ_i

3) $\mathbf{P_{i\leftrightarrow j}}$: Bidirectional transfer to and from L_i and L_j

$$\max \quad Y_u^{i\leftrightarrow j} = k_i G_{ij}^* \left[1 - \left(\frac{\zeta_j - \zeta_i + z_{ij}}{w_m}\right)^2\right] + k_i G_{ji}^* \left[1 - \left(\frac{\zeta_i - \zeta_j + z_{ji}}{w_m}\right)^2\right] - c_i(\zeta_i - \hat{\zeta}_i) \tag{7.33}$$

subject to (7.18), (7.19), (7.22), (7.23) and (7.31).

The global optimum is situated at $\dot{\zeta}_j = \zeta_i + \dfrac{G_{ji}^* z_{ji} - G_{ij}^* z_{ij}}{G_{ij}^* + G_{ji}^*}$ and the lower and upper boundary solutions are located at $\zeta_j' = \max\{(\zeta_i - z_{ij}), (\zeta_i + z_{ji} - w_m), 0\}$ and $\zeta_j'' = \min\{(\zeta_i + z_{ji}), (w_m - \zeta_i - z_{ij})\}$ respectively (Fig. 7.8).

7.3.2.3. Pruning at Levels 3 and 4

Pruning is achieved by functions $Q_\pi^i(u)$ and $Q_\pi^j(v)$. These functions are defined below.

Definitions: $Q_\pi^i(u)$ produces a value of ζ_i while satisfying $Y_\pi^u < \alpha Y^*$, for $u > u_a$. According to Fig. 7.9 and 7.10, the smallest feasible ζ_i for problems $\mathbf{P_{i\rightarrow j}}$,

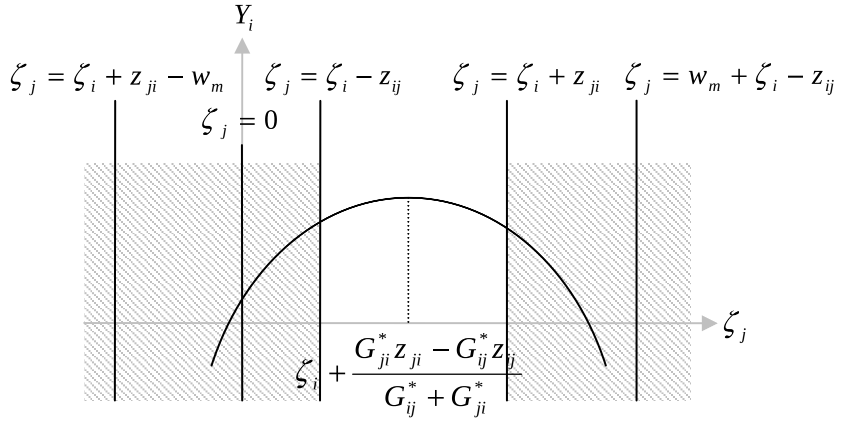


Fig. 7.8. Optimal and boundary solutions for bidirectional transfer at constant ζ_i

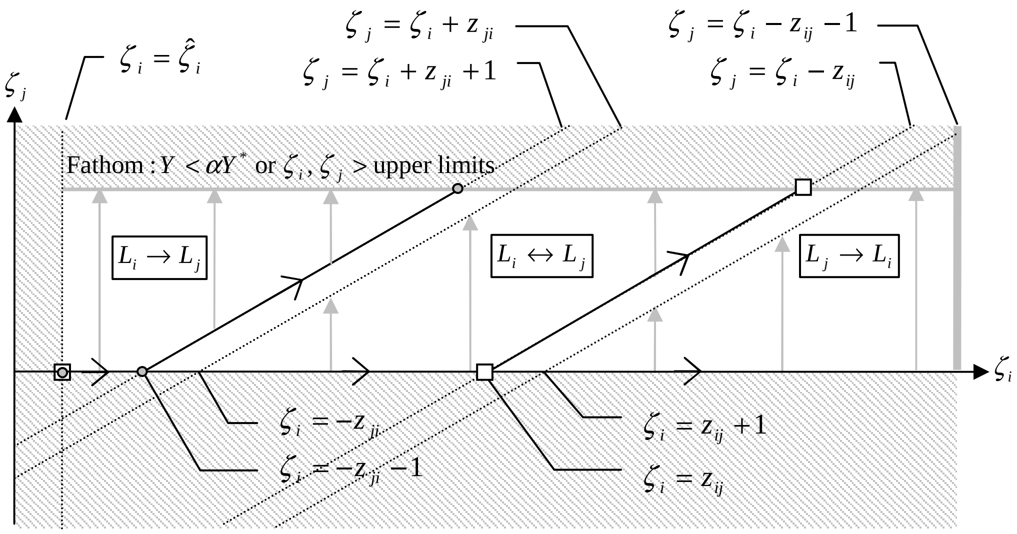


Fig. 7.9. Feasible regions when all three transfer situations are possible

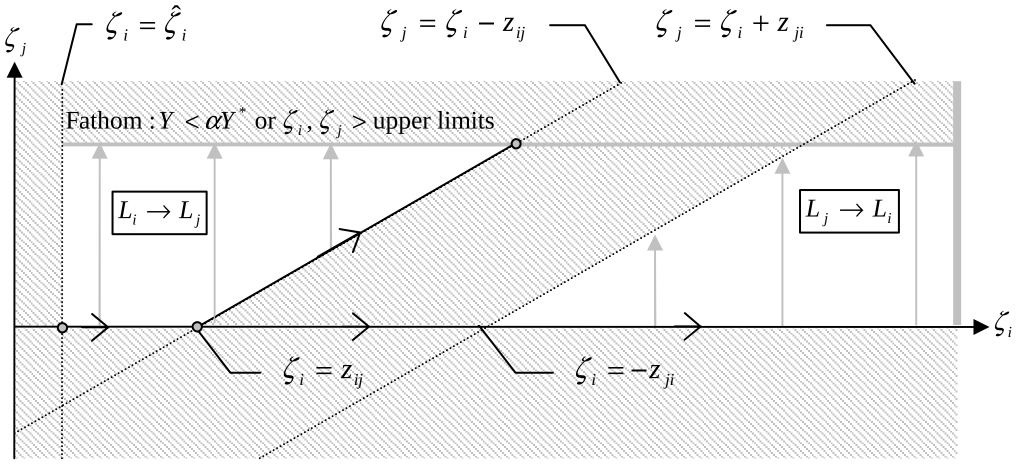
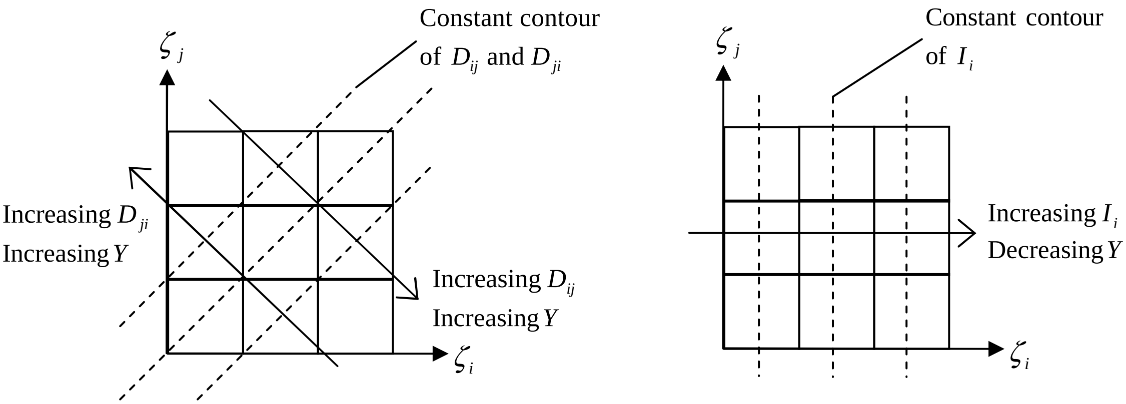


Fig. 7.10. Feasible regions when bidirectional transfer is not possible

$\mathbf{P_{j\to i}}$ and $\mathbf{P_{i\leftrightarrow j}}$ are $\zeta_i' = \hat{\zeta}_i$, $\zeta_i' = \max\{\hat{\zeta}_i, (z_{ij}+1), -z_{ji}\}$, and $\zeta_i' = \max\{\hat{\zeta}_i, -z_{ji}\}$ respectively.

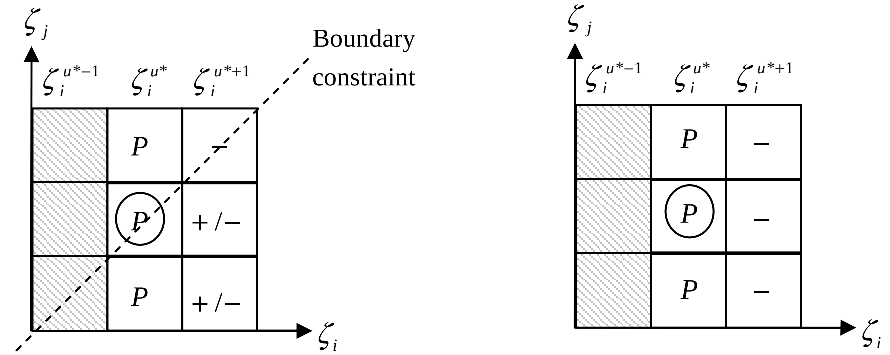
$\zeta_i^1 = \zeta_i' + p$, where $p = \{0, 1, ...\}$ is the first integer resulting to the satisfaction of $Y_\pi^u \geq \alpha \hat{Y}$ (the corresponding ζ_j^* is evaluated by checking the global optimum and boundary solutions discussed in Section 7.3.2.2). The subsequent values of $Q_\pi^i(u)$ can be determined by $Q_\pi^i(u) = \zeta_i^1 + u - 1$. Whenever $Y_\pi^u < \alpha Y^*$, the leaf nodes (for level 4) are pruned.

u_a may be determined as shown in Fig. 7.11. In Fig. 7.11a and Fig. 7.11b, the effects on Y by the passenger demands D_{ij} and D_{ji} and the idle cost I_i are isolated for consideration (refer to (7.8) for interpretation of the directions and contours). The resultant effects may increase (+) or decrease (–) Y. Suppose a current node at level 3 has been concluded to be pruned (denoted by P) with $\zeta_i = \zeta_i^u$. This means that the entire column of solutions has $Y < \alpha Y^*$. For problems $\mathbf{P_{i \to j}}$ and $\mathbf{P_{i \leftrightarrow j}}$, consider the centre box in Fig. 7.11c when $u = u^*$. Using the contours and directions in Fig. 7.11a-b, the cost at the upper left diagonal box is always lower because the demand(s) are constant and the idle cost is increasing. Although the change in the adjacent and lower diagonal boxes are uncertain, if these boxes are infeasible values (i.e. beyond the boundary constraints), then the columns beyond $u > u^*$ will not contain



(a) Trend of Y associated with D_{ij} and D_{ji}

(b) Trend of Y associated with I_i



(c) Effect on Y_i^{u+1} for $\mathbf{P_{i \to j}}$ and $\mathbf{P_{i \leftrightarrow j}}$

(d) Effect on Y_i^{u+1} for $\mathbf{P_{j \to i}}$

Fig. 7.11. Pruning conditions at Level 3

any solution with $Y \geq \alpha Y^*$. Hence $u_a = u^*$. Similarly, for problem $\mathbf{P_{j \to i}}$ (Fig. 7.11d), the entire column at $\zeta_i^{u^*+1}$ have their costs reduced, so that the columns beyond u^* can be pruned without the need of reaching the boundary constraint.

To prevent the evaluation of large number of nodes when α is small, the maximum number of nodes to be expanded by a parent at level 2 are limited by $\tilde{\zeta}_i$ (e.g. $\tilde{\zeta}_i = 60$), when the idle cost should dominate the objective functions.

Definitions: $Q_\pi^j(v)$ produces a feasible value of ζ_j while satisfying $Y_\pi^{uv} \geq \alpha Y^*$, for $v = 1, 2, ..., v_b$ and $Y_\pi^{uv} < \alpha Y^*$, for $v > v_b$. According to Fig. 7.6 to 7.8, $\zeta_j^1 = \zeta'_j + q$, where $q = \{0, 1, ...\}$ is the first integer resulting to the satisfaction of $Y \geq \alpha Y^*$. Then, the subsequent values of $Q_\pi^j(v)$ can be determined by $Q_\pi^j(v)$ $= \zeta_j^1 + v - 1$. On the other hand, v_b can be detected by the conditions of $Y_{v_b}^{\pi u} < \alpha Y^*$ or $\zeta_j^{v_b} > \zeta''_j$. As shown in Section 7.3.2.2, ζ''_j is determined by $w_m + \zeta_i - z_{ij}$, $\min\{(\zeta_i + z_{ji}), (\zeta_i - z_{ji} - 1)\}$ and $\min\{(\zeta_i + z_{ji}), (w_m - \zeta_i - z_{ij})\}$ for $\mathbf{P_{i \to j}}$, $\mathbf{P_{j \to i}}$ and $\mathbf{P_{i \leftrightarrow j}}$ respectively. As the practical values for the maximum weighting time (w_m), and the time lags required for maximum transfer demands (z_{ij} and z_{ji}) rarely exceed 60 minutes, the number of node expansions at level 4 is reasonable.

7.3.3. Algorithms for Strategy-MIN and Strategy-MAX

In these strategies, the initial offer proposed in round 1 can be generated using LCPA discussed in Section 7.3.1.3. However, to ensure that the resultant solution is feasible, the global optimal is evaluated by comparing the neighbouring solutions of $O_1 = \{Dn(\zeta_i), Dn(\zeta_j)\}$, $O_2 = \{Dn(\zeta_i) + 1, Dn(\zeta_j)\}$, $O_3 = \{Dn(\zeta_i), Dn(\zeta_j) + 1\}$, and

$O_4 = \{Dn(\zeta_i)+1, Dn(\zeta_j)+1\}$, where $Dn(\bullet)$ is a round-down-to-integer operator. In other words, $O^1 = O_i$, $i = \arg\{\max\{Y_i\} \mid i = \{1, 2, 3, 4\}\}$.

For the subsequent offers, the potential offer O^* can be obtained by comparing O' and O'', where O'' is found by the pruning tree algorithm at level 4, and O' is similarly found by holding ζ_j constant and the corresponding replacement of boundary constraints.

7.3.4. Computation Demand

The proposed algorithms for the three strategies are computationally efficient for solving the decision-making problems of the TSPs. Firstly, the quadratic programming problem is solved by LCPA which can be solved within a finite number of iterations (Bazaraa *et al.*, 1993). Also, since the optimisations at level 3 involves only the comparison of the global and boundary solutions of three quadratic functions, the computation demand is also not substantial. The only uncertainty in the algorithm is the number of nodes needed to be evaluated at the last two levels, which is controlled by α and $\tilde{\zeta}_i$ at level 3, and α and ζ''_j at level 4. However, as the values of $\tilde{\zeta}_i$ and ζ''_j are usually under 60, the concern for a high computation demand is not critical.

7.4. Simulation Setup

The case studies set up below are intended to examine the performance of the three negotiation strategies in terms of the optimality of resultant agreements (if any) and the length of negotiation. Ten simulation cases have been constructed (Table 7.3) and there are nine TSP-TSP negotiations in each case. If (S_1, S_2) denotes the strategies employed by TSP-1 and TSP-2 respectively, the combinations are generated by $S_1 \in \{S_{po},$

Table 7.3. Simulation Setup

Case	c_1	c_2	$\hat{\zeta}_1$	$\hat{\zeta}_2$	k_1	k_2	h_1	h_2	d_1	d_2	G^*_{12}	G^*_{21}	κ_{12}	κ_{21}	w_m
1	50	60	7	5	15	22	20	30	5	7	100	80	2	2	20
2	50	60	7	5	15	22	20	30	1	1	100	80	2	2	20
3	50	60	60	5	15	22	20	30	5	7	100	80	8	8	20
4	250	60	7	5	15	22	20	30	5	7	100	80	2	2	20
5	1	60	7	5	15	22	20	30	5	7	100	80	2	2	20
6	50	60	7	5	15	22	20	30	5	7	10	80	2	2	20
7	50	60	7	5	15	22	20	30	5	7	10	10	2	2	20
8	50	60	0	0	15	22	20	30	5	7	100	80	2	2	20
9	50	60	7	5	15	22	40	30	5	7	100	80	2	2	20
10	50	60	7	5	50	22	20	30	5	7	100	80	2	2	20

$S_{max}, S_{min}\} \times S_2 \in \{S_{po}, S_{max}, S_{min}\}$.

These cases represent a variety of scenarios so that the strategies are tested over a spectrum of extreme conditions. Case 1 is the example shown in Fig. 7.2. The earliest arrival times at the interchange station are 27 and 35 mins for TSP-1 and TSP-2 respectively, while the earliest departure times are 32 and 42 mins. In other words, without coordination, bidirectional transfer is impossible because the service operated by TSP-1 will depart from the interchange station 3 mins before the service of TSP-2 arrives at the platform. In case 2, the dwell times of both services are deliberately shortened to 1 min so that even when the schedules are coordinated, only unidirectional transfer can be achieved. In case 3, the release date of TSP-1 is postponed by almost an hour so as to resemble the scenario when the two TSPs begin the negotiation with substantial operational differences. Cases 4 and 5 are constructed to examine the consequence when the idle cost of rolling stock is high and low respectively. Afterwards, cases 6 and 7 are similarly paired up to study the situation when passenger demand is low in one and both direction(s) of transfer. In case 8, both services are available for operation at the beginning, and hence there are no release date constraints. In case 9, the journey time of TSP-1 is doubled so that the service of TSP-1 arrives later than the service of TSP-2. Finally, in case 10, the average charge for a transferring passenger of TSP-1 is increased by more than a factor of three.

7.5. Results and Findings

7.5.1. Agent Behaviour

Before the discussion on the performance of the strategies, it is useful to realise how the three strategies lead to an agreement between two TSP agents. To illustrate the mechanism of reaching an agreement, the negotiations of (S_{po}, S_{po}), (S_{min}, S_{po}) and (S_{max}, S_{po}) in case 1 are described in detail as follows.

7.5.1.1. Negotiation Pair (S_{po}, S_{po})

When using S_{po}, the potential offers are arranged in descending order of costs. These sequences are listed on the two columns under O^* for TSP-1 and TSP-2 respectively in Table 7.4.

At the beginning, TSP-1 initiates the negotiation by proposing {8, 0}, which is the best offer of TSP-1. The corresponding cost for TSP-2 is zero because the suggested commencement time is earlier than the release date. In replying to TSP-1, TSP-2 counter-proposes {16, 5} according to its sequence of potential offers.

From the perspective of TSP-1, the cost of {16, 5} is 87.0% (relative to the best offer proposed in round 1). Upon the arrival of the counteroffer, TSP-1 compares it

Table 7.4. Offers in Case 1 using (S_{po}, S_{po})

TSP-1					TSP-2				
Round	Counteroffer		O^*		Round	Counteroffer		O^*	
	$\{\zeta_1, \zeta_2\}$	Y_1/%	$\{\zeta_1, \zeta_2\}$	Y_1/%		$\{\zeta_1, \zeta_2\}$	Y_2/%	$\{\zeta_1, \zeta_2\}$	Y_2/%
0	-	-	{8, 0}	100.0	1	{8, 0}	0.0	{16, 5}	100.0
2	{16, 5}	87.0	{7, 0}	99.9	3	{7, 0}	0.0	{17, 5}	100.0
4	{17, 5}	84.9	{9, 0}	99.6	5	{9, 0}	0.0	{15, 5}	99.5
6	{15, 5}	88.5	{10, 0}	98.6	7	{10, 0}	0.0	{18, 5}	99.4
8	{18, 5}	82.3	{9, 1}	98.0	9	{9, 1}	0.0	{14, 5}	98.5
10	{14, 5}	89.4	{8, 1}	97.8	⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮	53	{14, 2}	0.0	{21, 8}	94.7
54	{21, 8}	76.2	{14, 4}	90.5	55	{14, 4}	0.0	{15, 7}	93.8
56	{15, 7}	85.8	{14, 1}	90.4	57	{14, 1}	0.0	{17, 8}	93.8
58	{17, 8}	83.3	{13, 5}	89.9	59	{13, 5}	97.0	{20, 9}	93.7
60	{13, 5}	89.9	{12, 5}	89.7					

with the cost of the second best offer {7, 0}, which has been found to be 99.9%. Since it is higher than 87.0%, {7, 0} is proposed by TSP-1 at round 3.

The corresponding cost of {7, 0} to TSP-2 is again zero. In fact, TSP-2 considers all counteroffers proposed by TSP-1 prior to round 59 are infeasible due to the release date constraint. As a result, TSP-2 continues suggesting its derived offers to TSP-1. On the other hand, even though the counteroffers received by TSP-1 are feasible, its potential offers are more favourable. Therefore, TSP-1 submits them to TSP-2. At round 59, TSP-1 proposes {13, 5}, which is the first feasible solution to TSP-2. In addition, since the cost (97%) is greater than O^* = {20, 9}, {13, 5} is accepted by TSP-2 in round 60. In round 61, TSP-1 confirms the acceptance by issuing an acknowledgment to TSP-2. According to the proof constructed in Section 7.2.2.2, such solution is Pareto-optimal.

7.5.1.2. Negotiation Pair (S_{min}, S_{po})

Despite the change from S_{po} to S_{min}, TSP-1 still proposes {8, 0} in round 1 as it is the best offer (Table 7.5). On the other hand, as TSP-2 employs the same strategy, its response remains unchanged. Upon the reception of {16, 5} in round 2, TSP-1

Table 7.5. Offers in Case 1 using (S_{min}, S_{po})

TSP-1							TSP-2				
Round	Counteroffer		O''		O'		Round	Counteroffer		O^*	
	$\{\zeta_1, \zeta_2\}$	Y_1/%	$\{\zeta_1, \zeta_2\}$	Y_1/%	$\{\zeta_1, \zeta_2\}$	Y_1/%		$\{\zeta_1, \zeta_2\}$	Y_2/%	$\{\zeta_1, \zeta_2\}$	Y_2/%
0	-	-	{8, 0}	100.0	-	-	1	{8, 0}	0.0	{16, 5}	100.0
2	{16, 5}	87.0	{8, 1}	97.8	{7, 0}	99.9	3	{7, 0}	0.0	{17, 5}	100.0
4	{17, 5}	84.9	{7, 1}	48.6	{9, 0}	99.6	5	{9, 0}	0.0	{15, 5}	99.5
6	{15, 5}	88.5	{9, 1}	98.0	{10, 0}	98.6	7	{10, 0}	0.0	{18, 5}	99.4
8	{18, 5}	82.3	{10, 1}	97.6	{11, 0}	97.1	9	{10, 1}	0.0	{14, 5}	98.5
10	{14, 5}	89.4	{10, 2}	95.9	{9, 1}	98.0	⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
22	{13, 5}	89.9	{10, 3}	93.8	{9, 2}	95.8	⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮	⋮	⋮	53	{14, 3}	0.0	{21, 8}	94.7
54	{21, 8}	76.2	{14, 2}	91.0	{15, 3}	89.0	55	{14, 2}	0.0	{15, 7}	93.8
56	{15, 7}	85.8	{14, 1}	90.4	{15, 2}	88.4	57	{14, 1}	0.0	{17, 8}	93.8
58	{17, 8}	83.3	{14, 5}	89.4	{15, 1}	87.3	59	{13, 5}	97.0	{20, 9}	93.7
60	{13, 5}	89.9	-	-	-	-					

generates $O'' = \{8, 1\}$ and $O' = \{7, 0\}$ from the most recently proposed offer and finds that Y' is greater than that of the counteroffer. As a consequence, O' becomes the potential offer and it is proposed to TSP-2 in round 3.

Since the offers submitted by TSP-1 during the first 58 rounds are all infeasible to TSP-2, TSP-2 rejects TSP-1 consistently. As the negotiation proceeds, the potential offer selected by TSP-1 is {9, 0} in round 4, {10, 0} in round 6, and so on. In round 22, TSP-1 receives {13, 5} from TSP-2 and determines that the cost (89.9%) is the highest among the counteroffers received. However, at that instance, the potential offers are still better than this solution, so the offer is not accepted yet. It is not until in round 58 that the cost of {14, 5} is smaller than 89.9%. TSP-1 then re-proposes {13, 5} to TSP-2 in round 59, and the solution is accepted by both parties.

7.5.1.3. Negotiation Pair (S_{max}, S_{po})

Similar to the negotiation using the pair (S_{min}, S_{po}), the first two offers of the agents are the same (Table 7.6). With the adoption of S_{max}, TSP-1 generates $O'' = \{16, 4\}$ and $O' = \{13, 5\}$ using the counteroffer {16, 5}. Since $Y' > Y''$, and it is also higher than the cost of the counteroffer, TSP-1 proposes O' to TSP-2. Such offer has a cost of 97.0% to TSP-2, which is lower than the cost of the second best offer {17, 5}. As a result, the offer {17, 5} is sent to the TSP-1.

The process iterates, where TSP-1 and TSP-2 propose alternately with offers {12,

Table 7.6. Offers in Case 1 using (S_{max}, S_{po})

TSP-1							TSP-2				
Round	Counteroffer		O''		O'		Round	Counteroffer		O^*	
	$\{\zeta_1, \zeta_2\}$	Y_1/%	$\{\zeta_1, \zeta_2\}$	Y_1/%	$\{\zeta_1, \zeta_2\}$	Y_1/%		$\{\zeta_1, \zeta_2\}$	Y_2/%	$\{\zeta_1, \zeta_2\}$	Y_2/%
0	-	-	{8, 0}	100.0	-	-	1	{8, 0}	0.0	{16, 5}	100.0
2	{16, 5}	87.0	{16, 4}	86.9	{13, 5}	89.9	3	{13, 5}	97.0	{17, 5}	100.0
4	{17, 5}	84.9	{17, 6}	84.9	{12, 5}	89.7	5	{12, 5}	97.0	{15, 5}	99.5
6	{15, 5}	88.5	{15, 4}	89.0	{14, 5}	89.4	7	{14, 5}	98.5	{18, 5}	99.4
8	{18, 5}	82.3	{18, 7}	82.9	{15, 5}	88.5	9	{15, 5}	99.5	{14, 5}	98.5

5}, {15, 5}, {14, 5}, {18, 5} and {15, 5}. In round 9, {15, 5} contributes a cost of 99.5% to TSP-2, which is higher than the cost of O^*={14, 5}, that is, 98.5%. TSP-2 therefore secures the agreement with {15, 5}. In this case, the number of negotiation rounds is substantially reduced, but the solution obtained is no longer Pareto-optimal.

7.5.2. Performances of Strategies

Table 7.7 shows the summary of results derived from the 10 scenarios. In each case, the result obtained by the negotiation pair (S_{po}, S_{po}) is considered as the reference solution, that is, the costs and the number of negotiation rounds are attributed to 100% while all other solutions derived from different negotiation pairs are computed against the reference.

Table 7.8 displays the frequency distribution of the solutions obtained by the negotiation pairs other than the reference case using (S_{po}, S_{po}). The first column describes the quality of solutions. Category (a) involves negotiations settling at the same agreement as the reference case (i.e. Pareto-optimal). Categories (b) and (c) contain cases reaching suboptimal solutions. In (b), the cost of one TSP is improved in the expense of the other one, while in (c), the costs of both TSPs are lower than that of the reference. Finally, cases in category (d) are terminated without making any agreement. The second column compares the negotiation rounds required. A (=) refers to the same number of rounds as the reference case. A (+) and (–) corresponds to needing more and fewer number of rounds respectively.

Table 7.7. Summary of Results

	S_1	S_{po}	S_{po}	S_{po}	S_{min}	S_{min}	S_{min}	S_{max}	S_{max}	S_{max}
Case	S_2	S_{po}	S_{min}	S_{max}	S_{po}	S_{min}	S_{max}	S_{po}	S_{min}	S_{max}
1	$\{\zeta_1, \zeta_2\}$	{13, 5}	{13, 5}	{13, 5}	{13, 5}	{13, 5}	{13, 5}	{15, 5}	{15, 5}	{14, 6}
	Y_1 /%	100.0	100.0	100.0	100.0	100.0	100.0	98.5	98.5	97.7
	Y_2 /%	100.0	100.0	100.0	100.0	100.0	100.0	102.6	102.6	98.4
	k /%	100.0	100.0	100.0	100.0	100.0	100.0	18.0	18.0	16.4
2	$\{\zeta_1, \zeta_2\}$	{8, 5}	{7, 5}	{7, 5}	{8, 5}	{7, 5}	{7, 5}	{11, 5}	{9, 5}	{7, 5}
	Y_1 /%	100.0	100.1	100.1	100.0	100.1	100.1	96.3	99.3	100.1
	Y_2 /%	100.0	96.4	96.4	100.0	96.4	96.4	107.4	103.0	96.4
	k /%	100.0	90.7	90.7	95.3	76.7	76.7	30.2	32.6	30.2
3	$\{\zeta_1, \zeta_2\}$	{60, 46}	{60, 46}	{60, 47}	{60, 46}	{60, 46}	{60, 47}	-	-	{60, 47}
	Y_1 /%	100.0	100.0	101.1	100.0	100.0	101.1	0	0	101.1
	Y_2 /%	100.0	100.0	98.5	100.0	100.0	98.5	0	0	98.5
	k /%	100.0	18.3	1.1	100.0	18.3	1.1	0.4	0.4	1.1
4	$\{\zeta_1, \zeta_2\}$	{12, 5}	{12, 5}	{12, 5}	{13, 5}	{12, 5}	{12, 5}	{14, 5}	{14, 5}	{7, 5}
	Y_1 /%	100.0	100.0	100.0	83.7	100.0	100.0	66.4	66.4	71.6
	Y_2 /%	100.0	100.0	100.0	102.2	100.0	100.0	103.8	103.8	35.2
	k /%	100.0	82.7	82.7	86.5	82.7	82.7	23.1	23.1	17.3
5	$\{\zeta_1, \zeta_2\}$	{16, 5}	{16, 5}	{16, 5}	{16, 5}	{16, 5}	{16, 5}	{16, 5}	{16, 5}	{16, 5}
	Y_1 /%	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	Y_2 /%	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	k /%	100.0	100.0	104.8	100.0	100.0	109.5	23.8	23.8	28.6
6	$\{\zeta_1, \zeta_2\}$	{12, 5}	{12, 5}	-	{12, 5}	{12, 5}	-	{12, 5}	{12, 5}	{13, 5}
	Y_1 /%	100.0	100.0	0	100.0	100.0	0	100.0	100.0	95.6
	Y_2 /%	100.0	100.0	0	100.0	100.0	0	100.0	100.0	100.2
	k /%	100.0	100.0	57.1	94.3	94.3	57.1	17.1	17.1	20.0
7	$\{\zeta_1, \zeta_2\}$	{12, 5}	{7, 5}	-	{12, 5}	{12, 5}	-	-	-	-
	Y_1 /%	100.0	333.2	0	100.0	100.0	0	0	0	0
	Y_2 /%	100.0	31.4	0	100.0	100.0	0	0	0	0
	k /%	100.0	58.7	26.7	57.3	57.3	26.7	14.7	14.7	8.0
8	$\{\zeta_1, \zeta_2\}$	{10, 0}	{10, 0}	{10, 0}	{10, 0}	{10, 0}	{10, 0}	{10, 0}	{10, 0}	{10, 0}
	Y_1 /%	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	Y_2 /%	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	k /%	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
9	$\{\zeta_1, \zeta_2\}$	{7, 13}	{7, 13}	{7, 14}	{7, 13}	{7, 13}	{7, 14}	{7, 13}	{7, 12}	{7, 14}
	Y_1 /%	100.0	100.0	101.1	100.0	100.0	101.1	100.0	98.4	101.1
	Y_2 /%	100.0	100.0	99.4	100.0	100.0	99.4	100.0	100.0	99.4
	k /%	100.0	81.8	10.2	100.0	81.8	10.2	100.0	81.8	10.2
10	$\{\zeta_1, \zeta_2\}$	{15, 5}	{15, 5}	{15, 5}	{15, 5}	{15, 5}	{15, 5}	{16, 5}	{16, 5}	{16, 5}
	Y_1 /%	100.0	100.0	100.0	100.0	100.0	100.0	99.9	99.9	99.9
	Y_2 /%	100.0	100.0	100.0	100.0	100.0	100.0	100.5	100.5	100.5
	k /%	100.0	100.0	100.0	69.8	69.8	69.8	13.2	13.2	24.5

Table 7.8. Frequency Distribution of Solutions

Cat.	Rnd.	(S_{min},S_{min})	(S_{po},S_{min})	(S_{min},S_{po})	(S_{min},S_{max})	(S_{max},S_{min})	(S_{po},S_{max})	(S_{max},S_{po})	(S_{max},S_{max})
(a)	(=)	3	5	5	2	1	3	2	1
(a)	(–)	6	3	4	2	2	1	2	1
(a)	(+)	0	0	0	1	0	1	0	0
(b)	(–)	1	2	1	3	5	3	4	5
(c)	(–)	0	0	0	0	0	0	0	2
(d)	(–)	0	0	0	2	2	2	2	1
Total		10	10	10	10	10	10	10	10

Note:
Cat. (a): solutions are identical to those obtained by (S_{po}, S_{po}) (i.e. Pareto-optimal)
Cat. (b): solutions are suboptimal (only one agent cost is lower than the Pareto-optimal solution)
Cat. (c): solutions are suboptimal (both agent costs are lower than the Pareto-optimal solution)
Cat. (d): no solutions (negotiation is terminated without reaching an agreement)
Rnd (=): negotiation requires equal number of rounds as in (S_{po}, S_{po})
Rnd (+): negotiation requires more number of rounds as in (S_{po}, S_{po})
Rnd (–): negotiation requires less number of rounds as in (S_{po}, S_{po})

7.5.2.1. Quality of Solutions

i) *Strategy-PO*: According to Table 7.8, there are no instances such that both TSPs achieve higher costs at the same time. Pareto-optimality obtained by (S_{po}, S_{po}) is thus supported by the simulation results. Nevertheless, the existence of category (a) implies that the involvement of S_{min} and S_{max} may still lead to the Pareto-optimal solution. Despite settling at the same solution, the solution paths (i.e. the sequence of proposed offers) are usually different when the negotiation involves other strategy. In fact, when using S_{po}, the cost of a TSP is always monotonically decreasing by definition. On the other hand, when S_{min} or S_{max} is employed, the costs are often rippling downwards (Fig. 7.12).

Table 7.9 compares the Pareto-optimal solutions between the 10 scenarios. In

Table 7.9. Comparison of Pareto-optimal Solutions

Case	1	2	3	4	5	6	7	8	9	10
$\{\zeta_1, \zeta_2\}$	{13, 5}	{8, 5}	{60, 46}	{12, 5}	{16, 5}	{12, 5}	{12, 5}	{10, 0}	{7, 13}	{15, 5}
(A_1, D_1)	(33, 38)	(28, 29)	(80, 85)	(32, 37)	(36, 41)	(32, 37)	(32, 37)	(30, 35)	(47, 52)	(35, 40)
(A_2, D_2)	(35, 42)	(35, 36)	(76, 83)	(35, 42)	(35, 42)	(35, 42)	(35, 42)	(30, 37)	(43, 50)	(35, 42)
w_{12} /min	7	6	1	8	4	8	8	5	1	5
w_{21} /min	1	-	7	0	4	0	0	3	7	3
$Y_1 /\$$	2213	1266	2549	1210	2583	1076	26	2079	2549	8198
$Y_2 /\$$	3686	2151	1279	3608	3801	1945	405	3783	3259	3783

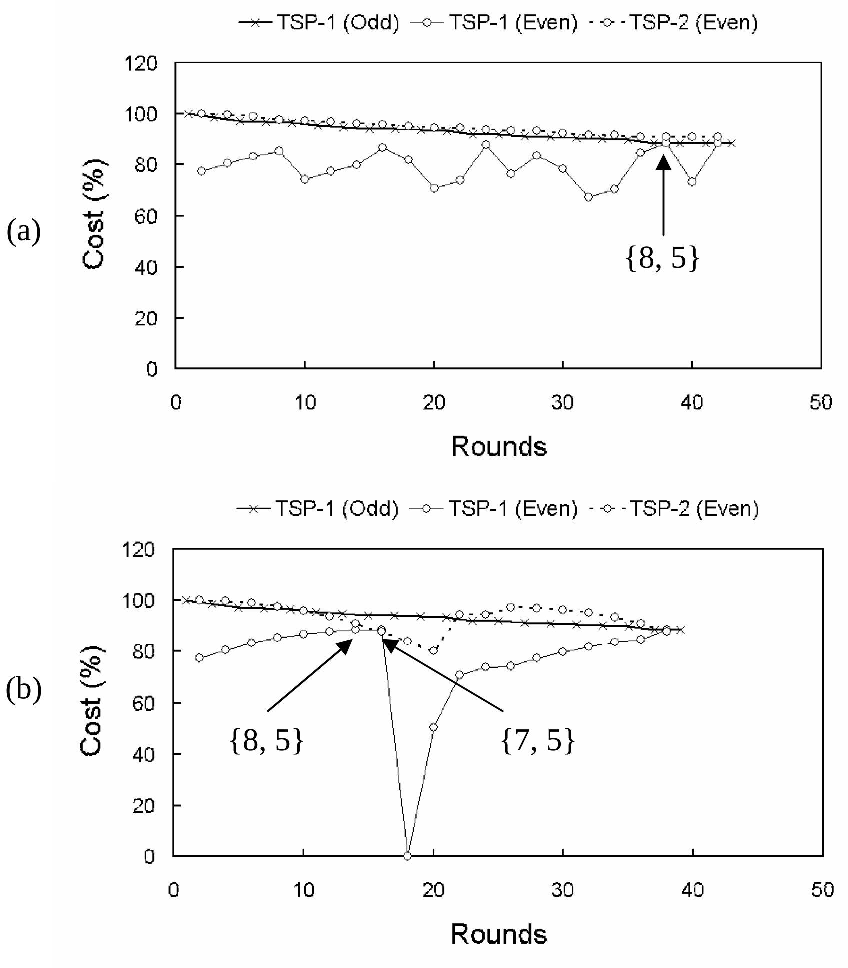


Fig. 7.12. Concession curves in case 2 (a) (S_{po}, S_{po}) (b) (S_{min}, S_{po})

each of the coordinated transfer, the passenger waiting time is maintained reasonably at less than 10 mins. In case 1, although the agents begin the negotiation with unidirectional transfer, the agents are able to settle at the more favourable bidirectional transfer. This is achieved by postponing the commencement time of L_1 from 7 mins to 13 mins. Although the Pareto-optimal costs obtained by the agents are both less than their corresponding optimal values, as mentioned in Section 7.2.2, these optimal solutions are either infeasible or causing a loss to the proponent. Achieving Pareto-optimality through S_{po} therefore provides a compromise between the objectives of the two parties.

In case 2, owing to the short dwell times for both TSPs, the negotiation can only enable unidirectional transfer from L_1 to L_2. As a result, the costs obtained by the two agents are significantly less than that in case 1.

In case 3, having increased the release date of TSP-1, it is now TSP-2 which changes its commencement time (from 5 mins to 46 mins). With such a large alteration, there is a high idle cost of rolling stock to offset the revenue collected by TSP-2.

Using different idle cost rates, the solutions obtained in cases 4 and 5 are different. As case 4 employs a high rate, TSP-1 is not willing to postpone the commencement time as late as in case 1. This indirectly increases the passenger waiting time from L_1 to L_2 by 1 min. On the contrary, since delaying the service only causes a slight rise in idle cost, TSP-1 has more capacity for making concession in case 5. At the end, TSP-1 agrees on starting the service at 16 mins, which not only is the Pareto-optimal solution for both agents, but also is the optimal solution for TSP-2.

In cases 6 and 7, the transfer demands are reduced. Both cases reach the same agreement, but their costs are significantly lower than case 1. In case 7, where the demands for both directions are low, the costs approach zero. In practice, if the simulation is used by the TSPs to evaluate whether the schedule coordination is beneficial, they may decide not to conduct the negotiation to avoid paying the transaction costs (e.g. costs of manpower and preparations of contractual documents) involved in the negotiation.

In case 8, when there are no release date constraints, both TSPs can start operating their services at an earlier time. In case 9, the effect of doubling the journey time of L_1 causes TSP-2 to postpone its commencement time to 13 mins. In case 10, when the passenger transfer charge rate is increased, the revenue intake of TSP-1 is increased accordingly. Moreover, for such a substantial gain, TSP-1 is willing to concede more on commencement time to 15 mins. Overall, the simulated results and agent behaviour are logical.

ii) *Strategy-MIN*: For negotiations employing S_{min}, the majority of negotiations are contained in category (a) while the remaining ones are mainly captured by category (b) (Table 7.8).

Comparing S_{po} and $S_{min,}$ the relaxation on forming a monotonic decreasing sequence of offers leads to the possible settlement on suboptimal agreements. In the operation of S_{min}, since the potential offers are generated by holding either ζ_1^{k-1} or ζ_2^{k-1} constant (the most recent offer proposed by the agent), the agent is now only able to search within a limited set of offers in each round. This contrasts to the operation of S_{po} which is capable of selecting the next best offer from the entire solution space. Owing to this restriction, S_{min} has the risk of proposing (or revealing) a less favourable solution during the negotiation. For example in case 2 (Fig. 7.12), a less favourable offer {7, 5} contributes a lower cost to TSP-2 (employing S_{min}), but the cost of its proponent is higher. As a result, the TSP-1 prefers (and accepts) the suboptimal offer rather than the Pareto-optimal one {8, 5}.

Nevertheless, the frequency of reaching a suboptimal offer (categories (b) and (c)) is not exceedingly high when compared with that associated with category (a). In addition, even if the negotiation ends with a suboptimal offer, the quality of solution is usually very close to that of the reference solution. In this aspect, S_{min} seems to be capable of approximating the operation of S_{po} in most scenarios. S_{min} can therefore be a good alternative to S_{po}.

iii) *Strategy-MAX*: According to Table 7.8, most negotiations fall in categories (a) and (b). There are only two negotiations in cases 1 and 4 whose agreements belong to category (c) and the negotiation pair is both (S_{max}, S_{max}). Moreover, there are several negotiations (in cases 3, 6 and 7) leading to solutions in category (d). In all these failed negotiations, S_{max} is employed by at least one of the agents. Therefore, the results suggest that S_{max} is less favourable than S_{po} and S_{min}, and it also has a higher risk of terminating the negotiations either with suboptimal solutions or without any agreements at all.

The reason for S_{max} leading to suboptimal solutions is identical to S_{min}. As both strategies employ similar operations in deriving their next potential offers (i.e. holding ζ_1 or ζ_2 constant), both S_{min} and S_{max} can only select the offer from a limited set of choices. This leads to a higher risk of proposing solutions that favours the proponent in the expense of the agent's benefit. Despite the similarity, there are now fewer negotiations of S_{max} leading to category (a). Since the only difference between the two strategies is the choice of O^{k-1} and O^k, it seems that the use of the proponent's offer has led to the observed results. In fact, since O^k is often generated from the proponent's benefit, the quality of the potential offer from such 'seed' is likely to be less favourable than that generated from O^{k-1}, which takes more consideration to the agent's advantage.

However, when using the negotiation pair (S_{max}, S_{max}), both agents may suffer from a reduction in cost because they are both manipulating the proponent's offer to generate their counteroffers. In other words, neither agent is consistently benefiting from the operation. Without any logical modification of the counteroffers, the final agreement may eventually be unfavourable to both parties.

On the other hand, termination with failures can be explained by Fig. 7.13. In the

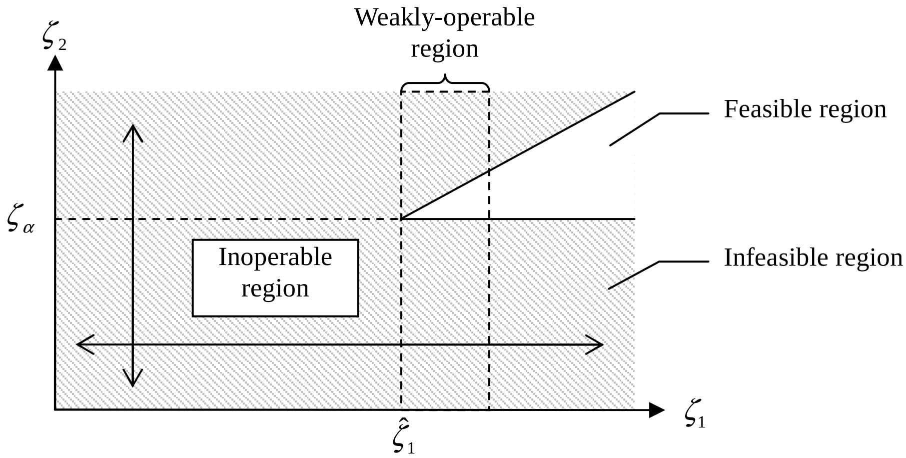


Fig. 7.13. Inoperable and weakly-operable regions of S_{max}

figure, $\hat{\zeta}_1$ represents the release date of the agent employing S_{max}. The corresponding earliest commencement time of the proponent's service is denoted by ζ_α. The inoperable region of S_{max} is defined by $\zeta_1 < \hat{\zeta}_1$ and $\zeta_2 < \zeta_\alpha$. If a counteroffer is proposed in this region, then by the definition of S_{max}, the strategy can never produce a feasible solution. The weakly-operable region is a band of solutions close to the $\hat{\zeta}_1$. Despite having a few feasible solutions, if the proponent is frequently proposing offers in this band of solutions, the feasible offers are quickly exhausted. Inspection of the progress of the negotiations confirms that the proponent's counteroffer is situated in the inoperable region in case 3 and in the weakly-operable region in cases 6 and 7.

As a result, unfavourable termination is likely to occur when the initial plans of the two agents (i.e. schedules departing from the release dates) are remotely compatible. In such case, the agent with the earlier commencement time will often propose offers in the inoperable or weakly operable region. In addition, termination can also occur when there are low passenger demands to recover the high idle cost on the rolling stock.

7.5.2.2. Length of Negotiation

i) *Strategy-PO*: Despite the observed benefit of S_{po} over S_{min}, and S_{po}, S_{min} over S_{max} regarding to the quality of solution, the advantage on fewer negotiation rounds favours in the reverse order. According to Table 7.7, negotiations employing S_{po} often require a substantial amount of rounds before the negotiation is settled. On the other hand, S_{min} may sometimes reduce the number of rounds, and S_{max} is even more likely to reduce the number considerably.

ii) *Strategy-MIN*: With the restriction in generating the potential offers, S_{min} may sometimes complete the transaction with fewer negotiation rounds. Since there is no need to propose the offers in decreasing order of costs, S_{min} may be able to skip some of the intermediate solutions while still being able to reach the Pareto-optimal or a suboptimal agreement. However, in many cases, S_{min} requires the same number of rounds as the reference negotiation. In the worst scenario, as demonstrated by one of the negotiation accompanying S_{max}, the transaction requires more rounds than (S_{po}, S_{po}). This is explained in the later discussion.

iii) *Strategy-MAX*: The reason that S_{max} requires fewer negotiation rounds resides to its method of generating the counteroffers. Given the most recent proponent's offer O^k, S_{max} first derives two potential offers O' and O'' by maximising the cost difference from the proponent's offer.

In deriving O', ζ_j is kept constant (equals to ζ_j^k). Since the value of ζ_j originates from the proponent's offer, O' is always a feasible offer to the proponent because it causes no violation to the release date constraint. As feasibility implies non-zero cost, the proponent is more likely to accept the counteroffer. Hence, the number of negotiation rounds may be lowered when O' is selected as counteroffer to the proponent.

However, there are exceptions in case 5, when (S_{po}, S_{max}) and (S_{min}, S_{max}) are employed. In these cases, S_{max} requires more negotiation rounds than the reference solution obtained by the pair (S_{po}, S_{po}). The reason can be explained in two-fold. Firstly, although the release date constraint of TSP-1 is $\zeta_1 \geq 7$, during the negotiation in the reference case, solutions that are outside the feasible range have not been visited. In other words, the advantage of generating feasible solutions by S_{max} is not applicable in this case. Secondly, the restriction of S_{max} (and S_{min}) to produce offers either vertically or horizontally from the counteroffer leads to a slight increase in the number of intermediate offers before the final solution {16, 5} is found (Fig. 7.14). As there is no such restriction for S_{po} (i.e. offers can be generated by moving diagonally), the

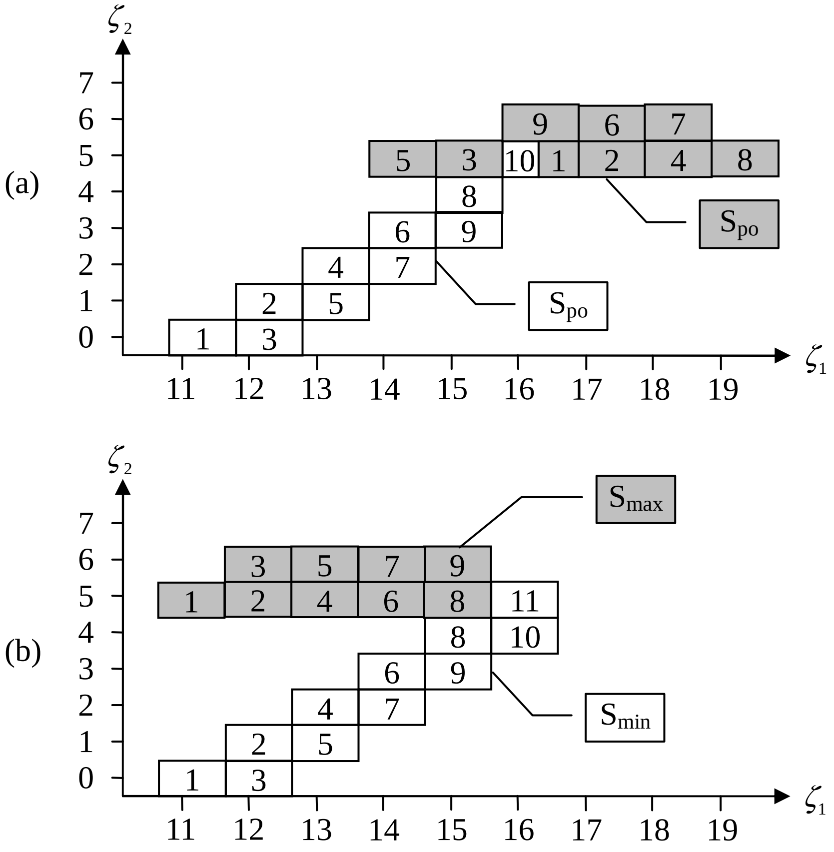


Fig. 7.14. Solution paths in case 5 (a) (S_{po}, S_{po}); (b) (S_{min}, S_{max})

negotiation can be completed with fewer negotiation rounds.

7.6. Remarks

This chapter has presented an agent model for the schedule coordination problem involving two TSP agents (the TSP-TSP transaction). The transaction employs a simple protocol allowing the agents to propose, accept or reject offers. The TSP agents are also able to incorporate one of the negotiation strategies S_{po}, S_{min} and S_{max}, which are devised to allow the agents to exhibit different behaviour. In order to propose offers during the negotiation, the agents are required to solve a quadratic programming problem and an offer generation problem. The former is solved using the Lemke's Complementary Pivoting Algorithm (LCPA) while the latter is resolved by devising a pruning tree searching algorithm. Through the agent negotiation process, the TSPs are enabled to decide whether coordinating the schedules between the train services is favourable.

Simulations have been conducted to evaluate the performances of the quality of the resultant solutions and the length of negotiation under a set of extreme scenarios. The findings confirm that S_{po} guarantees a Pareto-optimal agreement but requires the highest negotiation demand. S_{min} improves the transaction speed but occasionally suffers from reaching a suboptimal solution only. S_{max} is a fast means to complete a negotiation but it has a higher risk of termination without striking a deal.

The trade-off between optimality and computation speed varies with different stakeholders and it is up to the train planners to decide which strategies should be employed. However, in general, negotiation activities often have deadlines before which the transactions should be completed. As a result, train planners may decide to employ S_{po} when the negotiation is likely to finish before the deadline and use S_{min} or

S_{max} otherwise. It is also possible to begin the negotiation with S_{po} and switch to the other strategies when negotiation time seems inadequate. The incorporation of the BDI-model to aid such dynamic decision-making is certainly worth exploring in further works.

The study also generates further research opportunities. Firstly, the study here only examines the negotiation between two TSPs, but it does not consider the effect of IP on granting the required track and station capacity. Further works can be conducted to investigate the inclusion of the IP agent in the negotiation. Secondly, the model can be expanded to consider that the two TSPs are operating a set of regular services on different headways. For example, TSP-1 is operating its service on an hourly basis, while TSP-2 is dispatching a train every 20 minutes. In such case, the objective functions will then be more complicated, which may consider minimising the average passenger waiting time. Moreover, it is also possible to investigate the schedule coordination problems involving more than two TSPs (i.e. coordination three or more train schedules), which is likely to occur at large interchanging stations such as in London Euston, Manchester Piccadilly and Liverpool Lime Street.

Chapter 8

Applicability of MAS-ORAM in Real World Railways

Despite the substantial number of simulation cases in the previous three chapters, none of them was set up against the background of an existing open access market. As a result, this chapter is devoted to demonstrate how the developed simulation software may be employed as a tool for planning and evaluation in practice. However, it should be clearly stated that the simulation data used in this study is hypothetically created and not collected from any official organisations. Thus, the results demonstrate the applicability of the simulation models but not necessarily reflect any current situation. Through the application of the simulation software against a practical background, it is anticipated the readers may appreciate the potential benefits of the proposed MAS-ORAM.

This chapter is organised as follows. Some of the resource management problems occurring in the railway markets in Australia, the UK and China are first identified. Although the Chinese railway is currently not open to competition, the possibility of employing the simulation tool is discussed. According to the observed problems in the Australian and British railway markets, two special case studies are constructed and examined on MAS-ORAM.

8.1. Railway Markets in Selected Countries

8.1.1. Australia

Australia is the sixth largest country in the world covering a total area of 7.7 billion km^2 of land. It is a sparsely populated country where the majority of the population resides in the state capital cities (i.e. Sydney, Melbourne, Brisbane, etc.). As a result of the low population density (2.65/km^2) and long distances between the capital cities, Australia has a weak demand for passenger rail services.

Despite the unfavourable conditions for passenger railways, the abundance of natural resources and the necessary long distance transportation have led to a promising freight railway market. In particular, Australia has a large volume of coal deposits which are mostly located at the states of Queensland and New South Wales. In 2004, the annual coal output was 298 million tonnes (ABS, 2006), and it has been expected to increase by year in response to the growing international demand on coal. Being a solid bulk commodity, the large volume of coal cannot be transported to the exporting ports on the coast via pipelines, nor is it economical to be moved by road or air over long distances. Consequently, there is a high demand for coal transportation by railways in Australia.

Currently, the Australian interstate railway is open to competition. The infrastructure is managed by the Australian Rail Track Corporation (ARTC) which was formed in 1998 by the Commonwealth and State Governments to provide a ‘one-stop-shop’ for train operators to obtain capacity on the interstate rail network. As the Australians allow the freedom of local juridical control in individual States, the tracks under the management of ARTC are either self-owned or leased from the State Governments (Fig. 8.1).

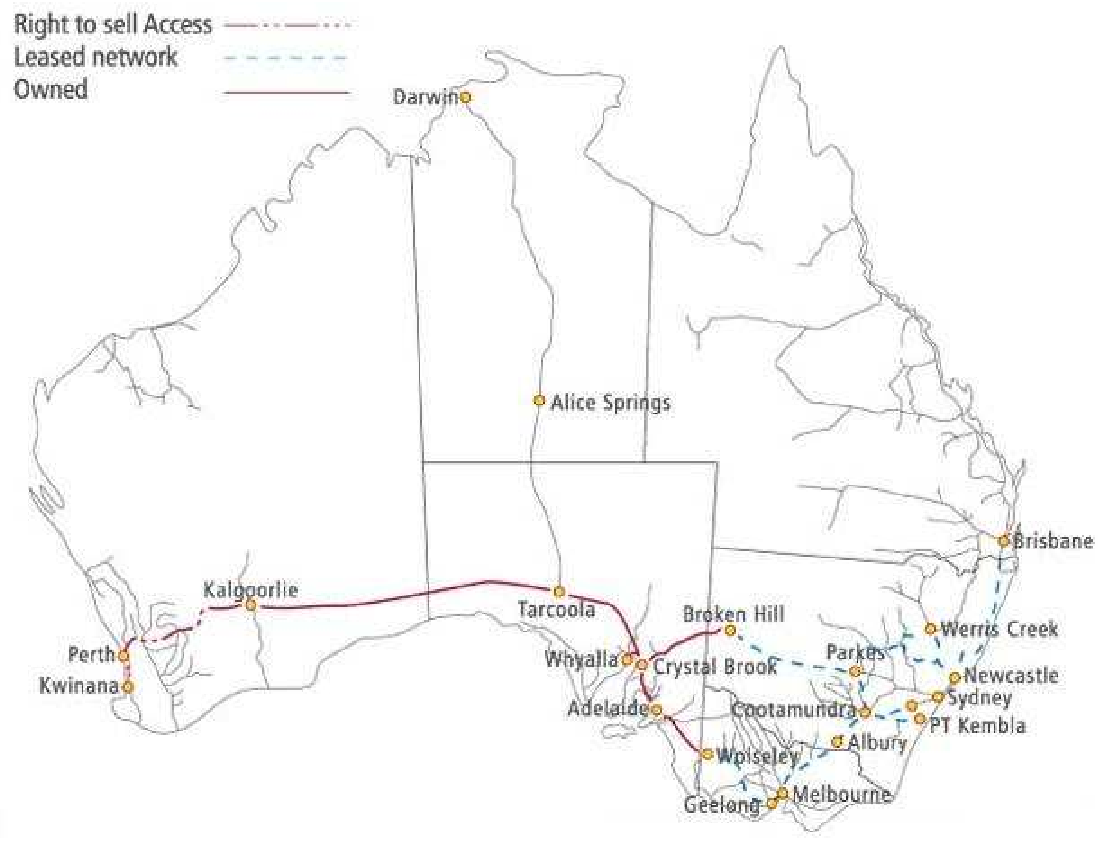


Fig. 8.1. Rail Network of the Australian Rail Track Corporation
Source: ARTC [Australian Rail Track Corporation] (2006) ‘Rail Network’. http://www.artc.com.au, accessed June 2006

There are nine major train operators seeking access on the ARTC network (ARTC, 2006). These include three passenger operators (CityRail, CountryLink, and Great Southern Railway) and six freight operators (Australian Southern Railroad, FreightLink, Pacific National, Patrick Rail Operations, Queensland Rail National, and Specialised Container Transport). However, not all of these operators are competing directly for customers. For example, FreightLink is specialised in freight transportation between Adelaide and Darwin while the other freight service providers are mainly operating in the Perth-Melbourne and the Perth-Sydney corridors.

Among the tracks under the control of ARTC, the Hunter Valley rail network (Fig. 8.2) in New South Wales is one of the railways experiencing strong on-rail competition. The network is composed of 452 km of tracks and it has a maximum speed limit of 60

Fig. 8.2. Coal transportation in the Hunter Valley

Source: QR National (2006) 'Coal Map'. http://www.qrnational.com.au, accessed June 2006

kph (ARTC, 2004). Along this railway line, coal is transported from the coalfields in the Hunter Valley to the port of Newcastle by two main competitors, Pacific National and Queensland Rail National (QR National). Unlike the freight transportation elsewhere in Australia, the haul distance in the Hunter Valley is relatively short (20-320 km) while the haul volume remains large (Pacific National, 2006). For example, in 2004, the annual coal export at Newcastle recorded 78 million tonnes and it has been forecasted to grow at a rate of 2.8% per year (ABARE, 2005).

Although the dominated traffic in the Hunter Valley is coal transportation, the track capacity is still shared with passenger and cargo services, both of which receive higher priority on track access than the coal trains. This has by no means hindered the growth of the industry. Moreover, further increase in traffic volume is limited by the capacity supported by the existing infrastructure. As a result, several construction plans have been devised to expand the rail capacity at the Hunter Valley. One of which is to increase the annual coal transportation capacity to 102 million tonnes by 2008 (ABARE, 2005). This will be mainly achieved by both upgrading the current infrastructure to increase the maximum speed limit to 80 kph and eliminating the identified bottlenecks (ARTC, 2004). Nevertheless, efficient allocation of capacity among the train operators is still an important issue before the additional capacity is made available. Moreover, as capacity cannot be increased indefinitely, capacity allocation is ultimately a long-term problem which requires much attention.

8.1.2. United Kingdom

In contrast to Australia, the geographic and population conditions in Britain are more favourable to passenger rail services. With only 224 thousand km^2 of land and a population of 60 millions (in 2005), the population density (243/km^2) in the UK is much higher than that in Australia. In addition, the majority of the population dwells in the major cities (i.e. London, Birmingham, Manchester, Leeds, etc.) with a gradual decrease in population density towards the outskirts. Such population distribution has encouraged the development of regional and intercity services in the UK. On the other hand, since natural resources such as coal are usually consumed within the country and the transportation distance is relatively short, the movement of freight commodities in the UK relies mainly on road transportation.

The national railway in Britain was privatised and it became an open access market in 1994 (ECMT, 1998). Network Rail is the current infrastructure provider and the rail network under its ownership is shown in Fig. 8.3. There are 25 franchised passenger train operating companies (i.e. TSPs) seeking access to this network. In addition to delivering train services to the public, these operators are also responsible for managing a total of 2500 local train stations (Network Rail, 2006). However, there remain 17 major stations under the management of Network Rail and a list of which is given in

Fig. 8.3. Railway network in the UK

Source: National Rail Enquiries (2006) ‘National Rail Network Maps’, Doe, B.S. http://www.nationalrail.co.uk, accessed June 2006

Table 8.1 (Network Rail, 2006; Wikipedia, 2006). These stations are the busiest in the UK where trains operated by different TSPs arrive to facilitate intra-modal transfer of passenger services.

An example is the railway station at Liverpool Lime Street at which there are five TSPs operating trains at this station. Intercity services are provided by TransPennine Express and Virgin Trains while regional services are offered by Central Trains, Northern Rail and Merseyrail. A simplified schematic for the lines of these operators (excluding Merseyrail which operates in the west of Liverpool) are shown in Fig. 8.4 (Central Trains, 2006; Northern Rail, 2006; TransPennine Express, 2006; Virgin Trains, 2006). TransPennine Express is competing with Virgins Trains in the northern

Table 8.1. List of Passenger Stations Operated by Network Rail in the UK

Station	Train operating company in service		
Birmingham New Street	Arriva Trains Wales	Central Trains	Virgin Trains
Edinburgh Waverley	First ScotRail	GNER	Virgin Trains
Gatwick Airport	First Capital Connect	First Great Western	Gatwick Express
	Southeastern	Southern	Virgin Trains
Glasgow Central	First ScotRail	GNER	Virgin Trains
Leeds	GNER	Midland Mainline	Northern Rail
	TransPennine Express	Virgin Trains	
Liverpool Lime Street	Central Trains	Merseyrail	Northern Rail
	TransPennine Express	Virgin Trains	
London Bridge	First Capital Connect	Southeastern	Southern
London Cannon Street	Southeastern		
London Charing Cross	Southeastern	Southern	
London Euston	First ScotRail	Silverlink	Virgin Trains
London Fenchurch Street	c2c		
London King’s Cross	First Capital Connect	GNER	Hull Trains
London Liverpool Street	one		
London Paddington	First Great Western	Heathrow Connect	Heathrow Express
London Victoria	Gatwick Express	Southeastern	Southern
London Waterloo	Eurostar	South West Trains	
Manchester Piccadilly	Arriva Trains Wales	Central Trains	Northern Rail
	TransPennine Express	Virgin Trains	

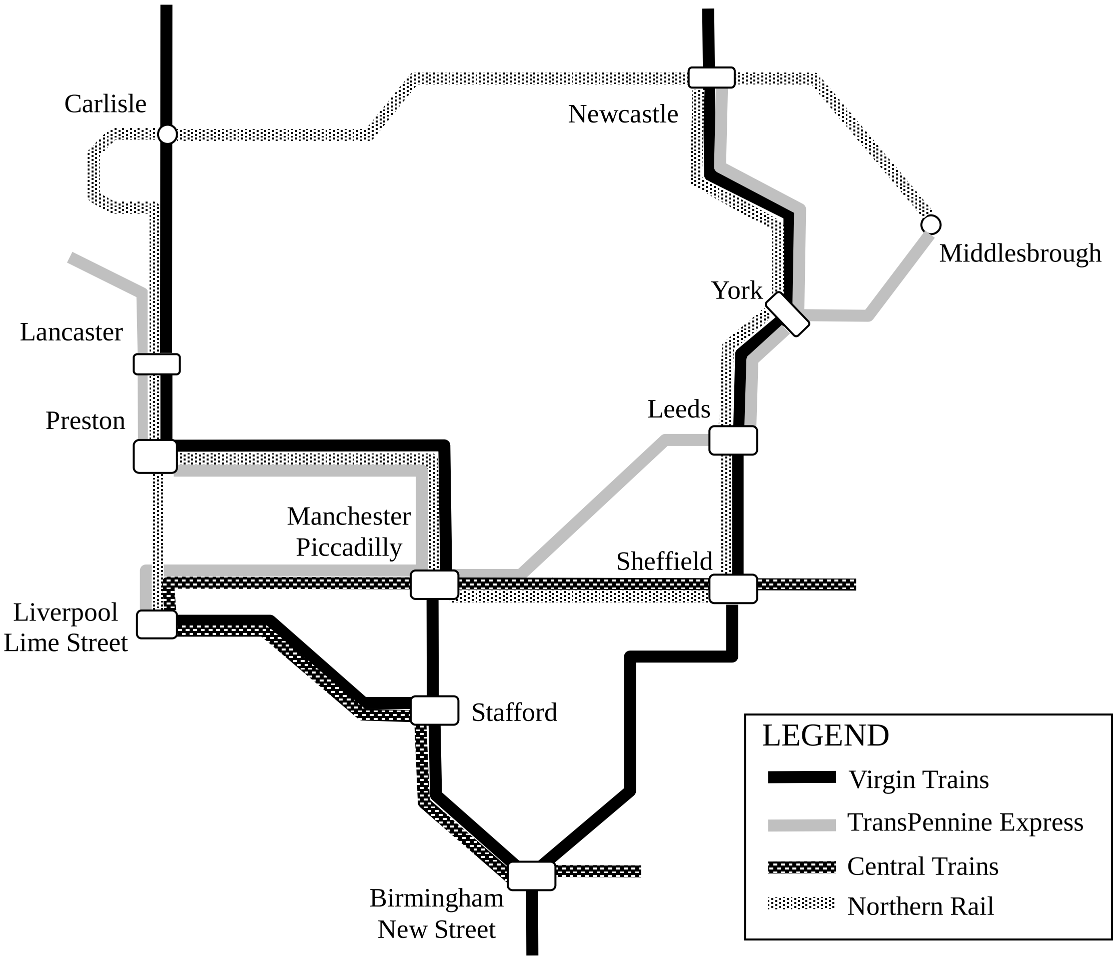


Fig. 8.4. Schematic diagram for major railway lines of four TSPs in the UK

England including major cities at Lancaster, Preston, Liverpool, Manchester, Sheffield, Leeds, York and Newcastle. On the other hand, Central Trains and Northern Rail have only limited competition at the Liverpool-Manchester-Sheffield corridor as they operate their services separately in the south and north of the corridor respectively.

As a consequence, the two regional service operators may consider coordinating their schedules to attract an additional demand for cross-regional services. This would create a yardstick competition with the seamless intercity services. For example, the journey from Preston to Birmingham via Virgin Trains takes about 1 hour 40 minutes while the trips from Preston to Liverpool via Northern Rail and Liverpool to Birmingham via Central Trains are approximately 1 hour and 1 hour 45 minutes respectively. In other words, the minimum journey time for the coordinated service is 2 hour 45 minutes, which is just about one hour longer than the intercity service. If the combined train fares for the regional services are lower than the intercity one, and the passenger waiting time is kept reasonably short, it is possible that some passengers will use the coordinated service instead of the seamless one.

8.1.3. People's Republic of China

The railway network in China is one of the largest in the world in terms of route length (about 72,000 km in 2005). The main network is illustrated in Fig. 8.5. Similar to Australia, China is a large country with abundance of natural resources. Coal and mineral deposits are found mainly in the northwest and transported both eastwards and southwards (Xue *et al.*, 2002a). On the other hand, since fertile agricultural lands are concentrated in the south, grains are mostly carried to the north (Xue *et al.*, 2002a). The uneven distribution of these bulk commodities has led to favourable conditions for freight transportation for the Chinese railways. However,

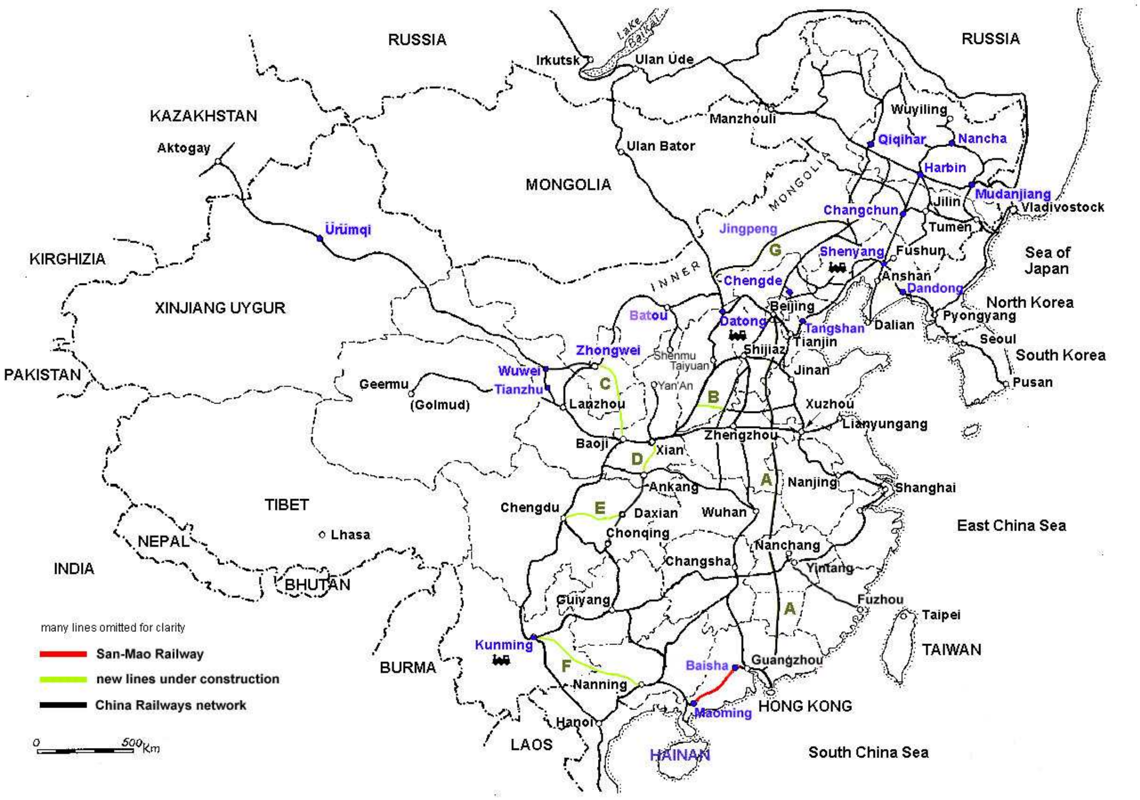


Fig. 8.5. Chinese Railways

Source: Gibbons, R.J. (2005) 'Railways of China'. http://www.railwaysofchina.com, accessed June 2006

unlike Australia, the passenger railways in China are in direct competition to road and air transportation. With a high population of 1.3 billion (in 2005), a low car ownership and a low Gross National Product (GNP) per capita, there is also a high demand for intercity and inter-regional passenger rail services (Wu & Nash, 2000).

In the past, the railways in China were owned entirely by the Central Government and were under the complete management of the Chinese Ministry of Railways (MOR). Being an agency that was responsible for both implementing railway policies for the government and managing the operations of the railway network, MOR had poor performance in providing market-oriented rail services. For example, as a result of capacity constraints, the rigid regulatory regime forbade freight transportation of less than 100 km by railways and the delivery time of goods was not provided to the consigners (Xie *et al.*, 2002). The failure to respond to the market demand has led to a progressive loss in market share to road transportation.

In response to the modal shift to road transportation, reformative acts have been conducted since the 1990s. The management and operation of the railways were first decentralised to 12 Regional Railway Administrations (Wu & Nash, 2002; Xue *et al.*, 2002b) and the number is currently increased to 18. The transfer of duties to these administrations has encouraged better planning and development of the local railways. In addition, there has been deregulation in central ownership through the possibility of constructing new infrastructure and services by joint ventures with foreign investors (Wu & Nash, 2002; Xie *et al.*, 2002). Although the central government is still the main shareholder, deregulation in ownership has raised funds and consequently expanded the rail capacity for the industry to improve the quality of railway services in China.

Although the reform has not led to an open railway access market yet, some small-scaled competition activities have been observed in China. For example, there is a limited yardstick competition between the national railways and the local ones constructed by joint ventures, where trains are operated on different infrastructures (Wu & Nash, 2002). Also, several joint ventures are allowed to operate a limited number of passenger and freight services on the national network according to the availability of line capacity (i.e. third-party access) (Wu & Nash, 2002). As a result, the proposed MAS-ORAM may be employed in this third-party access market to develop a proper access charge regime for these external operators. In addition, by conducting extensive studies for these local competition activities, insights may be drawn for the costs and benefits for introducing further competition in the Chinese railways.

8.2. Feasibility Studies

In the remaining part of this chapter, the capacity allocation problem at the Hunter Valley and the schedule coordination problem at the Liverpool Lime Street station are examined using the proposed MAS-ORAM in order to demonstrate its applicability on planning and evaluation in railways.

8.2.1. Allocation of Track Access Rights at the Hunter Valley

8.2.1.1. Problem Description and Simulation Setup

The problem on allocating track access rights by ARTC at the Hunter Valley is studied using the IP-TSP^n negotiation model developed in Chapter 6. The study presented here mainly aims to assist the train planners in ARTC to determine whether a proposed change in scheduling practice is beneficial.

Fig. 8.6 shows the schematic of a section of track between Muswellbrook and Maitland at the Hunter Valley. There are four intermediate stations and the total track length between Muswellbrook and Maitland is 120 km. The individual track lengths between the stations are summarised in Table 8.2. On this section of track, the traffic distribution is shown in Fig. 8.7. A daily intercity passenger service is operated by CountryLink from Werris Creek to Sydney (i.e. CL-01), while four regional services are provided by CityRail per day from Muswellbook to Newcastle (i.e. CR-01 to CR-04).

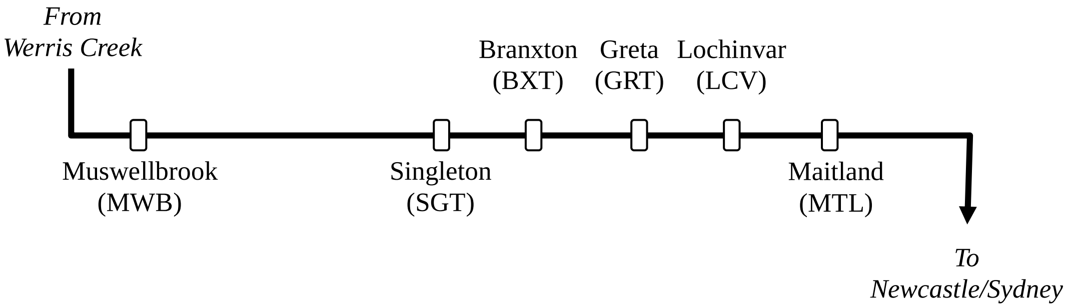


Fig. 8.6. Schematic diagram for a section of the Hunter Valley railway network

Table 8.2. Track and Station Data

Origin	Destination	Track Length (km)	Track Specific Constant (min)
MWB	SGT	60	1.0
SGT	BXT	27	1.0
BXT	GRT	6	1.0
GRT	LCV	12	1.0
LCV	MTL	15	1.0

Table 8.3. Definition of Rolling Stock

Type	Haul volume (ton)	Track usage rate c_1^{ω} (\$/ton·km)
ω_1	6500	0.005
ω_2	5500	0.004

Table 8.4. Definition of Flex Levels

Level	Flex time (min)	Discount factor
ϕ_1	0	1.00
ϕ_2	5	0.90
ϕ_3	10	0.80

For simplicity, the traffic in the opposite direction is neglected in the study. As mentioned, the passenger services have priority of access over the freight services provided by Pacific National and QR National.

It is assumed that the two freight operators have to compete for the remaining track capacity. There are two types of rolling stock (ω_1 and ω_2) available for selection (Table 8.3). ω_1 has a haul volume of 6500 tonnes and a track usage rate of \$0.005/ton·km. ω_2 has a lower haul volume of 5500 tonnes and a track usage rate of \$0.004/ton·km. It should be noted that dollar per tonne-kilometre is used instead of dollar per vehicle-kilometre when considering freight rail services.

In addition, there are three choices of flex levels as shown in Table. 8.4. ϕ_1 has a discount factor of 1.00 since it does not allow any adjustment in train schedule. When the freight operators accept ϕ_2, which allows a flex time of 5 minutes, the train service will be granted a 10% discount. ϕ_3 has a flex time of 10 minutes and thus has a higher discount of 20%.

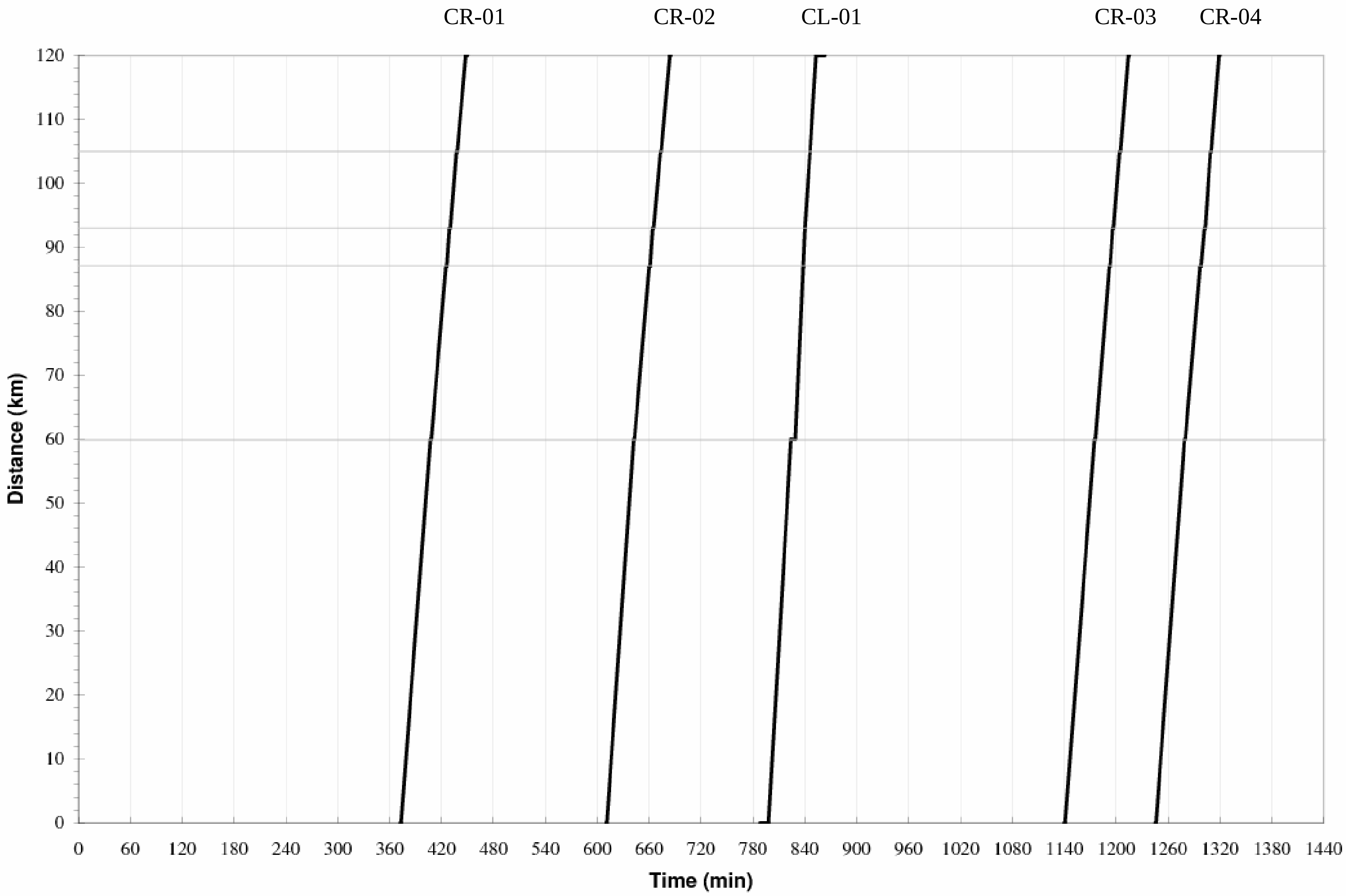


Fig. 8.7. Traffic distribution for passenger services between Muswellbrook and Maitland

The settings of ARTC in the simulation are shown in Table 8.5. As the freight trains are assumed to be powered by diesel engines, the rates for electricity (c_2) and peak demand (c_3) are set to zero. A high capacity weighting (w_η) of $8000 is used to differentiate track access rights of different capacity utilisation. The congestion charge rate (c_4) is set as $300/min. The adopted sequencing policy is FCFS.

With the above operating conditions, ARTC is investigating the consequence of the traffic distribution and capacity utilisation of two scheduling practices. In the first scenario, which supposedly reflects the normal situation, the freight operators are allowed to request capacity over the entire 24-hour interval. In such case, based on the previous scheduling experience, the desired service headways for Pacific National and QR National are assumed to be 60 minutes (i.e. 24 trains/day) and 72 minutes (i.e. 20 trains/day) respectively. In addition, the requirements for the first train service provided by the TSPs are summarised in Table. 8.6. TSP-PN and TSP-QR represent the settings of Pacific National and QR National respectively. The most preferable commencement times for Pacific National and QR National for their first services (i.e. TSP-PN-01 and TSP-QR-01) are 00:50 and 00:30 respectively. Since the freight operators are likely to transport as much coal as possible to the Newcastle port, they place a high priority to the inter-station runtimes and a lower priority to the track access charge. As it is not necessary for the coal trains to stop at the stations for loading or unloading, the most preferable dwell times are all zeros. Nevertheless, since the

Table 8.5. IP Definition

Attribute	Value
w_η ($)	8000
c_2 ($/kWh)	0
c_3 ($/MW)	0
c_4 ($/min)	300
Sequencing policy	FCFS

Table 8.6. TSP Definitions

Attribute	TSP-PN-01	TSP-QR-01
$\hat{\zeta}$ (hh:mm)	00:50	00:30
$\hat{T}_D$ (min)	{0, 0, 0, 0, 0, 0}	{0, 0, 0, 0, 0, 0}
$\hat{T}_R$ (min)	{50, 32, 8, 15, 20}	{55, 34, 10, 16, 22}
$\hat{c}$ ($)	4000	4250
z_{δ}	4	9
z_{T_D}	{9, 9, 9, 9, 9, 9}	{16, 16, 16, 16, 16, 16}
z_{T_R}	{25, 25, 25, 25, 25}	{36, 36, 36, 36, 36}
z_c	1	1
f_{ω_1}	1.0	0.0
f_{ω_2}	0.6	1.0
f_{ϕ_1}	1.0	1.0
f_{ϕ_2}	0.5	0.8
f_{ϕ_3}	0.0	0.4
τ	0.1	0.1

freight trains may need to give way to the passenger trains when there is a conflict in rights-of-way, the freight trains are willing to relax these constraints and stop at the stations.

As Pacific National owns both types of rolling stock, it is willing to accept offers with either type, but it prefers ω_1 over ω_2 due to the higher haul volume. On the other hand, since QR National only possesses ω_2, the preference of ω_1 is set to zero. As for the choice of flex levels, QR National is satisfied with all three levels, but Pacific National has stricter requirements of accepting only ϕ_1 and ϕ_2.

In the second scenario, there is a proposal in ARTC to reserve the time between the operation of CR-03 and CR04 for conducting track maintenance work. As a result, the 3-hour interval from 18:00 (i.e. 1080 mins) to 21:00 (i.e. 1260 mins) is not allowed for track access by the freight operators. When the information is provided to Pacific National and QR National, it is expected the operators will reduce the service headway time to 50 minutes and 60 minutes respectively to maintain the haul volume transported

Table 8.7. Settings of Two Scenarios

	Service Headway (min)	
	Scenario 1: 24-hour access	Scenario 2: 21-hour access
TSP-PN	60	72
TSP-QR	50	60

(i.e. the number of trains per day). The other operating characteristics are assumed to be unchanged. A summary for the two scenarios is shown in Table 8.7.

8.2.1.2. Results and Findings

The daily traffic distributions of the two scenarios are shown in Fig. 8.8 and Fig. 8.9 in which the speed profiles of trains beyond 1440 minutes (i.e. 24:00) are displayed back from 00:00. The key simulation results on the number of successful transactions, track access charge payments and capacity utilisation are shown in Table 8.8.

Scenario 1: According to Table 8.8, the majority of requests on track access rights by the freight operators are granted. Pacific National is able to secure 23 out of 24 schedules, while QR National is able to obtain 17 out of 20 schedules. Among these successful allocations of track capacity, Fig. 8.8 shows that the freight trains will give way to the passenger services when they encounter conflicts of rights-of-way at the intermediate train stations.

The train services that are not able to be scheduled on the track mostly occur between 18:00 (i.e. 1080 mins) to 21:00 (i.e. 1260 mins). Between this interval, capacity can only be allocated to one of the services operated by Pacific National (Fig. 8.8). The main reason for the difficulty in capacity allocation is the relatively close proximity between the two evening passenger trains (CR-03 and CR-04) where the 'knock-on' effect, as discussed in Chapter 6, imposes capacity constraints on the competing freight services.

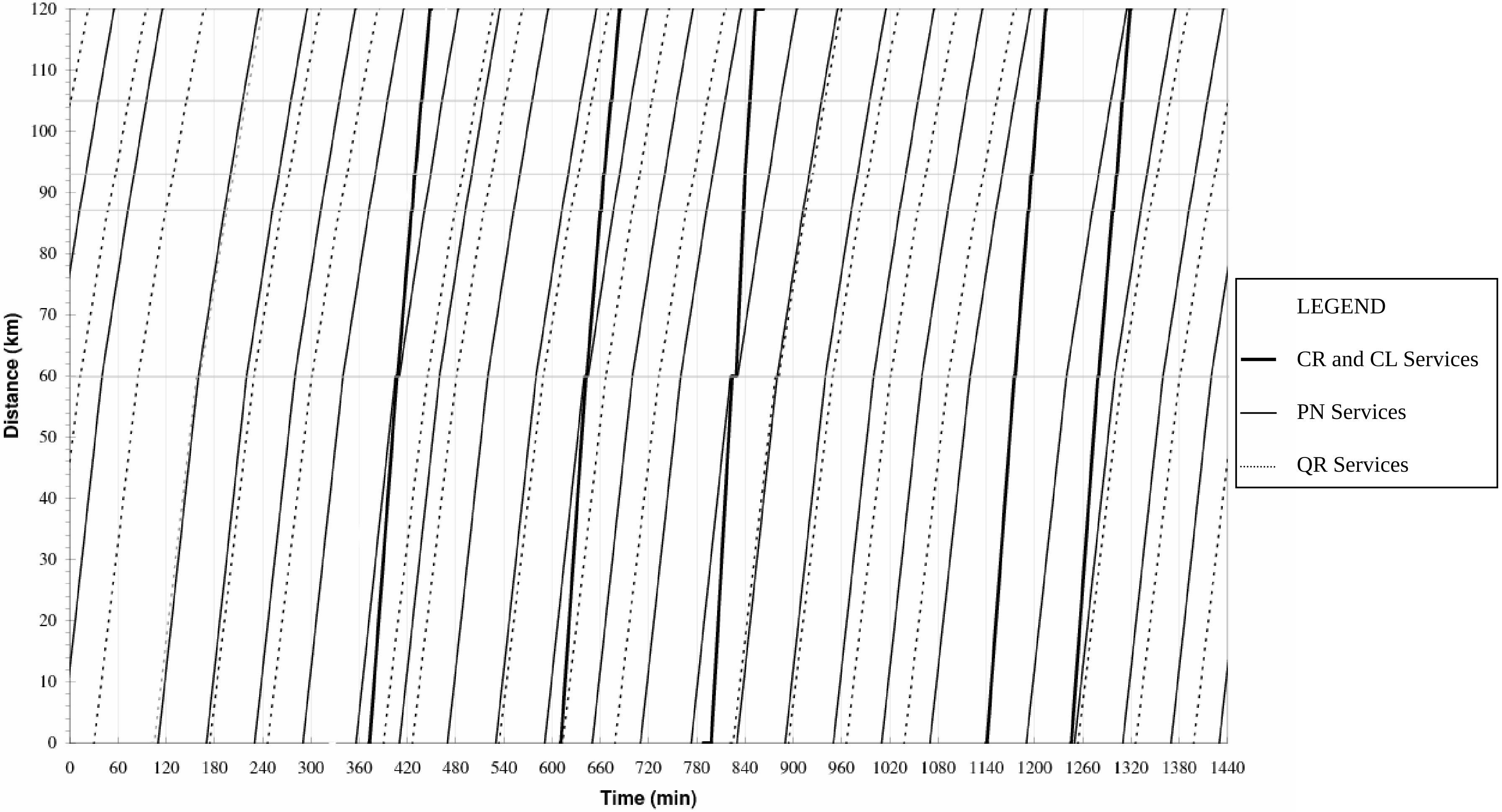


Fig. 8.8. Traffic distribution in scenario 1 (24-hour access)

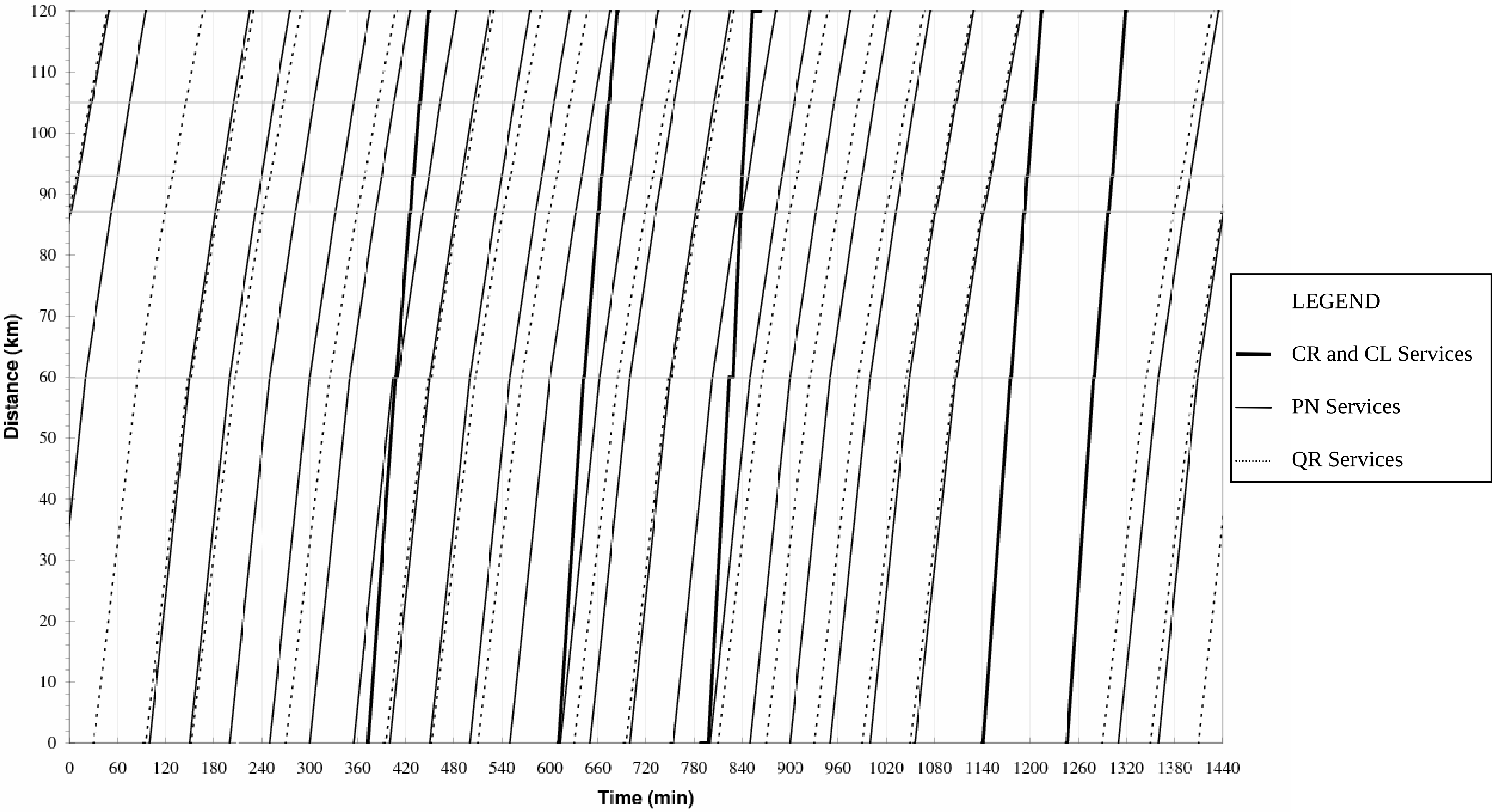


Fig. 8.9. Traffic distribution in scenario 2 (21-hour access)

Scenario 2: In this scenario, the time interval between 18:00 to 21:00 is reserved for maintenance work. The train that has been scheduled in this interval in scenario 1 is therefore absent in the train graph displayed in Fig. 8.9. From this figure, it can also be seen that the train services become more tightly packed due to the shorter headways. Moreover, according to Table 8.8, Pacific National is able to operate an additional service on the track which causes an increased in track access charge collection by ARTC and a reduction in capacity utilisation from 0.127 to 0.122.

The ability to schedule an additional service on the track is likely to be the result of the more uniform headways between the two train operators. In scenario 1, since the stakeholders are attempting to evenly distribute their trains over the 24-hour interval, the difference in desired headways is 12 minutes (60 mins and 72 mins). This difference is reduced to 10 minutes (50 mins and 60 mins) in scenario 2. With a more uniform headway, the constraints on capacity are relieved as the stakeholders may operate their trains more easily in an alternate manner.

Having an additional train on the track, one may expect the capacity utilisation will increase accordingly. However, there is an observed reduction in capacity utilisation in the simulation. This is mainly contributed by the reservation of track capacity between the two passenger trains which has indirectly eliminated the problem of heterogeneous traffic of different train speeds during the interval.

Conclusion: The simulation suggests that introducing a maintenance timeslot between 18:00 to 21:00 is beneficial to ARTC. This time interval is suitable for

Table 8.8. Summary of Results

	Scenario 1: 24-hour access			Scenario 2: 21-hour access		
	TSP-PN	TSP-QR	Total	TSP-PN	TSP-QR	Total
Number of successful transaction	23	17	40	24	17	41
TAC payments ($)	94,670	72,434	167,104	98,867	72,406	171,273
Capacity utilisation	-	-	0.127	-	-	0.122

conducting maintenance work rather than for train operation because capacity allocation during this interval appears difficult. By restricting track access during this interval, not only does it improve the track access charge collection and capacity utilisation, but the maintenance work also provides a safer transportation network. To obtain further improvement of capacity utilisation, ARTC is also recommended to persuade the freight operators to employ the same service headway wherever possible.

8.2.2. Schedule Coordination at the Liverpool Lime Street Station

8.2.2.1. Problem Description and Simulation Setup

The schedule coordination problem at the Liverpool Lime Street station involving Northern Rail and Central Trains is studied using the TSP-TSP transaction model presented in Chapter 7. In this study, it is assumed that the simulation is conducted from the perspective of Northern Rail whose train planners attempt to determine the possible operating conditions for its service from Preston to Liverpool if they are to coordinate the train schedule with Central Trains.

The background of the scheduling problem is illustrated in Fig 8.10. TSP-1 and TSP-2 represent Northern Rail and Central Trains respectively. Northern Rail is operating a service from Preston to Liverpool which requires a journey time of 60 minutes and a dwell time of 15 minutes at the Liverpool station. On the other hand,

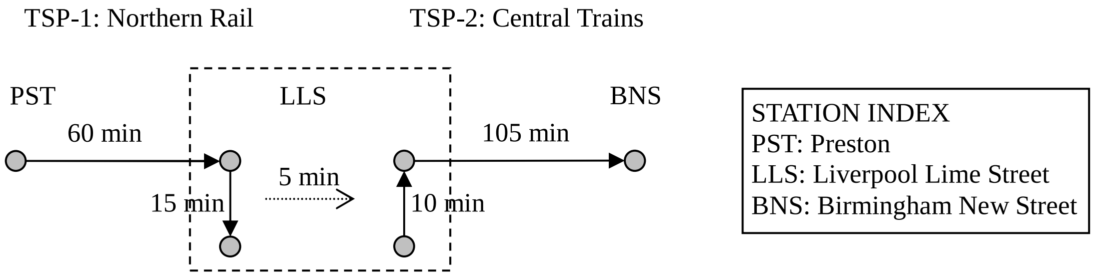


Fig. 8.10. Unidirectional transfer at Liverpool Lime Street station

the service provided by Central Trains from Liverpool to Birmingham consists of a journey time of 105 minutes and a dwell time of 10 minutes. The minimum transfer time between the two services (i.e. walking time between the two platforms) is 5 minutes. Since Liverpool Lime Street is the terminal station for the Northern Rail's service, the case shown in Fig 8.10 represents a unidirectional passenger transfer from Northern Rail to Central Trains. In addition, according to the past timetabling experience, the commencement time of the service operated by Central Trains is likely to be 70 minutes later than the commencement time of the Northern Rail's service. Therefore, the default passenger waiting time computed by (7.5) is 15 minutes.

Suppose the current average train fares for the Northern Rail and Central Trains services are £8.00 and £17.00 respectively. These train fares are expected to give rise to a maximum demand of 50 passengers when the waiting time is zero and the demand will cease when the waiting time exceeds 30 minutes. Moreover, the current estimation of idle costs for the rolling stock of Northern Rail and Central Trains are £20/min and £25/min respectively. The base case for the situation just described is denoted as case A in Table 8.9. Simulation of this case yields the probable outcome

Table 8.9. Simulation Setup for Schedule Coordination

Case	Description	Commencement time (min)		Average Train Fare (£/person)		Max. Demand (persons)		Idle Costs (£/min)	
		$\hat{\zeta}_1$	$\hat{\zeta}_2$	k_1	k_2	G^*_{12}	G^*_{21}	c_1	c_2
A	Unidirectional transfer *Default schedules lead to waiting time of 15 minutes*	0	70	8	17	50	0	20	25
B	Unidirectional transfer *Default schedules lead to waiting time of 25 minutes*	0	80	8	17	50	0	20	25
C	Unidirectional transfer *Reduced train fare to increase passenger demand*	0	70	6	17	70	0	20	25
D	Unidirectional transfer *Higher marginal cost for rolling stock*	0	70	8	17	50	0	40	25
E	Bidirectional transfer *Same rolling stock is used for the backward journey*	20	0	8	17	50	50	20	25

(i.e. train schedules) derived by negotiation.

Since the exact commencement time for Central Trains is not known, the train planners may like to further examine the consequence if it is postponed to a later time, say 80 minutes. This leads to a default waiting time of 25 minutes and the corresponding simulation parameters are given by case B. In addition, case C refers to the situation when Northern Rail attempts to increase the passenger demand by reducing the average train fare by £2.00. In case D, the stakeholder is considering upgrading the rolling stock for its service which leads to an increase in idle cost to £40/min. Finally, case E demonstrates an example of bidirectional transfer if the same set of rolling stock is used for the backward journey as shown in Fig. 8.11.

Having devised the situations intended for investigation, the train planners of Northern Rail can generate results using the simulator for TSP-TSP negotiation. During the simulation, the train planners may assume the proponent (i.e. Central Trains) to employ Strategy-PO (S_{po}), which has no intention to concede for the benefits of Northern Rail. On the other hand, Northern Rail may employ both Strategy-PO and Strategy-MAX (S_{max}) to examine the quality of the Pareto-optimal solution and the suboptimal one. The simulation results are summarised in Table 8.10.

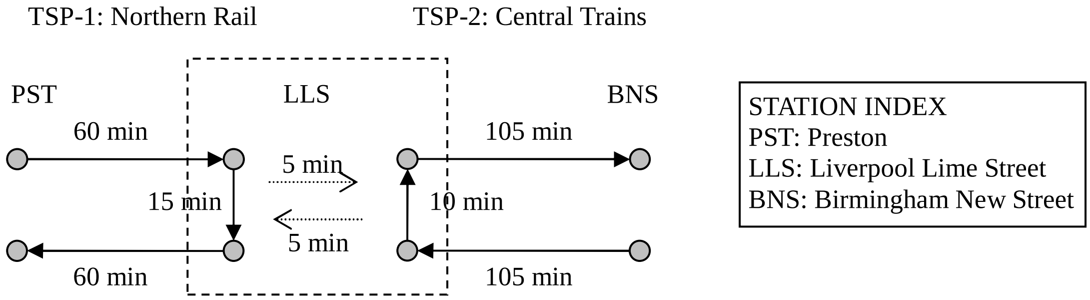


Fig. 8.11. Bidirectional transfer at Liverpool Lime Street station

Table 8.10. Simulation Results for Schedule Coordination

		Case A	Case B	Case C	Case D	Case E
Solution	(S_{po}, S_{po})	{2, 70}	{5, 80}	{3, 70}	{2, 70}	{39, 0}
$\{\zeta_1, \zeta_2\}$	(S_{max}, S_{po})	{7, 70}	{15, 80}	{7, 70}	{7, 70}	{40, 0}
Revenue gained by TSP-1	(S_{po}, S_{po})	289.89	122.22	292.80	244.89	359.11
Y_1 (£)	(S_{max}, S_{po})	231.56	55.56	250.13	91.56	344.44
Revenue gained by TSP-2	(S_{po}, S_{po})	789.56	755.56	999.60	690.39	1570.61
Y_2 (£)	(S_{max}, S_{po})	690.39	472.22	1105.38	789.56	1581.94
Number of negotiation	(S_{po}, S_{po})	147	504	167	133	19
rounds	(S_{max}, S_{po})	47	73	57	41	14
Waiting time $L_1 \rightarrow L_2$	(S_{po}, S_{po})	13	20	12	13	11
w_{12} (min)	(S_{max}, S_{po})	8	10	8	8	10
Waiting time $L_2 \rightarrow L_1$	(S_{po}, S_{po})	-	-	-	-	4
w_{21} (min)	(S_{max}, S_{po})	-	-	-	-	5
Demand for $L_1 \rightarrow L_2$	(S_{po}, S_{po})	40.6	27.7	58.8	40.6	42.3
D_{12} (persons)	(S_{max}, S_{po})	46.4	44.5	65.0	46.4	44.4
Demand for $L_2 \rightarrow L_1$	(S_{po}, S_{po})	-	-	-	-	49.1
D_{21} (persons)	(S_{max}, S_{po})	-	-	-	-	48.6

8.2.2.2. Results and Findings

Case A: In Table 8.10, the solution obtained in this case using the negotiation pair (S_{po}, S_{po}) is {2, 70}. The Pareto-optimal solution has reduced the passenger waiting time by 2 minutes (from 15 to 13 minutes) when Northern Rail (i.e. TSP-1) postponed its commencement time from 0 to 2 minutes during the negotiation. With the balance between the income generated from a passenger demand of 40.6 and the 2-minute idle cost of rolling stock, the overall revenue gained by Northern Rail is found to be £289.89. On the other hand, the solution obtained from the negotiation pair (S_{max}, S_{po}) is {7, 70}. As S_{max} aims to reduce the negotiation time by sacrificing Pareto-optimality, the commencement time for Northern Rail is further delayed to 7 minutes which leads to a higher idle cost. Although the passenger demand has been increased to 46.4 due to a shorter waiting time of 8 minutes, the overall revenue gained is lowered to £231.56. Nevertheless, since both simulated negotiations lead to a considerable gain in revenue, conducting a negotiation with Central Trains in practice is likely to be beneficial.

Case B: Despite the possible benefits in case A, if the commencement time for the service of Central Trains is changed to 80 minutes, Northern Rail suffers from a significant reduction in revenue collection. The corresponding values for (S_{po}, S_{po}) and (S_{max}, S_{po}) are £122.22 and £55.56 respectively. Under these circumstances, Northern Rail may consider withdrawing from the negotiation because the accompanied transaction costs (e.g. manpower, preparation of legal contracts, etc.) may swallow the monetary gained by the coordinated schedule.

Case C: In contrast to the results obtained in case B, the reduction of train fare has increased the revenue of Northern Rail to £292.80 in the negotiation (S_{po}, S_{po}) and £250.13 in the negotiation (S_{max}, S_{po}). This may suggest the stakeholder to indeed lower the train fare. However, according to the simulation results, the expected gain is not substantial (only £3 - £20) and thus the stakeholder may retain the basic train fare to avoid the additional administration cost of modifying the charging scheme.

Case D: With the improvement of the quality of rolling stock, the revenue collection of Northern Rail is reduced. The reduction is relatively small in the (S_{po}, S_{po}) negotiation, but considerably large in the (S_{max}, S_{po}) negotiation. With a higher idle cost for the rolling stock, if the stakeholder concedes easily for the benefits of the proponent (as in the case of employing S_{max}), the cumulative loss for the delay in commencement time will be increased dramatically.

Case E: The possibility of bidirectional transfer has provided a reasonable increase in revenue for Northern Rail. With an additional demand of almost 50 passengers in the backward journey, Northern Rail is willing to postpone the commencement time by about 20 minutes (from 20 minutes to 39 and 40 minutes) instead of only several minutes in case A.

Conclusion: Based on the simulation settings and results, Northern Rail should explore the possibility of schedule coordination with Central Trains using the current set of rolling stock. Preferably, the rolling stock should also be used for the backward journey. However, the stakeholder should pay serious attention to the possible errors in their estimation or prediction (e.g. commencement time for the proponent's service, passenger demand, etc.). It is also recommended to Northern Rail to negotiate in a cautious manner if adequate time is available for negotiation.

8.3. Remarks

This chapter has described some of the resource management problems experienced by the railway markets in Australia, the UK and China. Since the problems in Australia and the UK are the direct consequence of restructuring into open access markets, MAS-ORAM clearly has promising potential in studying these railways. On the other hand, MAS-ORAM may offer as a useful tool to determine whether a widespread competition is suitable for the Chinese railways.

Two hypothetical resource management problems in Australia and the UK have been studied to demonstrate how MAS-ORAM may be employed in practical planning and evaluation by the stakeholders before initiating any negotiation. The simulation results are hugely valuable in predicting the possible outcomes of different scenarios and they therefore suggest the appropriate objectives and actions to the railway stakeholders. Since the studies are all conducted in a virtual environment, such evaluation tool offers a cost-effective means to assist the train planners/service managers in making decisions with respect to the best interest of their employers.

Despite the demonstration of the usefulness of MAS-ORAM in the hypothetical case studies, it is still important and beneficial to validate the negotiation models on existing problems using actual data. Although the acquisition of data from industries is difficult since these data are usually confidential, it is anticipated that further research may involve the collaboration of the railway stakeholders to enhance the capabilities of the agents.

Chapter 9

Conclusions and Further Works

This chapter is organised into two sections. In the first part, the achievements presented in the thesis are summarised. In the second part, the further research spun off from the work is addressed.

9.1. Summary of Achievements

The achievements of the work in the thesis are summarised in this section. The major contributions associated with the MAS-ORAM (Multi-agent System for Open Railway Access Market) architecture, the transactions of IP-TSP, IP-TSP^n, TSP-TSP, and the application of the MAS-ORAM in resource planning are discussed.

9.1.1. MAS-ORAM Architecture

The thesis has presented a MAS framework to model an open railway access market. MAS-ORAM considers railway stakeholders as a group of loosely-coupled software agents which possess their individual objectives and are capable of conducting negotiation to resolve their conflicts in resource planning and management. Unlike the conventional modelling approaches which regard the entire system as a central decision-making unit, MAS-ORAM takes into consideration of the distributed nature of open markets as a result of the separation of ownership and responsibilities. This enables the examination of the local resource management problems and their interactions under a computer simulation environment. The idea has opened up a new means to conduct useful hypothetical studies in railway resource planning in open access markets.

9.1.2. IP-TSP Transaction

A bilateral negotiation on track access rights between an IP and a TSP has been modelled. By employing BSBP (Buyer and Seller Behaviour Protocol), the TSP agent can express its requirements by submitting a set of crisp constraints, while the responsibility of the IP agent is to generate offers for the TSP agent in consideration of the submitted constraints and its internal costs and benefits.

When making concession during the negotiation, the decision of the TSP on selecting an attribute is modelled as a PFCS (Prioritised Fuzzy Constraint Satisfaction) problem resolved by a rule-based system. By assigning different sets of effective priority values, the model allows the TSP agents to exhibit various conceding behaviour on commencement time, station dwell times, inter-station runtimes and track access charge. The TSP-model has also enabled the representation of non-monotonic change in satisfaction on these attributes by using fuzzy membership functions.

On the other hand, the ability of the IP agent to generate offers is enabled by a BNB (Branch-and-Bound) algorithm. The algorithm has the advantage of guaranteeing the optimal solution for the IP in each round of negotiation. Such solution takes into consideration of the overall balance among the costs on track usage, traction energy, peak demand, congestion and capacity utilisation. Since the algorithm suffers from a high computation demand, three heuristic rules have been proposed to reduce the simulation time. The resulting algorithm is recommended for conducting small-scale studies (e.g. less than 10 stations) for strategic and tactical planning.

Simulation results have also shown that the IP and TSP agents are able to negotiate for track access rights autonomously according to their pre-assigned operational objectives. When a deal is identified in the negotiation process, the solution is Pareto-optimal. The IP agent is also capable of resolving conflicts of rights-of-way and deriving offers for TSP agents in response to their willingness-to-pay and incentives in capacity requirements.

9.1.3. IP-TSP^n Transaction

A multilateral negotiation on track access rights between a single IP and a group of TSP has been modelled. MBSBP (Multiple Buyers and Seller Behaviour Protocol) enables the IP agent to negotiate with a set of TSP agents that are operating train services within the same CTP (Commencement Time Interval). The ability to conduct more than one IP-TSP negotiation allows the examination of the effect of competition on capacity utilisation and quality of train services.

FCFS (First-Come-First-Serve) and HW2PF (Highest-Willingness-to-Pay-First) have been proposed as the policies adopted by the IP agent to handle the negotiations in a sequential manner. Through t-tests and hypothesis testing, results have suggested HW2PF is more favourable to the IP (and express train services) when the traffic is dominated by fast trains, while FCFS tends to benefit the IP (and regular services) otherwise. In railway networks that support a mixed mode of freight and passenger traffic, a short CTP is also recommended for the freight services in order to introduce sensible competition in the railway market.

9.1.4. TSP-TSP Transaction

A bilateral negotiation on schedule coordination between two TSPs has been modelled. The negotiation employs a simple protocol that consists of communicative acts of proposing, accepting and rejecting offers. The objectives of the TSP agents are thus to generate an appropriate offer to the proponent and decide whether the counteroffers received should be accepted or rejected.

The offer generation problem has been modelled as a tree searching problem which composes of a quadratic programming problem and a pruning decision problem. LCPA (Lemke's Complementary Pivoting Algorithm) has been used to tackle the quadratic programming problem and heuristic rules have been devised to prune the search tree. The combined algorithm for generating offers in the schedule coordination problem is fast and efficient.

On the other hand, the decision on accepting and rejecting an offer is governed by one of the three proposed negotiation strategies. S_{po} is specially devised to steer the negotiation towards a Pareto-optimal solution. S_{min} and S_{max} are intended to represent the behaviour of a cautious and a desperate TSP respectively. Simulation results have shown that S_{min} is a good approximation to S_{po} in terms of quality of solution and negotiation time, while S_{max} is able to shorten the negotiation time but more likely to lead to lower quality of solution.

9.1.5. Applications

The thesis has also applied MAS-ORAM in railway resource planning under the backgrounds of two practical open access markets in Australia and the UK. The studies have demonstrated the advantage and potential use of the software prior to conducting the real (human-human) negotiation. The stakeholders may use the software to analyse and predict the possible negotiation outcomes in the planning stage which helps the decision makers to improve their negotiation power and avoid conducting unprofitable negotiations. MAS-ORAM also has potential applications in determining whether open access should be introduced in conventional railway markets such as the mainline railways in China.

9.2. Further Works

The presented work is a pioneer work on applying agent modelling to resolve modern resource management problems in open railway markets. The work is therefore a useful catalyst for further research in the discipline. Described below are two possible directions for further works on capability enhancement of agents and extension on transactions. In all cases, it is important to validate the applicability of the models in real-life situations. Close collaborations with the industry is vital.

9.2.1. Capability Enhancement of Agents

Being the initial work of agent modelling in open railway access markets, the agents modelled in this study are relatively primitive. In all negotiations, the agents are only given the ability to respond rigidly to the communicative acts of other agents. In other words, the agents may be regarded as reactive agents and they lack the ability to learn and/or initiate proactive acts. For example, in the IP-TSP negotiation, the TSP may not want to adopt a fixed operational strategy (i.e. passenger-oriented or expenditure-reducing), but desire to determine the behaviour according to the availability of track capacity supply. By analysing the replies from the IP agent, the TSP agents may attempt to deduce whether the required track capacity has been occupied. In such case, a passenger-oriented TSP agent will have no reason to insist on its requirements and may opt for reducing the expenditure. On the other hand, the IP agent may learn from the TSP agent's response and promote its idle track capacity proactively by lowering the relevant charge rates. The BDI (Belief-Desire-Intention) agent modelling paradigm is a potential means to capture such proactive behaviour.

In addition, the agents, especially the IP, can benefit from a faster algorithm in decision-making. Although the BNB algorithm for capacity allocation is able to derive the optimal solution for the IP in each negotiation round, the algorithm suffers from a high computational complexity. Even if equipped with the proposed heuristic rules, the required simulation time becomes impractical when the problem involves large number of stations (or sidings). Such issue is especially apparent when negotiation is an iterative process in which the algorithm will be frequently reused. As a result, further research may be conducted to explore other algorithms, such as GA, which derive near-optimal solution with shorter simulation time.

Furthermore, since the negotiation models derived in this study assume that the railway stakeholders are free from regulations on track capacity allocation. It is thus worth enhancing the models by considering the equity issue on capacity allocation. One means to measure equity is using Gini Coefficient, which can be incorporated in the objective function of the IP. The imposition of a minimum level of equity by the railway regulatory bodies may be represented by a constraint in the in the combinatorial optimisation problem.

9.2.2. Extension on Transactions

The schedule coordination problem between two TSPs modelled in the study has assumed that the IP will grant the necessary track access rights to the TSPs. However, in practice, the TSPs are uncertain if they can indeed obtain the required capacity. Therefore, additional work can be performed to coordinate the TSP-TSP negotiation with the IP-TSP negotiation. A possible approach is to first perform a TSP-TSP negotiation to produce the desired schedules for the TSPs. Having obtained these schedules, the TSPs initiate a negotiation with the IP to resolve the capacity constraints (if any). However, as the resultant schedules may deviate from the desired ones, the TSPs may need to negotiate among themselves again to refine the commitment associated with the schedule coordination. This may in turn lead to an update arrangement with the IP. In some cases, the negotiation process may iterate a few times before all stakeholders settle at a final agreement.

While the major transaction problems involving IP and TSPs have been studied, other negotiations are also worth examining. Apart from the track access rights allocation problem between IP and TSP, platform allocation at railway stations is also an important problem. When many trains are arriving at a busy station, the available platforms may become scarce and the IP will attempt to reduce the station dwell times of the TSPs' train services. On the other hand, owing to a schedule coordination commitment with other TSPs, some of them may prefer a longer dwell time at a particular platform for better transfer arrangements (e.g. aiding bidirectional transfer). This type of negotiation concerning platform utilisation is also valuable in railway resource planning.

Moreover, the possibility of adopting an auction approach to allocate track capacity should not be overlooked. Although the current access pricing regimes are dominated by posted pricing and negotiation, many policy makers are still striving to explore better regimes, and auctioning is one of the promising approaches. It is therefore desirable to evaluate the complexity to set up an auctioning system in open railway markets and its efficiency in capacity utilisation when compared with the negotiation approach. Results obtained from such study can be of great advantage in suggesting an appropriate pricing regime in various open access markets.

Appendix A

Complexity Analysis of BNB Algorithm

The BNB algorithm shown in Fig. 5.11 is repetitive and it terminates when all nodes in $LIST$ have been evaluated. The exact number of node evaluations may vary, but in the worst scenario, the entire tree is expanded to yield the maximum number $N_T = n_\phi + n_\phi n_\omega + n_\phi n_\omega n_\zeta + n_\phi n_\omega n_\zeta n_{D1} + n_\phi n_\omega n_\zeta n_{D1} n_{R1} + n_\phi n_\omega n_\zeta n_{D1} n_{R1} n_{D2} + \ldots + n_\phi n_\omega n_\zeta n_{Dn_s} \prod_{i=1}^{n_s-1} n_{Di} n_{Rj}$, where n_ϕ, n_ω and n_ζ are the number of feasible flex levels, rolling stock and commencement times respectively; n_{Di} and n_{Rj} are the numbers of feasible dwell times at station i and feasible runtimes in inter-station j; and n_s is the number of visiting stations. In other words, the algorithm has a time complexity of $O(n_\phi n_\omega n_\zeta n_{Dn_s} \prod_{i=1}^{n_s-1} n_{Di} n_{Ri}) = O(n_\phi n_\omega n_\zeta {n_D}^{n_s} {n_R}^{n_s})$, where $n_D = \max\limits_{1 \le i \le n_s} (n_{Di})$ and $n_R = \max\limits_{1 \le j \le n_s - 1} (n_{Rj})$.

However, this assumes that one unit of computation time is used to evaluate the utility value U of a node, which is unreasonable. To obtain a more realistic estimation for the time complexity, the computation requirements for the sub-charges and capacity utilisation are summarised in Table A.1. It can be seen that PDC is the dominating factor over the five components so that when its computational complexity is considered with the worst-case expansion scenario for the BNB algorithm, the overall

Table A.1. Time Complexity of Components in Objective Function

Term	Complexity	Description of derivation
TUC	$O(n_\omega)$	Among the available rolling stock, find the maximum charge rates. This requires n_ω data retrievals.
TEC	$O(n_\omega \prod_{j=1}^{n_s-1} n_{Rj}) = O(n_\omega {n_R}^{n_s})$	Among the available rolling stock, find the maximum cumulative energy consumption for the inter-station runs. This requires a maximum of n_R retrievals of the energy data within each inter-station run, and hence ${n_R}^{n_s-1}$ computations of cumulative energy consumption for each ω.
PDC	$O(n_\omega n_\zeta \prod_{j=1}^{n_s-1} n_{Dj} n_{Rj}) = O(n_\omega n_\zeta {n_D}^{n_s} {n_R}^{n_s})$	Among the available rolling stock, find the schedule (combination of commencement time, dwell times and runtimes) which causes the highest increase in peak demand. For each inter-station run, there is a maximum of $n_D n_R$ combinations of dwell times and runtimes. This leads to a possibility of evaluating n_ω schedules for each ω.
CGC	$O(n_\phi + n_\omega n_\zeta \prod_{j=1}^{n_s-1} n_{Dj}) = O(n_\omega n_\zeta {n_D}^{n_s})$	Among the available flex levels, find the flex level with the highest discount factor. The search requires n_ϕ data retrievals. In addition, each ω has $n_\zeta {n_D}^{n_s-1}$ schedules for evaluations. This term is independent on n_R because congestion charge is maximised by directly retrieving the maximum runtime which consumes only 1 unit of computation.
CPU	$O(n_\omega n_\zeta \prod_{j=1}^{n_s-1} n_{Dj}) = O(n_\omega n_\zeta {n_D}^{n_s})$	Each ω has $n_\zeta {n_D}^{n_s-1}$ schedules for evaluations. This term is independent on n_R because incremental capacity utilisation is minimised by directly retrieving the minimum runtime which consumes only 1 unit of computation.

time complexity is given by $O(n_\phi n_\omega n_\zeta {n_D}^{n_s} {n_R}^{n_s}) \times O(n_\omega n_\zeta {n_D}^{n_s} {n_R}^{n_s})$ $= O(n_\phi {n_\omega}^2 {n_\zeta}^2 {n_D}^{2n_s} {n_R}^{2n_s})$.

To illustrate the computation demand, consider $n_\phi = 5$ and $n_\omega = 90$ (in the UK, there are about 90 classes of rolling stock). Suppose the negotiation adopts a time-window of 30 mins so that the maximum value of $n_\zeta = 30$, and assume $n_D = 10$ and $n_R = 20$. If the number of visited stations n_s is 4, the order of computation becomes 5.83×10^{16}, and if n_s is doubled, the order increases substantially to 9.33×10^{25}.

The complexity in fact increases in exponential order for n_D and n_R with respect to the number of stations n_s, and in polynomial order for n_ϕ, n_ω and n_ζ. This means the algorithm is computational expensive and it quickly becomes impractical as the number of stations increases. Nevertheless, since the above estimation considers the worst scenario when the entire tree is expanded, the computation is less demanding in cases when intermediate nodes are pruned.

Appendix B

Statistical Tables

Table B.1. Table of the Student's t-distribution

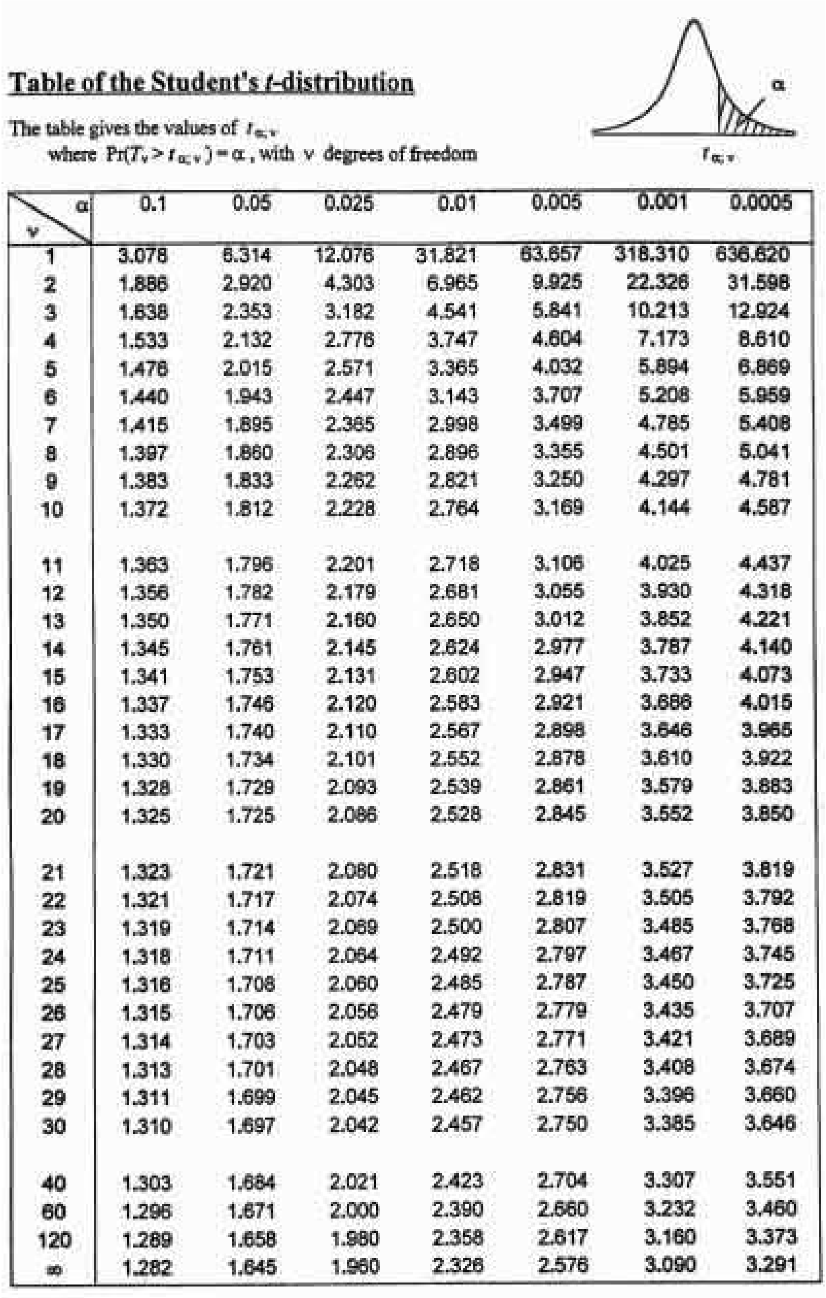
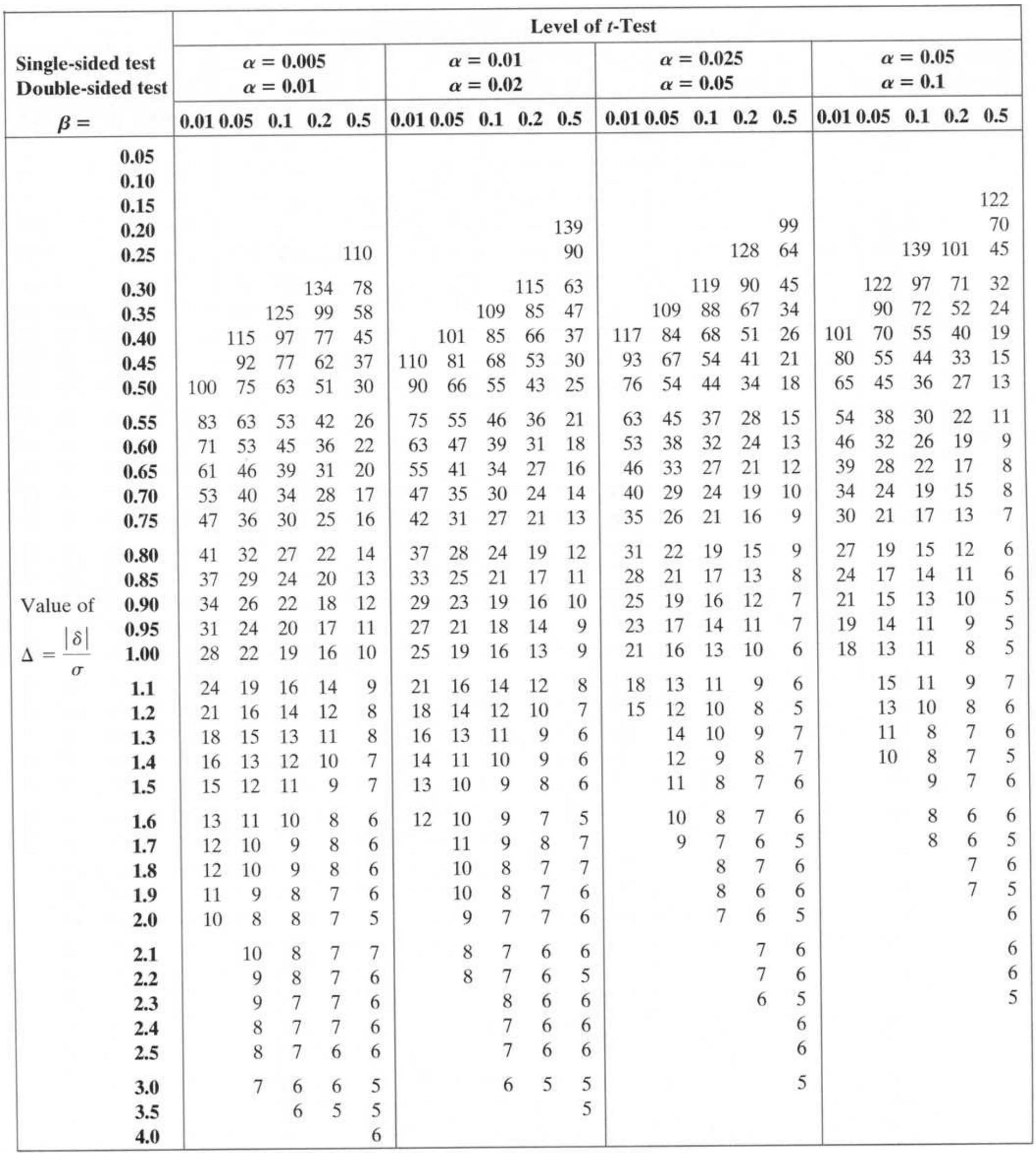


Table of the Student's *t*-distribution

The table gives the values of $t_{\alpha;\nu}$
where $\Pr(T_\nu > t_{\alpha;\nu}) = \alpha$, with ν degrees of freedom

ν \ α	0.1	0.05	0.025	0.01	0.005	0.001	0.0005
1	3.078	6.314	12.076	31.821	63.657	318.310	636.620
2	1.886	2.920	4.303	6.965	9.925	22.326	31.598
3	1.638	2.353	3.182	4.541	5.841	10.213	12.924
4	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	1.476	2.015	2.571	3.365	4.032	5.894	6.869
6	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	1.319	1.714	2.069	2.500	2.807	3.485	3.768
24	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	1.314	1.703	2.052	2.473	2.771	3.421	3.689
28	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	1.311	1.699	2.045	2.462	2.756	3.396	3.660
30	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	1.296	1.671	2.000	2.390	2.660	3.232	3.460
120	1.289	1.658	1.980	2.358	2.617	3.160	3.373
∞	1.282	1.645	1.960	2.326	2.576	3.090	3.291

Table B.2. Estimation of Sample Size for Two-sample Hypothesis Test of Unknown Variance



		Level of *t*-Test																			
Single-sided test		$\alpha = 0.005$					$\alpha = 0.01$					$\alpha = 0.025$					$\alpha = 0.05$				
Double-sided test		$\alpha = 0.01$					$\alpha = 0.02$					$\alpha = 0.05$					$\alpha = 0.1$				
$\beta =$		**0.01**	**0.05**	**0.1**	**0.2**	**0.5**	**0.01**	**0.05**	**0.1**	**0.2**	**0.5**	**0.01**	**0.05**	**0.1**	**0.2**	**0.5**	**0.01**	**0.05**	**0.1**	**0.2**	**0.5**
Value of $\Delta = \frac{\lvert\delta\rvert}{\sigma}$	**0.05**																				
	0.10																				
	0.15																				122
	0.20										139					99					70
	0.25					110					90				128	64			139	101	45
	0.30				134	78				115	63			119	90	45		122	97	71	32
	0.35			125	99	58			109	85	47		109	88	67	34		90	72	52	24
	0.40		115	97	77	45		101	85	66	37	117	84	68	51	26	101	70	55	40	19
	0.45		92	77	62	37	110	81	68	53	30	93	67	54	41	21	80	55	44	33	15
	0.50	100	75	63	51	30	90	66	55	43	25	76	54	44	34	18	65	45	36	27	13
	0.55	83	63	53	42	26	75	55	46	36	21	63	45	37	28	15	54	38	30	22	11
	0.60	71	53	45	36	22	63	47	39	31	18	53	38	32	24	13	46	32	26	19	9
	0.65	61	46	39	31	20	55	41	34	27	16	46	33	27	21	12	39	28	22	17	8
	0.70	53	40	34	28	17	47	35	30	24	14	40	29	24	19	10	34	24	19	15	8
	0.75	47	36	30	25	16	42	31	27	21	13	35	26	21	16	9	30	21	17	13	7
	0.80	41	32	27	22	14	37	28	24	19	12	31	22	19	15	9	27	19	15	12	6
	0.85	37	29	24	20	13	33	25	21	17	11	28	21	17	13	8	24	17	14	11	6
	0.90	34	26	22	18	12	29	23	19	16	10	25	19	16	12	7	21	15	13	10	5
	0.95	31	24	20	17	11	27	21	18	14	9	23	17	14	11	7	19	14	11	9	5
	1.00	28	22	19	16	10	25	19	16	13	9	21	16	13	10	6	18	13	11	8	5
	1.1	24	19	16	14	9	21	16	14	12	8	18	13	11	9	6		15	11	9	7
	1.2	21	16	14	12	8	18	14	12	10	7	15	12	10	8	5		13	10	8	6
	1.3	18	15	13	11	8	16	13	11	9	6		14	10	9	7		11	8	7	6
	1.4	16	13	12	10	7	14	11	10	9	6		12	9	8	7		10	8	7	5
	1.5	15	12	11	9	7	13	10	9	8	6		11	8	7	6			9	7	6
	1.6	13	11	10	8	6	12	10	9	7	5		10	8	7	6			8	6	6
	1.7	12	10	9	8	6		11	9	8	7		9	7	6	5			8	6	5
	1.8	12	10	9	8	6		10	8	7	7			8	7	6				7	6
	1.9	11	9	8	7	6		10	8	7	6			8	6	6				7	5
	2.0	10	8	8	7	5		9	7	7	6			7	6	5					6
	2.1		10	8	7	7		8	7	6	6				7	6					6
	2.2		9	8	7	6		8	7	6	5				7	6					6
	2.3		9	7	7	6			8	6	6				6	5					5
	2.4		8	7	7	6			7	6	6					6					
	2.5		8	7	6	6			7	6	6					6					
	3.0		7	6	6	5			6	5	5					5					
	3.5			6	5	5					5										
	4.0					6															

Source taken from:
Davies, L. (ed.) (1956) ‘Design and analysis of industrial experiments’ (Edinburgh: Oliver & Boyd), reproduced by Walpole, R.E., Myers, R.H., and Myers, S.L. (1998) ‘Probability and statistics for engineers and scientists’ (Upper Saddle River, NJ: Prentice Hall)

Appendix C

Partitioning and Transformation of the Objective Functions

In the following, let $\mathbf{x} = \begin{bmatrix} \zeta_i \\ \zeta_j \end{bmatrix}$.

1) $\mathbf{P_{i \to j}}$: Unidirectional transfer from L_i to L_j

The objective function is obtained from (2.8) with the exclusion of $k_i G(\zeta_j, \zeta_i)$ which is shown in (C1).

$$\max \quad Y_i = k_i G_{ij}^* \left[1 - \left(\frac{\zeta_j - \zeta_i + z_{ij}}{w_m} \right)^2 \right] - c_i (\zeta_i - \hat{\zeta}_i) \tag{C1}$$

With expansion and rearrangement using matrix notation, (C1) can expressed by (C2).

$$\begin{aligned} \max \quad Y_i = {} & k_i G_{ij}^* \left(1 - \frac{{z_{ij}}^2}{{w_m}^2} \right) + c_i \hat{\zeta}_i + \frac{2 k_i G_{ij}^* z_{ij}}{{w_m}^2} \begin{bmatrix} 1 - c_i {w_m}^2 / 2 k_i G_{ij}^* z_{ij} \\ -1 \end{bmatrix}^T \mathbf{x} \\ & - \frac{1}{2} \frac{2 k_i G_{ij}^*}{{w_m}^2} \mathbf{x}^T \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \mathbf{x} \end{aligned} \tag{C2}$$

Since the constant terms in an objective function can be eliminated from the optimisation problem, the objective function can be reduced to (C3) without affecting the optimal solution.

$$\max \quad Y_i' = \frac{2k_iG_{ij}^*z_{ij}}{w_m^{\ 2}}\begin{bmatrix}1-c_iw_m^{\ 2}/2k_iG_{ij}^*z_{ij}\\ -1\end{bmatrix}^T\mathbf{x}-\frac{1}{2}\frac{2k_iG_{ij}^*}{w_m^{\ 2}}\mathbf{x}^T\begin{bmatrix}1 & -1\\ -1 & 1\end{bmatrix}\mathbf{x} \qquad \text{(C3)}$$

Also, any maximisation function can be converted to the equivalent minimisation form by the multiplication of the objective function by –1. This is shown in (C4).

$$\min \quad Z_i = -Y_i' = \frac{2k_iG_{ij}^*z_{ij}}{w_m^{\ 2}}\begin{bmatrix}c_iw_m^{\ 2}/2k_iG_{ij}^*z_{ij}-1\\ 1\end{bmatrix}^T\mathbf{x}+\frac{1}{2}\frac{2k_iG_{ij}^*}{w_m^{\ 2}}\mathbf{x}^T\begin{bmatrix}1 & -1\\ -1 & 1\end{bmatrix}\mathbf{x} \qquad \text{(C4)}$$

2) $\mathbf{P}_{\mathbf{j}\to\mathbf{i}}$: Unidirectional transfer from L_j to L_i

This problem can be similarly transformed to a minimisation problem with (C1)-(C4) replaced by (C5)-(C8).

$$\max \quad Y_i = k_iG_{ji}^*\left[1-\left(\frac{\zeta_i-\zeta_j+z_{ji}}{w_m}\right)^2\right]-c_i(\zeta_i-\hat{\zeta}_i) \qquad \text{(C5)}$$

$$\begin{aligned}\max \quad Y_i = {} & k_iG_{ji}^*\left(1-\frac{z_{ji}^{\ 2}}{w_m^{\ 2}}\right)+c_i\hat{\zeta}_i+\frac{2k_iG_{ji}^*z_{ji}}{w_m^{\ 2}}\begin{bmatrix}-1-c_iw_m^{\ 2}/2k_iG_{ji}^*z_{ji}\\ 1\end{bmatrix}^T\mathbf{x}\\ & -\frac{1}{2}\frac{2k_iG_{ji}^*}{w_m^{\ 2}}\mathbf{x}^T\begin{bmatrix}1 & -1\\ -1 & 1\end{bmatrix}\mathbf{x}\end{aligned} \qquad \text{(C6)}$$

$$\max \quad Y_i' = \frac{2k_iG_{ji}^*z_{ji}}{w_m^{\ 2}}\begin{bmatrix}-1-c_iw_m^{\ 2}/2k_iG_{ji}^*z_{ji}\\ 1\end{bmatrix}^T\mathbf{x}-\frac{1}{2}\frac{2k_iG_{ji}^*}{w_m^{\ 2}}\mathbf{x}^T\begin{bmatrix}1 & -1\\ -1 & 1\end{bmatrix}\mathbf{x} \qquad \text{(C7)}$$

$$\min \quad Z_i = -Y_i' = \frac{2k_iG_{ji}^*z_{ji}}{w_m^{\ 2}}\begin{bmatrix}c_iw_m^{\ 2}/2k_iG_{ji}^*z_{ji}+1\\ -1\end{bmatrix}^T\mathbf{x}+\frac{1}{2}\frac{2k_iG_{ji}^*}{w_m^{\ 2}}\mathbf{x}^T\begin{bmatrix}1 & -1\\ -1 & 1\end{bmatrix}\mathbf{x} \qquad \text{(C8)}$$

3) $\mathbf{P}_{\mathbf{i}\leftrightarrow\mathbf{j}}$: Bidirectional transfer to and from L_i and L_j

This problem is similarly transformed to a minimisation problem with (C1)-(C4) replaced by (C9)-(C12).

$$\max \quad Y_i = k_i G_{ij}^{*}\left[1-\left(\frac{\zeta_j - \zeta_i + z_{ij}}{w_m}\right)^2\right] + k_i G_{ji}^{*}\left[1-\left(\frac{\zeta_i - \zeta_j + z_{ji}}{w_m}\right)^2\right] - c_i(\zeta_i - \hat{\zeta}_i) \tag{C9}$$

$$\begin{aligned}
\max \quad Y_i &= k_i\left[G_{ij}^{*}\left(1-\frac{{z_{ij}}^2}{{w_m}^2}\right) + G_{ji}^{*}\left(1-\frac{{z_{ji}}^2}{{w_m}^2}\right)\right] + c_i \hat{\zeta}_i \\
&+ \quad \frac{2k_i}{{w_m}^2}\begin{bmatrix} G_{ij}^{*} z_{ij} - G_{ji}^{*} z_{ji} - \dfrac{c_i {w_m}^2}{2k_i} \\ G_{ji}^{*} z_{ji} - G_{ij}^{*} z_{ij} \end{bmatrix}^T \mathbf{x} \\
&- \quad \frac{1}{2}\frac{2k_i}{{w_m}^2}\mathbf{x}^T \begin{bmatrix} G_{ij}^{*} + G_{ji}^{*} & -(G_{ij}^{*} + G_{ji}^{*}) \\ -(G_{ij}^{*} + G_{ji}^{*}) & G_{ij}^{*} + G_{ji}^{*} \end{bmatrix}\mathbf{x}
\end{aligned} \tag{C10}$$

$$\begin{aligned}
\max \quad Y_i' &= \frac{2k_i}{{w_m}^2}\begin{bmatrix} G_{ij}^{*} z_{ij} - G_{ji}^{*} z_{ji} - \dfrac{c_i {w_m}^2}{2k_i} \\ G_{ji}^{*} z_{ji} - G_{ij}^{*} z_{ij} \end{bmatrix}^T \mathbf{x} \\
&- \quad \frac{1}{2}\frac{2k_i}{{w_m}^2}\mathbf{x}^T \begin{bmatrix} G_{ij}^{*} + G_{ji}^{*} & -(G_{ij}^{*} + G_{ji}^{*}) \\ -(G_{ij}^{*} + G_{ji}^{*}) & G_{ij}^{*} + G_{ji}^{*} \end{bmatrix}\mathbf{x}
\end{aligned} \tag{C11}$$

$$\begin{aligned}
\min \quad Z_i = -Y_i' &= \frac{2k_i}{{w_m}^2}\begin{bmatrix} G_{ji}^{*} z_{ji} - G_{ij}^{*} z_{ij} + \dfrac{c_i {w_m}^2}{2k_i} \\ G_{ij}^{*} z_{ij} - G_{ji}^{*} z_{ji} \end{bmatrix}^T \mathbf{x} \\
&+ \quad \frac{1}{2}\frac{2k_i}{{w_m}^2}\mathbf{x}^T \begin{bmatrix} G_{ij}^{*} + G_{ji}^{*} & -(G_{ij}^{*} + G_{ji}^{*}) \\ -(G_{ij}^{*} + G_{ji}^{*}) & G_{ij}^{*} + G_{ji}^{*} \end{bmatrix}\mathbf{x}
\end{aligned} \tag{C12}$$

Appendix D

Conditions for Optimality

A solution $\hat{\mathbf{x}} = (\hat{x}_1, \hat{x}_2, ..., \hat{x}_n)$ is an optimal solution to $\min\{f(\mathbf{x})\}$ subject to $g_i(\mathbf{x})$, $i = \{1, 2, ..., m\}$ if $f(\mathbf{x})$ is a convex function, $g_i(\mathbf{x})$ are concave functions, and the following Karush-Kuhn-Tucker (KKT) conditions are satisfied (Hillier & Lieberman, 1995):

$$\frac{\partial f}{\partial x_i} - \sum_{i=1}^{m} u_i \frac{\partial g_i}{\partial x_j} \le 0 \quad \text{at} \quad \mathbf{x} = \hat{\mathbf{x}}, \text{ for } \quad j = 1, 2, ..., n \tag{D1}$$

$$\hat{x}_j \left(\frac{\partial f}{\partial x_i} - \sum_{i=1}^{m} u_i \frac{\partial g_i}{\partial x_j} \right) = 0 \quad \text{at} \quad \mathbf{x} = \hat{\mathbf{x}}, \text{ for } \quad j = 1, 2, ..., n \tag{D2}$$

$$g_i(\hat{\mathbf{x}}) - b_i \le 0 \quad \text{for} \quad i = 1, 2, ..., m \tag{D3}$$

$$u_i[g_i(\hat{\mathbf{x}}) - b_i] = 0 \quad \text{for} \quad i = 1, 2, ..., m \tag{D4}$$

$$\hat{x}_j \ge 0 \quad \text{for} \quad j = 1, 2, ..., n \tag{D5}$$

$$u_i \ge 0 \quad \text{for} \quad i = 1, 2, ..., m \tag{D6}$$

For the quadratic programming problem in (D7), it can be shown that the KKT conditions can be used to construct the linear system (D8)-(D10) with the complementary constraints (D11)-(D12).

$$\min\{f(\mathbf{x}) = \mathbf{c}^T \mathbf{x} + \frac{1}{2} \mathbf{x}^T \; \mathbf{H}\mathbf{x} : \mathbf{A}\mathbf{x} \le \mathbf{b}, \mathbf{x} \ge \mathbf{0}\} \tag{D7}$$

$$\mathbf{Ax} + \mathbf{y} = \mathbf{b} \tag{D8}$$

$$-\mathbf{Hx} - \mathbf{A}^{\mathbf{T}}\mathbf{u} + \mathbf{v} = \mathbf{c} \tag{D9}$$

$$\mathbf{x}, \mathbf{y}, \mathbf{u}, \mathbf{v} \geq \mathbf{0} \tag{D10}$$

$$\mathbf{x}^{\mathbf{T}}\mathbf{v} = 0 \tag{D11}$$

$$\mathbf{u}^{\mathbf{T}}\mathbf{y} = 0 \tag{D12}$$

where $\mathbf{u}$ and $\mathbf{v}$ are the Lagrangian multiplier vectors for $\mathbf{Ax} \leq \mathbf{b}$ and $\mathbf{x} \leq \mathbf{0}$ respectively, and $\mathbf{y}$ is the vector for the slack variables.

Further, let $\mathbf{M} = \begin{bmatrix} \mathbf{0} & -\mathbf{A} \\ \mathbf{A}^{\mathbf{T}} & \mathbf{H} \end{bmatrix}$, $\mathbf{q} = \begin{bmatrix} \mathbf{b} \\ \mathbf{c} \end{bmatrix}$, $\mathbf{w} = \begin{bmatrix} \mathbf{y} \\ \mathbf{v} \end{bmatrix}$ and $\mathbf{z} = \begin{bmatrix} \mathbf{u} \\ \mathbf{x} \end{bmatrix}$, then (D8)-(D12) can be simplified by (D13)-(D15).

$$\mathbf{w} - \mathbf{Mz} - \mathbf{1}z_0 = \mathbf{q} \tag{D13}$$

$$\mathbf{w}^{\mathbf{T}}\mathbf{z} = 0 \tag{D14}$$

$$\mathbf{w}, \mathbf{z} \geq \mathbf{0} \tag{D15}$$

z_0 is introduced as a dummy variable, so that an initial solution can be easily obtained by setting $z_0 = \max\{-q_i : 1 \leq i \leq m\}$, $\mathbf{z} = \mathbf{0}$ and $\mathbf{w} = \mathbf{q} + \mathbf{1}z_0$. The original problem in (D7) is solved when z_0 is driven to zero while satisfying (D13)-(D15).

If $\mathbf{H}$ is positive semi-definite, $f(\mathbf{x})$ is a convex function. Moreover, $g_i(\mathbf{x})$ are linear which are both concave and convex. Therefore, the resolution of (D13) subject to (D14) and (D15) will be optimal for the original problem in (D7).

Appendix E

Verification of Positive Semi-definiteness of H-Matrix

A $m \times m$ matrix $\mathbf{H}$ is positive definite if $\mathbf{x}^{\mathbf{T}}\mathbf{H}\mathbf{x} > 0$ and it is positive semi-definite if $\mathbf{x}^{\mathbf{T}}\mathbf{H}\mathbf{x} \geq 0$, for all $\mathbf{x}$. Positive definiteness can also be verified by the value of the determinant of $\mathbf{H}$, denoted by $\det(\mathbf{H})$. If $\det(\mathbf{H}) > 0$, it is positive definite, and if $\det(\mathbf{H}) \geq 0$, it is positive semi-definite.

1) $\mathbf{P}_{\mathbf{i}\to\mathbf{j}}$: Unidirectional transfer from L_i to L_j

According to (C4), $\mathbf{H}_{\mathbf{i}\to\mathbf{j}} = \dfrac{2k_i G_{ij}^*}{{w_m}^2}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$. In other words, $\det(\mathbf{H}_{\mathbf{i}\to\mathbf{j}})$ $= \dfrac{2k_i G_{ij}^*}{{w_m}^2}\begin{vmatrix} 1 & -1 \\ -1 & 1 \end{vmatrix} = \dfrac{2k_i G_{ij}^*}{{w_m}^2}\left[(1)^2 - (-1)^2\right] = 0$. $\mathbf{H}_{\mathbf{i}\to\mathbf{j}}$ is therefore positive semi-definite.

2) $\mathbf{P}_{\mathbf{j}\to\mathbf{i}}$: Unidirectional transfer from L_j to L_i

According to (C8), $\mathbf{H}_{\mathbf{j}\to\mathbf{i}} = \dfrac{2k_i G_{ji}^*}{{w_m}^2}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$. In other words, $\det(\mathbf{H}_{\mathbf{j}\to\mathbf{i}})$ $= \dfrac{2k_i G_{ji}^*}{{w_m}^2}\begin{vmatrix} 1 & -1 \\ -1 & 1 \end{vmatrix} = \dfrac{2k_i G_{ji}^*}{{w_m}^2}\left[(1)^2 - (-1)^2\right] = 0$. $\mathbf{H}_{\mathbf{j}\to\mathbf{i}}$ is therefore positive semi-definite.

3) $\mathbf{P}_{\mathbf{i}\leftrightarrow\mathbf{j}}$: Bidirectional transfer to and from L_i and L_j

According to (C12), $\mathbf{H}_{\mathbf{i}\leftrightarrow\mathbf{j}} = \dfrac{2k_i}{{w_m}^2}\begin{bmatrix} G_{ij}^* + G_{ji}^* & -(G_{ij}^* + G_{ji}^*) \\ -(G_{ij}^* + G_{ji}^*) & G_{ij}^* + G_{ji}^* \end{bmatrix}$. In other words,

$$\det(\mathbf{H}_{\mathbf{i}\leftrightarrow\mathbf{j}}) = \frac{2k_i}{{w_m}^2}\begin{vmatrix} G_{ij}^* + G_{ji}^* & -(G_{ij}^* + G_{ji}^*) \\ -(G_{ij}^* + G_{ji}^*) & G_{ij}^* + G_{ji}^* \end{vmatrix} = \frac{2k_i}{{w_m}^2}\left[(G_{ij}^* + G_{ji}^*)^2 - (G_{ij}^* + G_{ji}^*)^2\right] = 0 \ .$$

$\mathbf{H}_{\mathbf{i}\leftrightarrow\mathbf{j}}$ is therefore positive semi-definite.

References

ABARE [Australian Bureau of Agricultural and Resource Economics] (2005) ‘Infrastructure issues in the Hunter Valley coal supply chain’, *ABARE Report to the Australian Government Department of Industry, Tourism and Resources*, (Canberra: ABARE)

ABS [Australian Bureau of Statistics] (2006) ‘Year Book Australia 2006 (Ch. 16) - mining: profile of major minerals, oil and gas’. http://www.abs.gov.au, accessed June 2006

Al-Agtash, S.Y., and Al-Fahoum, A.A. (2005) ‘An evolutionary computation approach to electricity trade negotiation’, *Advances in Engineering Software*, **36**, (3), pp. 173-179

ARTC [Australian Rail Track Corporation] (2004) ‘$152 million upgrade for the Hunter Valley rail freight network’. http://www.artc.com.au, accessed June 2006

ARTC [Australian Rail Track Corporation] (2006) ‘About ARTC’. http://www.artc.com.au, accessed June 2006

Ayyub, B.M., and McCuen, R.H. (2003) ‘Probability, statistics, and reliability for engineers and scientists’ (Boca Raton, FL: Chapman & Hall)

Bagnall, A.J., and Smith, G.D. (2005) ‘A multiagent model of the UK market in electricity generation’, *IEEE Transactions on Evolutionary Computation*, **9**, (5), pp. 522-536

Bazaraa, M.S., Sherali, H.D., and Shetty, C.M. (1993) ‘Nonlinear programming: theory and algorithms’ (New York: Wiley)

Bellifemine, F., Poggi, A., and Rimassa, G. (1999) ‘JADE - A FIPA-compliant agent framework’, *Proceedings of the 4th International Conference and Exhibition on the Practical Application of Intelligent Agents and Multi-Agent Technology* (PAAM’99), London, UK, pp. 97-108

Bellifemine, F., Caire, G., Poggi, A., and Rrimassa, G. (2003) ‘JADE - a white paper’, *EXP in Search of Innovation*, **3**, (3), pp. 6-19

Bellifemine, F., Caire, G., Trucco, T., and Rimassa, R. (2003) ‘JADE programmer’s guide’. http://sharon.cselt.it, accessed April 2006

Böcker, J., Lind, J., and Zirkler, B. (2001) ‘Using a multi-agent approach to optimise the train coupling and sharing system’, *European Journal of Operational Research*, **131**, (2), pp. 242-252

Bond, A.H., and Gasser, L.G. (1988) ‘An analysis of problems and research in DAI’. In Bond, A.H., and Gasser, L.G. (eds.) ‘Readings in distributed artificial intelligence’ (San Mateo, CA: Morgan Kaufmann), pp. 3-35

Boyer, K.D. (1998) ‘Principles of transportation economics’ (Reading, MA: Addison-Wesley)

Brazier, F.M.T., Cornelissen, F., Gustavsson, R., Jonker, C.M., Lindeberg, O., Polak, B., and Treur, J. (2002) ‘A multi-agent system performing one-to-many negotiation for load balancing of electricity use’, *Electronic Commerce Research and Applications*, **1**, (2), pp. 208-224

Brucker, P., Burkard, R.E., and Hurink, J. (1990) ‘Cyclic schedules for r irregularly occurring events’, *Journal of Computational and Applied Mathematics*, **30**, (2), pp. 173-189

BTRE [Bureau of Transport and Regional Economics] (2003) ‘Rail infrastructure pricing: principles and practice’, *Report 109* (Canberra ACT: BTRE)

Burkard, R.E. (1986) ‘Optimal schedules for periodically recurring events’, *Discrete Applied Mathematics*, **15**, (2), pp. 167-180

Caire, G. (2001) ‘JADE tutorial: application-defined content languages and ontologies’. http://sharon.cselt.it, accessed April 2006

Campos J., and Cantos, P. (1999) ‘Regulating privatized rail transport’, *Policy Research Working Paper Series 2064* (Geneva: World Bank)

Cantos, P., and Maudos, J. (2001) ‘Regulation and efficiency: the case of European railways’, *Transportation Research Part A: Policy and Practice*, **35**, (5), pp. 459-472

Central Trains (2006) ‘Your journeys – stations’. http://www.centraltrains.co.uk, accessed July 2006

Covington, M.A. (1998): ‘Speech acts, electronic commerce, and KQML’, D*ecision Support Systems*, **22**, (3), pp. 203-211

Crompton, G., and Jupe, R. (2003) ‘Such a silly scheme: the privatisation of Britain's Railways 1992-2002’, *Critical Perspectives on Accounting*, **14**, (6), pp. 617-645

Cuppari, A., Guida, P.L., Martelli, M., Mascardi, V., and Zini, F. (1999) ‘Prototyping freight trains traffic management using multi-agent systems’, *Proceedings of the 1999*

International Conference on Information Intelligence and Systems (ICIIS’99), Bethesda MD, USA, pp. 646-653

Dia, H. (2002) ‘An agent-based approach to modelling driver route choice behaviour under the influence of real-time information’, *Transportation Research Part C: Emerging Technologies*, **10**, (5), pp. 331-349

Dodgson, J. (1994) ‘Access pricing in the railway system’, *Utilities Policy*, **4**, (3), pp. 205-213

DOTARS [Department of Transport and Regional Services] (2003) ‘History of Rail in Australia’. http://www.dotars.gov.au, accessed February 2003

Durfee, E.H., Lesser, V.R., and Corkill, D.D. (1989) ‘Trends in cooperative distributed problem solving’, *IEEE Transactions on Knowledge and Data Engineering*, **1**, (1), pp. 68-68

EC [The European Commission] (2006) ‘Rail transport and interoperability: overview’. http://europa.eu.int, accessed March 2006

ECMT [European Conference of Ministers of Transport] (1998) ‘Rail restructuring in Europe’ (Paris: OECD Publication Service)

ECMT [European Conference of Ministers of Transport] (2001) ‘Railway reform: regulation of freight transport markets’ (Paris: OECD Publication Service)

ECMT [European Conference of Ministers of Transport] (2005) ‘Railway reform & charges for the use of infrastructure’ (Paris: OECD Publication Service)

Ehtamo, H., Verkama, M., and Hamalainen, R.P. (1996) ‘On distributed computation of Pareto solutions for two decision makers’, *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, **26**, (4), pp. 498-503

Emorphia (2001) ‘FIPA-OS developers guide’. http://fipa-os.sourceforge.net, accessed April 2006

Faratin, P., Sierra, C., and Jennings, N.R. (1998) ‘Negotiation decision functions for autonomous agents’, *Robotics and Autonomous Systems*, **24**, (4), pp. 159-182

Faratin, P., Sierra, C., and Jennings, N.R. (2002) ‘Using similarity criteria to make issue trade-offs in automated negotiations’, *Artificial Intelligence*, **142**, (2), pp. 205-237

Finin, T., Fritzson, R., McKay, D., and McEntire, R. (1994) ‘KQML as an agent communication language’, *Proceedings of the 3rd International Conference on Information and Knowledge Management* (CIKM'94), Gaithersburg MD, USA, pp. 456-463

FIPA [Foundation for Intelligent Physical Agents] (2005) ‘Specifications’. http://www.fipa.org, accessed April 2006

Franklin, S., and Graesser, A. (1996) ‘Is it an agent, or just a program?: a taxonomy for autonomous agent’, *Proceedings of the 3rd International Workshop on Agent Theories, Architectures, and Languages* (ATAL’96), Budapest, Hungary, pp. 21-26

Ghosh, S. (1999) ‘Fundamental issues in intelligent transportation systems’, *Proceedings of the IME - Part F: Journal of Rail and Rapid Transit*, **213**, (2), pp. 125-131

Gibson, S., Cooper, G., and Ball, B. (2002) ‘Developments in transport policy: the evolution of capacity charges on the UK rail network’, *Journal of Transport Economics and Policy*, **36**, (2), pp. 341-354

Gibson, S. (2003) ‘Allocation of capacity in the rail industry’, *Utilities Policy*, **11**, (1), pp. 39-42

Godward, E. (1998) ‘The privatization of British Rail 1994-1997’, *Proceedings of the IME - Part F: Journal of Rail and Rapid Transit*, **212**, (3), pp. 191-200

Goodman, C.J., Siu, L.K., and Ho, T.K. (1998) ‘A review of simulation models for railway systems’, *Proceedings of the IEE International Conference on Developments in Mass Transit Systems*, London, UK, pp. 80-85

Grassie, S.L. (2005) ‘Rolling contact fatigue on the British railway system: treatment’, *Wear*, **258**, (7), pp. 1310-1318

Hallé, S., and Chaib-draa, B. (2005) ‘A collaborative driving system based on multiagent modelling and simulations’, *Transportation Research Part C: Emerging Technologies*, **13**, (4), pp. 320-345

Harris, N.G. (1999) ‘Competitive strategies for railways in the UK: a corporate perspective’, *Transport Reviews*, **19**, (2), pp. 191-202

Hill, R.J. (1995) ‘Electric railway traction. IV. signalling and interlockings’, *Power Engineering Journal*, **9**, (4), pp. 201-206

Hillier, F.S., and Lieberman, G.J. (1995) ‘Introduction to mathematical programming’ (New York: McGraw-Hill)

Ho, T.K., Norton, J.P., and Goodman, C.J. (1997) ‘Optimal traffic control at railway junctions’, *IEE Proceedings - Electric Power Applications*, **144**, (2), pp. 140-148

Holland, J.H. (1992) ‘Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence’ (Cambridge, MA: MIT Press)

Ilic, M., Hsieh, E., and Ramanan, P. (2003) ‘Transmission pricing of distributed multilateral energy transactions to ensure system security and guide economic dispatch’, *IEEE Transactions on Power Systems*, **18**, (2), pp. 428-434

Jennings, N.R., and Bussmann, S. (2003) ‘Agent-based control systems: why are they suited to engineering complex systems?’, *IEEE Control Systems Magazine*, **23**, (3), pp. 61-73

Jennings, N.R., Norman, T.J., Faratin, P., O’Brien, P., and Odgers, B. (2000) ‘Autonomous agents for business process management’, *Applied Artificial Intelligence*, **14**, (2), pp. 145-189

Jennings, N.R., Norman, T.J., Faratin, P., O’Brien, P., Odgers, B., and Alty, J.L. (2000) ‘Implementing a business process management system using ADEPT: a real-world case study’, *Applied Artificial Intelligence*, **14**, (5), pp. 421-463

Jensen, A. (1998) ‘Competition in railway monopolies’, *Transportation Research Part E: Logistics and Transportation Review*, **34**, (4), pp. 267-287

Johansson, P., and Nilsson, J.E. (2004) ‘An economic analysis of track maintenance costs’, *Transport Policy*, **11**, (3), pp. 277-286

Krishna, V., and Ramesh, V.C. (1998) ‘Intelligent agents for negotiations in market games - part 1: model’, *IEEE Transactions on Power Systems*, **13**, (3), pp. 1103-1108

Krishna, V., and Ramesh, V.C. (1998): ‘Intelligent agents for negotiations in market games - part 2: application’, *IEEE Transactions on Power Systems*, **13**, (3), pp. 1109-1114

Krovi, R., Graesser, A.C., and Pracht, W.E. (1999) ‘Agent behaviours in virtual negotiation environments’, *IEEE Transactions on Systems, Man, and Cybernetics - Part C: Applications and Reviews*, **29**, (1), pp. 15-25

Liu, J., and You, J. (2003) ‘Smart shopper: an agent-based web-mining approach to internet shopping’, *IEEE Transactions on Fuzzy Systems*, **11**, (2), pp. 229-237

Luo, X., Jennings, N.R., Shadbolt, N., Leung, H., and Lee, J.H. (2003) ‘A fuzzy constraint based model for bilateral, multi-issue negotiations in semi-competitive environments’, *Artificial Intelligence*, **148**, (1), pp. 53-102

McCarthy, J. (2004) ‘What is artificial intelligence?’. http://www-formal.stanford.edu/jmc, accessed April 2006

Minsky, M.L., and Papert, S. (1988): ‘Perceptrons: an introduction to computational geometry' (Cambridge, MA: MIT Press)

Mizutani, F. (1999) 'An assessment of the Japan railway companies since privatization: performance, local rail service and debts', *Transport Reviews*, **19**, (2), pp. 117-139

Moore, S.A. (1999) 'KQML & FLBC: contrasting agent communication languages', *Proceedings of the 32nd Annual Hawaii International Conference on System Sciences* (HICSS-32), Maui HI, USA, [CD-ROM] pp. 1-10

Mulley C., and Nelson, J.D. (1997) 'Interoperability and transport policy: the impediments to interoperability in the organisation of trans-European transport systems', *Journal of Transport Geography*, **7**, (2), pp. 93-104

Murata, S., and Goodman, C.J. (1998) 'Optimally regulating disturbed metro traffic with passenger inconvenience in mind', *Proceedings of the International Conference on Developments in Mass Transit Systems*, London, UK, pp. 86-91

Nachtigall, K. (1996) 'Periodic network optimization with different arc frequencies', *Discrete Applied Mathematics*, **69**, (1), pp. 1-17

Nachtigall, K., and Voget, S. (1996) 'A genetic algorithm approach to periodic railway synchronization', *Computers & Operations Research*, **23**, (5), pp. 453-463

Network Rail (2004) 'Type and volume of traffic per train'. http://www.networkrail.co.uk, accessed March 2006

Network Rail (2004) '2003 Business Plan - Section 6: Operational Planning'. http://www.networkrail.co.uk, accessed March 2006

Network Rail (2006): 'Stations run by Network Rail'. http://www.networkrail.co.uk, accessed June 2006

Nilsson, J.E. (2002) 'Towards a welfare enhancing process to manage railway infrastructure access', *Transportation Research Part A: Policy and Practice*, **36**, (5), pp. 419-439

Northern Rail (2006) 'Network Map'. http://www.northernrail.org, accessed July 2006

Ossowski S., Hernández, J.Z., Belmonte, M., Fernández, A., García-Serrano, A., Pérez-de-la-Cru, J., Serrano, J., and Triguero F. (2005) 'Decision support for traffic management based on organisational and communicative multiagent abstractions', *Transportation Research Part C: Emerging Technologies*, **13**, (4), pp. 272-298

Pacific National (2006) 'Coal'. http://www.pacificnational.com.au, accessed June 2006

Papadimitriou, C.H., and Steiglitz, K. (1998) 'Combinatorial optimization: algorithms and complexity' (Mineola, NY: Dover Publications)

Poslad, S., Buckle, P., and Hadingham, R. (2000) 'The FIPA-OS agent platform: open source for open standards', *Proceedings of the 5th International Conference and Exhibition on the Practical Application of Intelligent Agents and Multi-Agent Technology* (PAAM'2000), Manchester, UK, pp. 355-368

Pokahr, A., and Braubach, L. (2005) 'Jadex User Guide'. http://vsis-www.informatik.uni-hamburg.de/projects/jadex, accessed July 2006

Rao, A.S., and Georgeff, M. (1995) 'BDI agents: from theory to practice', *Proceedings of the 1st International Conference on Multi-agent Systems* (ICMAS'95), San Francisco, USA, pp. 312-319

Rossetti, R.J.F., Bordini, R.H., Bazzan, A.L.C., Bampi, S., Liu, R. and Van Vliet D. (2002) 'Using BDI agents to improve driver modelling in a commuter senario', *Transportation Research Part C: Emerging Technologies*, **10**, (5), pp. 373-398.

Rothengatter, W. (1991) 'Deregulating the European railway industry: Theoretical background and practical consequences', *Transportation Research Part A: General*, **25**, (4), pp. 181-191

Russell, S.J., and Norvig, P. (1994) 'Artificial intelligence: a modern approach' (Englewood Cliffs, NJ: Prentice Hall)

Shaw, J. (2000) 'Competition, regulation and the privatisation of British Rail' (Aldershot, England: Ashgate)

Shaw, J. (2001) 'Competition in the UK passenger railway industry: prospects and problems', *Transport Reviews*, **21**, (2), pp. 195-216

Shaw, M.J., Gardner, D.M., and Thomas, H. (1997) 'Research opportunities in electronic commerce', *Decision Support Systems*, **21**, (3), pp. 149-156

Sim, K.M., and Wong, E. (2001) 'Toward market-driven agents for electronic auction', *IEEE Transactions on Systems, Man and Cybernetics - Part A: Systems and Humans*, **31**, (6), pp. 474-484

Sim, K.M., and Choi, C.Y. (2003) 'Agents that react to changing market situations', *IEEE Transactions on Systems, Man and Cybernetics - Part B: Cybernetics*, **33**, (2), pp. 188-201

Sim, K.M., and Wang, S.Y. (2004) 'Flexible negotiation agent with relaxed decision rules', *IEEE Transactions on Systems, Man and Cybernetics - Part B: Cybernetics*, **34**, (3), pp. 1602-1608

Smith, R.G. (1980) ‘The contract net protocol: high-level communication and control in a distributed problem solver’, *IEEE Transactions on Computers*, **C-29**, (12), pp. 1104-1113

Teodorović, D. (2003) ‘Transport modeling by multi-agent systems: a swarm intelligence approach’, *Transportation Planning and Technology*, **26**, (4), pp. 289-312

TransPennine Express (2006) ‘Network & stations’. http://www.tpexpress.co.uk, accessed July 2006

Virgin Trains (2006) ‘Network Map’. http://www.virgintrains.co.uk, accessed July 2006

Walpole, R.E., Myers, R.H., and Myers, S.L. (1998) ‘Probability and statistics for engineers and scientists’ (Upper Saddle River, NJ: Prentice Hall)

Watkins, A.E., Scheaffer, R.L., and Cobb, G.W. (2004) ‘Statistics in action: understanding a world of data’ (Emeryville, CA: Key Curriculum Press)

Watson, R. (2001) ‘The effect of railway privatization on train planning: a case study of the UK’, *Transport Reviews*, **21**, (2), pp. 181-193

Wikipedia (2006) ‘Network Rail – Railway stations’. http://en.wikipedia.org, accessed June 2006

Willmott, S., Calisti, M., Faltings, B., Macho-Gonzalez, S., Belakhdar, O., and Torrens, M. (2000) ‘CCL: expressions of choice in agent communication’, *Proceedings of the 4th International Conference on Multi-Agent Systems* (ICMAS'2000), Boston MA, USA, pp. 325-332

Winograd, G.I., and Kumar, P.R. (1996) ‘The FCFS service discipline: stable network topologies, bounds on traffic burstiness and delay, and control by regulators’, *Mathematical and Computer Modelling*, **23**, (11), pp. 115-129

Wooldridge, M.J. (1997) ‘Agent-based software engineering’, *IEE Proceedings - Software Engineering*, **144**, (1), pp. 26-37

Wooldridge, M.J., and Jennings, N.R. (1999) ‘Software engineering with agents: pitfalls and pratfalls’, *IEEE Internet Computing Magazine*, **3**, (3), pp. 20-27

Wooldridge, M.J. (2002) ‘An introduction to multiagent systems’ (Chichester: Wiley).

Wu, J.H., and Nash, C. (2000) ‘Railway reform in China’, *Transport Reviews*, **20**, (1), pp. 25-48

Xie, R., Chen, H., and Nash, C. (2002) ‘Migration of railway freight transport from command economy to market economy: the case of China’, *Transport Reviews*, **22**, (2), pp. 159-177

Xue, X., Schmid, F., and Smith, R.A. (2002) ‘An introduction to China’s rail transport - part 1: history, present and future of China’s railways’, *Proceedings of the IME - Part F: Journal of Rail and Rapid Transit*, **216**, (3), pp. 153-163

Xue, X., Schmid, F., and Smith, R.A. (2002) ‘An introduction to China’s rail transport - part 2: urban rail transit systems, highway transport and the reform of China’s railways’, *Proceedings of the IME - Part F: Journal of Rail and Rapid Transit*, **216**, (3), pp. 165-174

Yaskawa, S., and Sakata, A. (2003): ‘The application of intelligent agent technology to simulation’, *Mathematical and Computer Modelling*, **39**, (9), pp. 1083-1092

Zambonelli, F., and Luck, M. (2005) ‘The agent hell: a scenario of worst-practices in agent-based software engineering’, *IEEE Potentials*, **24**, (2), pp. 23-26

Zhang, L., and Levinson, D. (2005) ‘Balancing efficiency and equity of ramp meters’, *Journal of Transportation Engineering*, **131**, (6), pp. 477-481