

Dynamic road pricing incorporating dynamic traffic assignment

Zhang, Chun Rong

2008

Zhang, C. R. (2008). Dynamic road pricing incorporating dynamic traffic assignment. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/13077>

<https://doi.org/10.32657/10356/13077>



**NANYANG
TECHNOLOGICAL
UNIVERSITY**

**DYNAMIC ROAD PRICING INCORPORATING DYNAMIC
TRAFFIC ASSIGNMENT**

ZHANG CHUNRONG

SCHOOL OF CIVIL & ENVIRONMENTAL ENGINEERING

2008

DYNAMIC ROAD PRICING INCORPORATING DYNAMIC TRAFFIC ASSIGNMENT

ZHANG CHUNRONG

2008

Dynamic Road Pricing Incorporating Dynamic Traffic Assignment

Zhang Chunrong

School of Civil & Environmental Engineering

A thesis submitted to the Nanyang Technological University
in fulfilment of the requirement for the degree of
Doctor of Philosophy

2008

ACKNOWLEDGEMENTS

I would like to sincerely thank my supervisor, Dr. Lum Kit Meng, for his support throughout this research, especially when I was in the hard time of my life.

My life in Nanyang Technological University has been fruitful. I have to thank Dr. Mak Ching Long for being such a good exercise mate. Thanks are also extended to my friends: Ms. Ng Choo Hiang, Mr. Liew Kai Liang, Ms. Liu Li, and Mr. Li Shoujie for their helpful minds.

Special thanks go to Prof. Huang Hai Jun for the discussion on his original dynamic traffic assignment model and to Prof. Michael C. Ferris for his enlightments on the complementarity problems.

Finally, I wish to express my gratitude to my parents for their soul encouragement.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	ii
TABLE OF CONTENTS	iii
LIST OF TABLES	viii
LIST OF FIGURES	ix
LIST OF ABBREVIATIONS.....	xii
ABSTRACT	xiii
CHAPTER ONE INTRODUCTION.....	1
1.1 BACKGROUND	1
1.2 PROBLEM STATEMENT.....	3
1.3 OBJECTIVES AND SCOPE OF THE RESEARCH	4
1.4 PERSPECTIVES AND CONTRIBUTIONS.....	5
1.5 STRUCTURE OF THE THESIS	7
CHAPTER TWO ROAD PRICING AND TRAFFIC	
ASSIGNMENT REVIEW.....	9
2.1 INTRODUCTION	9
2.2 ROAD PRICING IN THEORY	9
2.2.1 Static Road Pricing Studies.....	10

2.2.2	Dynamic Road Pricing Studies	13
2.2.3	Social Acceptability Studies.....	13
2.3	ROAD PRICING IN PRACTICE.....	14
2.3.1	City Center Congestion Pricing	15
2.3.2	Toll Ring	16
2.3.3	Single Facility Congestion Pricing	16
2.3.4	Area-wide Congestion Pricing.....	18
2.4	WARDROP'S PRINCIPLE AND STATIC TRAFFIC ASSIGNMENT.....	19
2.4.1	Wardrop's Principle.....	20
2.4.2	Mathematical Programming Formulation.....	21
2.4.3	Variational Inequality Problems.....	24
2.4.4	Nonlinear Complementarity Problems	26
2.5	DYNAMIC WARDROP'S PRINCIPLE AND DYNAMIC TRAFFIC ASSIGNMENT.....	27
2.5.1	Dynamic User Equilibrium.....	28
2.5.2	FIFO Constraints.....	29
2.5.3	Mathematical Programming Approach.....	29
2.5.4	Optimal Control Theory Approach	31
2.5.5	Variational Inequality Theory Approach.....	31
2.5.6	Combined Approach	33
2.6	SUMMARY	34

CHAPTER THREE OPTIMAL TOLL FORMULATION

AND ALGORITHMS.....36

3.1	INTRODUCTION	36
3.2	MPEC FORMULATION OF OPTIMAL TOLL	36

3.3	REVIEW OF ALGORITHMS FOR MPEC PROBLEMS	38
3.3.1	NLP Techniques Based Method.....	39
3.3.2	Smoothing Method.....	39
3.3.3	Penalty and Interior Point Based Method	40
3.3.4	Implicit Programming Method	40
3.4	HEURISTICS ALGORITHMS	41
3.4.1	Evolutionary Algorithm Based Method (DE-Toll)	41
3.4.2	Pattern Search Algorithm (PS-Toll)	46
3.5	NUMERICAL EXAMPLE	48
3.6	SUMMARY	54

CHAPTER FOUR DYNAMIC TRAFFIC ASSIGNMENT

	MODEL AND ALGORITHM	55
4.1	INTRODUCTION	55
4.2	DTA MODEL DESCRIPTION.....	55
4.2.1	Problem Formulation	56
4.2.2	Link Travel Time.....	58
4.2.3	Representation of Queuing	58
4.2.4	Link Flow Conservation	60
4.2.5	Path Cost	60
4.3	EXISTENCE OF EQUILIBRIUM	61
4.4	MODEL MODIFICATION.....	64
4.5	ROUTE/TIME SWAPPING ALGORITHM	66
4.6	NUMERICAL EXAMPLE FOR “ROUTE/TIME” SWAPPING ALGORITHM.....	68
4.7	A PRIME PROJECTION-BASED ALGORITHM	72
4.7.1	Algorithm Description	72

4.7.2	The Selection of Tolerance ϵ	74
4.7.3	The Selection of Step Length τ	75
4.8	PROJECTION-BASED ALGORITHM WITH MSA AND EXAMPLE.....	77
4.9	ALGORITHMS COMPARISON	81
4.10	SUMMARY	82

CHAPTER FIVE DYNAMIC ROAD PRICING MODEL.....83

5.1	INTRODUCTION	83
5.2	MPEC FORMULATION	83
5.3	NUMERICAL EXAMPLE	86
5.3.1	Application of DE-Toll	86
5.3.2	Application of PS-Toll	91
5.3.3	Sensitivity Analysis.....	92
5.4	SUMMARY	93

CHAPTER SIX MULTI-OBJECTIVE PROBLEM FOR

DYNAMIC ROAD PRICING95

6.1	INTRODUCTION	95
6.2	PROBLEM SETTING	95
6.3	EVOLUTIONARY-BASED APPROACHES FOR MULTI-OBJECTIVE PROBLEM.....	97
6.3.1	Approaches That Use Aggregating Functions	97
6.3.2	Other Not Pareto-based Approaches.....	99
6.3.3	Pareto-based Approaches	102
6.3.4	NSGA-II Algorithm	103
6.3.5	A Fast Non-dominated Sorting	105

6.3.6 Crowded-comparison Operator (\prec_n) 107

6.4 MULTI-OBJECTIVE OPTIMAL TOLL DESIGN AND EXAMPLE 109

6.5 SUMMARY 111

CHAPTER SEVEN CONCLUSIONS AND FUTURE

RESEARCH 113

7.1 PRACTICAL ROAD PRICING 113

7.2 THEORETICAL APPROACH TO SOLVING DYNAMIC ROAD
PRICING INCORPORATING DTA 114

7.3 SOME THOUGHTFUL RECOLLECTIONS 116

7.4 LIMITATIONS AND RECOMMENDATIONS 116

REFERENCES 121

APPENDIX: MATHEMATICAL NOTATIONS

LIST OF TABLES

Table 2.1	Overview of Norway’s Toll Ring (Source: May and Sumalee, 2003)	17
Table 2.2	Public Attitudes towards Toll Rings (Source: Odeck and Brathen, 2002)	18
Table 3.1	Elastic Demand Function of Each OD Pair	50
Table 3.2	Parameters of Link Travel Time Function	50
Table 3.3	Parameters for DE-Toll	51
Table 3.4	Comparison of NLPEC, DE-Toll and PS-Toll	51
Table 4.1	Link Inflow Rate	65
Table 4.2	The Resulting Queue and Travel Time	66
Table 4.3	Link Exit Flow	66
Table 4.4	Link Parameters	68
Table 4.5	Path List	69
Table 5.1	Link Parameters for MPEC	87
Table 5.2	Sensitivity Analysis of Pre-determined	

Parameters93

LIST OF FIGURES

Figure 1.1	Structure of the Thesis	7
Figure 2.1	Illustration of Marginal Cost Pricing Theory	10
Figure 3.1	Differential Evolutionary Algorithm for MPEC ...	42
Figure 3.2	An Example of Mutation in Two Dimensional Setting	44
Figure 3.3	An Example of the Crossover Process for $D = 6$ Parameters	44
Figure 3.4	Pattern Search Algorithm for MPEC.....	46
Figure 3.5	Pattern and Mesh Illustration.....	47
Figure 3.6	The Nine Node Network.....	48
Figure 3.7	Total Travel Time Delay Using DE-Toll	52
Figure 3.8	Total Travel Time Delay using PS-Toll with Starting Point 0	53
Figure 3.9	Total Travel Time Delay Using PS-Toll with Starting Point 5	53

Figure 4.1 **An Example of Linear Interpolation**64

Figure 4.2 **Grid Network**68

Figure 4.3 **Convergence of Route/time Swapping Algorithm**.....70

Figure 4.4 **Path Inflow Rate and Travel Cost by Using Route/time Swapping Algorithm (Path 1 and 2)**71

Figure 4.5 **Path Inflow Rate and Travel Cost by Using Route/time Swapping Algorithm (Path 7)**.....71

Figure 4.6 **Queue Length by Using Route/time Swapping Algorithm**.....72

Figure 4.7 **Value of Convergence Indicator and $W(f)$ under too Small Tolerance ϵ** 75

Figure 4.8 **Convergence Indicator with Different Step Length (1)**76

Figure 4.9 **Convergence Indicator with Different Step Length (2)**77

Figure 4.10 **Convergence Indicator by Using Projection-based Algorithm with MSA**79

Figure 4.11 **Path Inflow Rate and Travel Cost by Using Projection-based Algorithm with MSA (Path 1**

	and 2).....	79
Figure 4.12	Path Inflow Rate and Travel Cost by Using Projection-based Algorithm with MSA (Path 7)	80
Figure 4.13	Queue Length by Using Projection-based Algorithm with MSA	80
Figure 5.1	Queue Length on the Network without Tolling...	87
Figure 5.2	Convergence of Total Queue Volume by Using DE-Toll	88
Figure 5.3	Queue Length under Tolling	89
Figure 5.4	Link Inflow Rate before and after Tolling (Link 1 and 7).....	90
Figure 5.5	Convergence of Total Queue Volume by Using PS-Toll from Starting Point (0,0,0).....	91
Figure 5.6	Convergence of Total Queue Volume by Using PS-Toll from Starting Point (1,1,1).....	92
Figure 6.1	An Example of Pareto Optimum	96
Figure 6.2	Schematic of VEGA Selection.....	100
Figure 6.3	Procedure of NSGA-II	105
Figure 6.4	An Example of Non-dominated Sorting	106
Figure 6.5	A Fast Non-dominated Sorting Method	106

Figure 6.6	An Example of Crowding-distance Calculation	107
Figure 6.7	Discrete Results of Bi-objective Road Pricing	110
Figure 6.8	Pareto Front Generated by NSGA-II	111

LIST OF ABBREVIATIONS

ALS	Area Licensing Scheme
CBD	Central Business District
CTM	Cell Transmission Model
DE	Differential Evolutionary
DTA	Dynamic Traffic Assignment
DUE	Dynamic User Equilibrium
ERP	Electronic Road Pricing
FIFO	First-In-First-Out
GA	Generic Algorithm
GADS	Genetic Algorithm and Direct Search
GPS	Generalized Pattern Search
ITS	Intelligent Transportation Systems
MADS	Mesh-Adaptive Direct Search
MCP	Mixed Complementarity Problem
MFCQ	Mangasarian Fromovitz Constraint Qualification
MOGA	Multiple Objective Genetic Algorithm
MPEC	Mathematical Program with Equilibrium Constraints
MSA	Method of Successive Averages
NCP	Nonlinear Complementarity Problem
NLP	Nonlinear Programming
NLPEC	Non-linear Problem with Equilibrium Constraints

NPGA	Niched Pareto Genetic Algorithm
NSGA	Non-dominated Sorting Genetic Algorithm
O-D	Origin-Destination
PAES	Pareto Archived Evolutionary Strategy
PIPA	Penalty Interior Point Algorithm
PS	Pattern Search
RZ	Restricted Zone
SPEA	Strength Pareto Evolutionary Algorithm
SQP	Sequential Quadratic Programming
VEGA	Vector Evaluated Genetic Algorithm
VI	Variational Inequality

ABSTRACT

Road pricing has been widely advocated as an efficient instrument to alleviate traffic congestion. Current static pricing schemes are designed for long term transportation planning as traffic conditions are assumed not to vary dramatically. When dealing with short time period, especially during the peak period, such assumption cannot be valid due to variable conditions of traffic over the whole network. Thus, pricing should be dynamic to accommodate the dynamic traffic evolution. This research aims to develop a theoretical dynamic road pricing model incorporating dynamic traffic assignment.

The framework for the optimal road pricing problem incorporating traffic assignment on a network was a Stackelberg game. Road pricing schemes have to be announced firstly and the travellers' behaviour were assumed to follow the concept of Wardrop's user equilibrium. This problem can also be expressed as a Mathematical Problem with Equilibrium Constraints (MPEC). Traffic assignment model under fixed tolling level was expressed as the constraint. The toll level was expected to have an effect on travel time. After reviewing various algorithms for the MPEC problems, three different algorithms were tested for a static road pricing problem: Differential Evolutionary (DE), Pattern Search (PS), and Non-linear Problem with Equilibrium Constraints (NLPEC) solver. For this well-defined problem, NLPEC solver showed to be the most efficient in computation time for a given

hypothetical nine-node network.

To study the dynamic road pricing scheme, promising dynamic traffic evolution should allow for route choosing, departure time selection, and queue size on the network to vary with the tolling level. A suitable dynamic traffic assignment (DTA) model from Huang and Lam (2002) was adopted as the base model. A new procedure to compute some inter-variables was proposed to represent traffic evolution more appropriately. The behaviour of travellers followed the dynamic extension of Wardrop's user equilibrium, which is based on predictive travel time. First-in-first-out (FIFO) conditions in the dynamic situations are satisfied. Different from the original algorithm from Huang and Lam (2002), a new algorithm based on the projection method was proposed resulting in a smoother traffic flow.

The DTA model was set as the equilibrium constraint of the dynamic road pricing problem which was in the form of a dynamic MPEC. For this dynamic road pricing problem with a single objective, Differential Evolutionary (DE) algorithm and Pattern Search (PS) algorithm were tested and DE algorithm was found to be superior as compared to PS algorithm. To evaluate the objective of the MPEC model in the process of applying DE and PS algorithm, the resulting parametric DTA problem was solved using the algorithm based on the projection method proposed in this thesis. In this research, the amount of toll depends on the length of stay on the tolled link.

Lastly, multi-objective dynamic road pricing model applying the Non-dominated Sorting Genetic Algorithm (NSGA-II) was successfully illustrated to show how the problem would behave when confronted by the decision maker in reality. The numerical example suggested that when

considering queue optimization, the decision maker can take into account the tradeoffs with revenue generation.

The results for the dynamic road pricing model and multi-objective dynamic road pricing model were based on the assumptions of the DTA model involved. These assumptions include simplified link travel time function, idealistic travel behaviour modeling, pre-determined input parameters and illustration on a small size network.

CHAPTER 1 INTRODUCTION

1.1 BACKGROUND

With globalization of businesses, demands for travel are growing substantially. The ever-growing demand has outpaced the supply of new roads, thus, leading to serious traffic congestion which has become one of the dominant characteristics of large urban centres. Effects of traffic congestion are often intolerable and frustrating. It causes not only disruptions to traffic flow, delay to passengers and environmental pollution, but also greatly impact on the efficiency of the economic systems. Berg and Bhatt (1998) reported that the 1994 annual cost of urban traffic delay and associated excess fuel consumption was US\$ 53 billion for fifty large and medium-size urban areas in USA. British economists estimated that congestion on the M25 London motorway alone costs around £ 15 billion a year (LTA, 1996). Singapore, being an island city, is also not spare from such a congestion problem. Particularly, the car population had increased 53% from 273,127 in 1990 to 417103 in 2004 (LTA, 1996 and 2005), and the number of vehicular trips grew annually by 7%; that is, from 2.7 million trips per day in 1981 to about 7 million trips per day in 1996.

Traditional way of tackling the congestion problem is to build more road infrastructures to cater to rising demand. However, in some urban areas, it may not be feasible to enlarge the existing road network. One alternative way to manage congestion is through fiscal measures, such as road pricing which was proposed more than eighty years ago (Pigou, 1920; Walters, 1961; Vickrey, 1969). Economists assume that motorists would not always consider the associated costs such as congestion, emission and accidents generated by them. With this in mind, the concept of marginal cost pricing theory is to charge every motorist the difference between marginal social and marginal private costs in the form of toll. This is also referred to as the first-best tolling schemes which can cause a serious public outcry and a political acceptability problem. In reality, practical road pricing schemes tend to practice the second-best solutions.

Authorities are concerned with more specific objectives and public acceptance problems when implementing road pricing. These specific objectives can be to control total traffic on a network or to generate revenue. The first practical implementation of a road pricing scheme in Singapore was the Area Licensing Scheme (ALS) (Holland and Watson, 1978). The main objective was to reduce traffic entering into the central business district. However, Norway's toll ring schemes are the main source of income to finance transport infrastructure. The pioneering work on Electronic Road Pricing (ERP) in Hong Kong had demonstrated the importance of public acceptance as it had made use of different charging cordons and screen lines which were too complicated to implement without causing the invasion of privacy (Harrison, 1986).

Theoretically, the earliest work on road pricing by Pigou (1920) was based on economic theory of supply and demand equilibrium. It is the first-best pricing scheme which can be used to impose tolls on all links. However, this pricing scheme was rather impractical to adopt as it ignored the complexity and imperfect network environment in reality. Researchers are more interested to consider road pricing on a whole network with some untolled links. This is often called the second-best pricing scheme. In the network environment, the concept of Wardrop's user equilibrium (Wardrop, 1952) is normally adopted to represent the route choice behaviour of the motorists. With Wardrop's user equilibrium, motorists are expected to choose their routes so as to minimise the generalized travel cost they incurred. The Wardropian equilibrium condition is a special case of Nash's equilibrium (Nash, 1951). Marchand (1968) may be the first to analyze the road pricing problem on a two-link network with one untolled link.

The problem of determining an optimal road pricing charge can be framed as a Stackelberg game (Stackelberg, 1952) that involves the leader and the follower in the game. The leader has the propriety to choose his/her decisions, while the follower is assumed to follow Nash's equilibrium given the leader's decision. The leader has the ability to anticipate the reaction of the follower to the decision the leader had imposed onto the follower. For the optimal road pricing problem, the leader can be assumed as the authority responsible for the setting

of the toll level. The follower can be the motorists whose travel behaviour follows Wardrop's equilibrium. The authority always knows exactly how the toll level has impacted on the motorists' route choice behaviour. As the motorists (the follower) are following the equilibrium conditions, there is no necessity to explicitly write out the objective function. In essence, the Stackelberg game is a special case of the Mathematical Program with Equilibrium Constraints, MPEC (Luo et al., 1996).

As static road pricing and route choice decision do not consider the variable traffic conditions on the network, especially during the peak period, dynamic road pricing in a dynamic network has attracted some research interests. Arnott et al. (1990) studied the dynamic road pricing problem on a one Origin-Destination (O-D) pair network with two parallel routes. Travellers were allowed to change their departure time given the toll level. Wie and Tobin (1998) investigated the optimal dynamic road pricing problem on a continuous network. The Dynamic Traffic Assignment (DTA) which represents dynamic route choice behaviour was formulated as an optimal control problem. Recently, Joksimovic et al. (2005) explored dynamic road pricing on a dynamic network. However, the toll levels were set as discrete variables and brute-force method was adopted in their algorithm design.

1.2 PROBLEM STATEMENT

Hitherto, there are very few research studies relating to dynamic road pricing modelling on a dynamic network. Part of the reason is that there is no general mathematical model to express dynamic network. To describe the dynamic network, more issues have to be carefully considered such as dynamic route choice behaviour, first-in-first-out (FIFO) condition, link travel time function, and possible queuing on the network. Besides, if the dynamic road pricing problem is seen as an extension of the static road pricing problem, the computation algorithm is another main concern due to the complexity of DTA. This research is to model dynamic road pricing problem theoretically as a MPEC on a dynamic network with queue. A revised DTA from previous

research (Huang and Lam, 2002) is used and a new algorithm applying projection method is proposed to tackle this DTA. Road pricing is treated as a continuous variable. The resulting MPEC is solved by applying the Differential Evolutionary (DE) algorithm. In this research, the amount of toll depends on the length of stay on the tolled link. For the road pricing policy maker, multi-objectives are not rare in reality. Sometimes, these objectives may conflict with each other. Decision makers are faced with the trade-off between these objectives. This research uses the Non-dominated Sorting Genetic Algorithm-II (NSGA-II) to demonstrate the trade-off between various objectives quantitatively.

1.3 OBJECTIVES AND SCOPE OF THE RESEARCH

The main intent of this research is to develop a model for the dynamic road pricing problem on a dynamic network with queues. It helps the authority to understand how the dynamic pricing level has impacted onto the whole network performance. It also gives the trade-off between the different objectives of the pricing schemes. The specific objectives are:

- To investigate the computation algorithms for optimal static road pricing problem;
- To revise a DTA model from a previous research (Huang and Lam, 2002) and to propose a new algorithm for the DTA
- To develop a mathematical approach to design a dynamic road pricing problem with DTA;
- To develop a quantitative approach to support the decision making on the trade-off between different design objectives, and
- To test the algorithms developed on a small size network and to show the potential application to the real-world problem.

The scope of this research is limited to the use of a second-best dynamic road pricing scheme; that is, applicable to a general network with some untolled links. The main assumptions in modelling include the static and dynamic extension of Wardrop's equilibrium and point queuing on the dynamic network. Motorists are also assumed to have perfect information about the network. The model also assumes a single user class and a single travel mode.

1.4 PERSPECTIVES AND CONTRIBUTIONS

In perspective, this research firstly reviews the traffic assignment and road pricing in theory and practice. Secondly, a static optimal road pricing designed to incorporate traffic assignment on a network is formulated as a MPEC. In this formulation, link travel time has continuous derivatives over traffic flow. To solve this well-defined problem, three different algorithms: DE, Pattern Search (PS) and Non-linear Problem with Equilibrium Constraints (NLPEC) solver are used to show their respective efficiency. It shows that when the problem is well defined, NLPEC can be the most efficient in terms of computation time.

Thirdly, a dynamic DTA on a network with queue from a previous research (Huang and Lam, 2002) is employed and revised. This model allows travellers to select their own departure time. It also simulates the queue phenomenon on the network. Different to the "route/time swapping" algorithm in the literature, a new algorithm applying the projection method is proposed. Simulation shows that this new algorithm can represent smoother traffic flows.

Fourthly, dynamic optimal road pricing model is then formulated on this dynamic network with queue. Toll link is predetermined and toll level is set as a continuous variable dependent upon time dimension. DE shows its superiority for this not well-defined MPEC. At last, NSGA-II algorithm is adopted to show the multi-objective road pricing problem encountered in reality.

In essence, the computational efficiencies and difficulties encountered using the various mathematical algorithms (DE, NLPEC and NSGA-II) to solve the

various relaxations (static versus dynamic traffic assignment, static versus dynamic road pricing and queues) of a well-defined problem help to demonstrate the complexities of deriving solutions to the problems at hand. This would, in fact, contribute to the broadening of knowledge and make potential applications of dynamic road pricing under various constraints to be mathematically feasible.

In short, the contributions can be stated as follows:

- Different algorithms including DE, Pattern Search and NLPEC solver are applied to solve the static road pricing problem. The pros and cons of these algorithms in solving the static road pricing problem are highlighted. In essence, NLPEC is found to be the fastest algorithm.
- A DTA model on a network with queue from a previous research (Huang and Lam, 2002) is adopted as an equilibrium constraint in the dynamic road pricing problem development. A new procedure to compute some inter-variables is proposed.
- A new algorithm based on the projection method is proposed to solve the revised DTA model. A numerical example shows that this algorithm can generate smoother traffic flow than the original route/time swapping algorithm proposed by Huang and Lam (2002).
- This study establishes that dynamic road pricing problem incorporating DTA can be formulated as a MPEC problem. DE approach is demonstrated to be superior as compared to PS algorithm through a numerical example.
- A multi-objective road pricing problem is set up for consideration. This study also demonstrates that NSGA-II can give satisfactory results in searching for solutions. The importance of considering multi-objectives is shown through a numerical example.

1.5 STRUCTURE OF THE THESIS

The aim of this thesis is to develop dynamic road pricing models on a network. On one hand, it involves the description of traffic distribution over the network. On the other hand, it is restricted to more than one designing objective in reality. This thesis is structured into three parts which is shown in Figure 1.1.

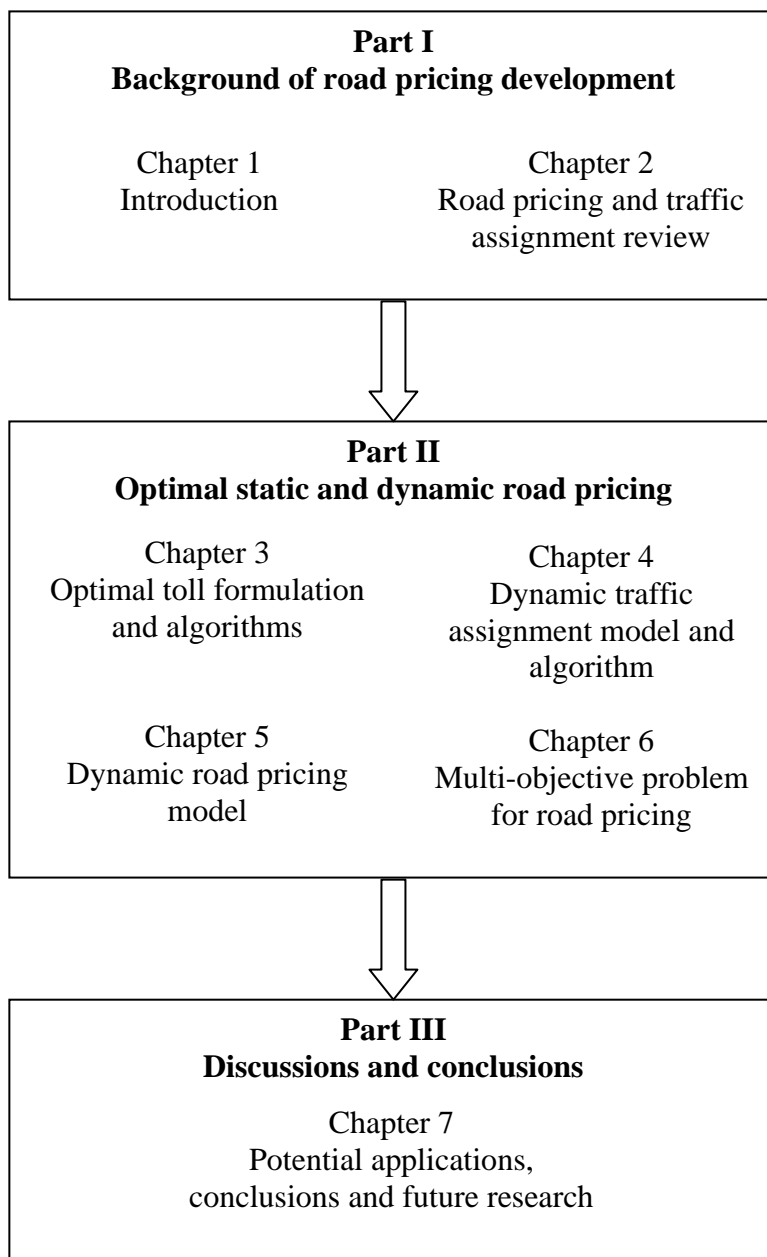


Figure 1.1 Structure of the Thesis

Part I which comprises Chapters 1 and 2 concentrates on the background of road pricing development. Chapter 2 reviews road pricing from the economic theory to the practical implementation in the world. As the modelling of road pricing is based on a network, development of static and dynamic traffic assignment is also introduced.

Part II comprising Chapters 3 to 6 focuses on the development of static and dynamic road pricing models. Road pricing on a network is formulated as a MPEC problem. As this is a tremendously difficult problem, a review of algorithms for MPEC is provided. Different algorithms including DE, PS (Pattern Search), and NLPEC solver are tested with a numerical example in Chapter 3. In order to derive road pricing designed for considering traffic impact on a network, a suitable DTA model is needed to describe traffic evolution. In Chapter 4, a DTA model on a network with queue from previous researchers is revised. A new algorithm is proposed for this DTA model. Setting this DTA model as an equilibrium constraint, the dynamic road pricing model is structured as a MPEC problem in Chapter 5. For the multi-objective road pricing problem encountered in reality, Chapter 6 introduces NSGA-II algorithm and shows the application of the algorithm with a numerical example.

The last part of the thesis (Part III) as given in Chapter 7 is dedicated to the discussions and conclusions. It also concludes this thesis and suggests some future research issues.

CHAPTER 2 ROAD PRICING AND TRAFFIC ASSIGNMENT REVIEW

2.1 INTRODUCTION

This chapter begins the review of the road pricing issues in the theoretical modelling methods. Road pricing originates from the economic concept of marginal cost theory. It is designed to toll on a single link. Later, researchers consider the tolling strategy on the entire network. The toll level is treated as a static variable. After that, both of these two tolling strategies are extended to a dynamic environment where the toll level is set to vary with time. After the social acceptability of pricing policy is discussed, the implementation of road pricing in practice is reviewed. The second part of this chapter focuses on static traffic assignment and dynamic traffic assignment as these will be extensively employed in the following chapters. Wardrop's (1952) principle to describe travellers' behaviour and response to road pricing will be addressed. The whole pricing problem incorporating Wardrop's principle written as a Mathematical Problem with Equilibrium Constraints (MPEC) will also be discussed.

2.2 ROAD PRICING IN THEORY

The theoretical road pricing studies can fall into 3 main categories: static road pricing, dynamic road pricing and acceptability studies. In static road pricing studies, toll level does not vary with time. Generally, it assumes that the traffic distribution is static for the long term. While in dynamic road pricing, the traffic distribution is treated as time-dependent, and toll level is set to vary with time to accommodate this situation. Social acceptability studies focus on the public acceptance of pricing policy.

2.2.1 Static Road Pricing Studies

The classical road pricing model is based on the marginal cost pricing theory which can be traced back to Pigou (1920). It is shown in Figure 2.1 (cited in Yang and Huang (1998)).

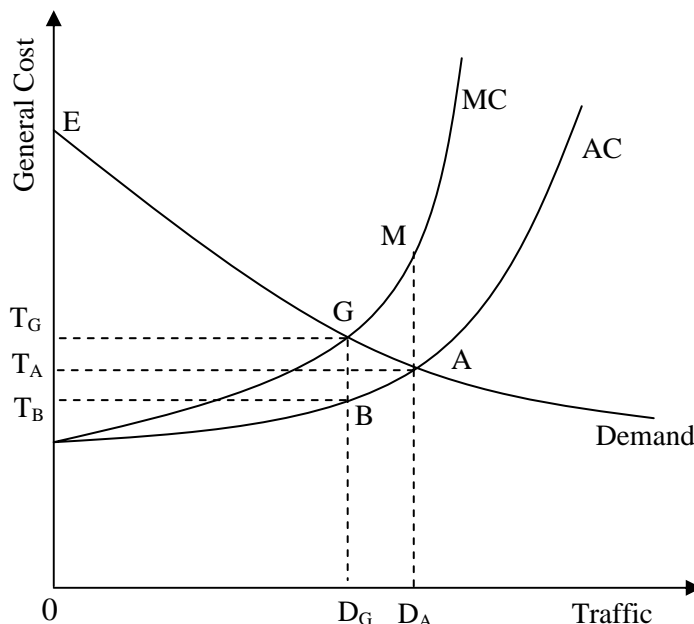


Figure 2.1 Illustration of Marginal Cost Pricing Theory

In Figure 2.1, AC represents the average cost of congestion at each level of demand. MC represents the marginal cost which is always larger than AC because of the assumption: each traveller not only incurs the average private cost himself, but also bears an additional cost as the one who makes the road more congested. This difference between the AC curve and the MC curve reflects the economic costs of congestion at that demand (Button, 1993). The systematic equilibrium is at point “G” and the optimal flow is D_G . However, the actual flow tends to be D_A as travellers are assumed to be non-cooperative and do not take into account other additional costs such as those associated with the environmental emission, noise and etc. From Figure 2.1, the actual demand D_A is excessive because the D_A^{th}

user is only enjoying a benefit of $\overline{D_A A}$, but is imposing the cost $\overline{D_A M}$. The additional traffic beyond the optimal D_G can be seen to be generating a cost equals to the area $D_A M G D_G$, but only the benefit equals to the area $D_A A G D_G$ resulting a deadweight welfare loss equals to the area AMG . A demand level lower than D_G is also not optimal because the potential consumer surplus gained from trip-making is not fully exploited. Therefore, the optimal toll is equal to \overline{BG} . Under this toll charge, the economic benefit given by the area $BGET_B$ (*total user benefit minus total social cost*) will be maximised.

Past research studied pricing problem based on a simple network with only one O-D pair and one route as reviewed by Alfa (1986). Yang and Huang (1998) checked the marginal cost pricing theory by applying elastic demand to a road network. They found that when an additional user is added to the network, traffic distribution will change to effect a change in the optimal toll. They implied that the detailed network structure needs to be considered when determining the optimal toll.

The classical marginal cost pricing theory has two limitations. First, it assumes perfect equilibrium exists between demand and cost function. For the relaxation of using the demand function, Walters (1961) linked the speed-flow relationships with congestion pricing theory. Vickrey (1993) and Downs (1993) proposed a trial-and-error method to derive the optimal pricing instead of using the demand function. The basic idea of using a trial-and-error method is that given a trial set of link tolls, traffic distribution can be determined and link tolls are then derived to satisfy some objectives. The new tolls derived are used as the next trial for iteration. Following their directions, Li (1999, 2002) used traffic count data to estimate the optimal pricing for a practical implementation purpose. In Li's (2002) paper, a bisection iterative procedure was proposed to estimate the congestion toll based on the speed-flow relationship for a single road. Yang et al. (2004) pointed out that for a general network, this bisection iterative procedure may not converge. They

proposed a new trial-and-error implementation scheme based on the method of successive averages and applied it onto the general network. The toll is levied on each link of the network to achieve the maximum net economic benefit.

The other limitation of the marginal cost pricing theory is that it is usually called the first-best tolling strategy as travellers on each link over the entire network will endure marginal cost. As tolling on the whole network may not be practical due to politics or public acceptance, research have concentrated on the second-best road pricing strategy which means that toll will only be applied to certain roads or set of links on the network. Early studies on the second-best pricing problem only employed a two-route network, where an untolled alternative road is parallel to a toll road. This simple situation was studied by Marchand (1968), Braid (1996), and Liu and McDonald (1998, 1999). For the general network, Yang and Lam (1996) formulated the optimal pricing problem as a bi-level problem for fixed demand. Optimal pricing strategy was determined to satisfy some objectives from the decision maker while making the total traffic distribution followed a static traffic equilibrium. Hearn and Yildirim (1999) followed the same way to investigate optimal pricing satisfying different objectives. More recent studies have been done by Verhoef (2002) to incorporate general network with elastic traffic demand. Lo and Szeto (2005) suggested that static road pricing could worsen the congestion problem on a hyper-congestion network. By assuming fixed demand, Stewart (2007) studied link-toll problem to meet the minimal of total network travel cost with stochastic user equilibrium.

All the previous research has focused on theoretical modelling of static road pricing problem. As discussed above, they attempted to describe the dynamics of traffic over network as realistic as possible. These effort lies in the extension of one O-D pair to general network, fixed demand function to elastic demand function, determined user equilibrium to stochastic user equilibrium, etc. However, practical application is seldom reported.

2.2.2 Dynamic Road Pricing Studies

In contrast to static road pricing studies, there were very few research studies on dynamic road pricing. Dynamic pricing assumes that pricing varies with time due to changing traffic conditions on a network.

Arnott et al. (1990) considered a simple network with only one O-D pair and parallel routes. Departure time and route choice were taken into account simultaneously. In the model, queuing phenomenon was simulated and it was shown that dynamic pricing strategy can reduce queuing by altering departure times. Later, Carey and Srinivasan (1993) extended the marginal cost theory to the dynamic situations. The traffic demand was assumed to vary over time. Using the Kuhn-Tucker conditions, they derived dynamic pricing based on the level of congestion and the rate of increase or decrease of congestion. The dynamic traffic distribution was modelled by mathematical programming. Yang and Huang (1997) and then Yang and Meng (1998) considered a dynamic pricing model for a bottleneck with elastic demand by applying the optimal control theory. Employing the same optimal control theory, Wie and Tobin (1998) investigated time continuous road pricing by applying the dynamic traffic assignment technique. The dynamic optimal road pricing was derived by applying the Pontryagin's necessary conditions. The previous research has focused on first-best pricing scheme which is to toll on each link. This may encounter difficulty in reality. Further, mathematical programming method and optimal control theory have their limitations. They cannot model some traffic phenomenon explicitly such as queuing.

2.2.3 Social Acceptability Studies

Different from studies to design optimal road pricing based on mathematical simulation of network traffic distribution, a lot of research work has focused on the

social acceptability of pricing policy. This critical issue emphasizes on investigating the relationship between public acceptance and other factors such as: social norms, knowledge/information about options, perceived effectiveness of the proposed pricing systems, equity/fairness, etc. Hau (1992a) explained why road pricing always encountered difficulty in reality by using welfare analysis for different types of travellers. All parties except the government are seemed worse off under road pricing. Hau (1992b) also investigated different price charging mechanisms including manual tollbooths, cordon pricing, Area Licensing Schemes, and Electronic Road Pricing. The conclusion was that Electronic Road Pricing is superior to others if budgets allow. Jones (2002) listed a few design parameters when considering equity in relation to road pricing. Such design parameters include the basis of charging, the area covered by the charge, the time periods covered by the charge, discounts or exemptions, linkages to other transport charges, etc. According to the analysis of Schade and Schlag (2003), social norm, personal outcome expectations and perceived effectiveness are positively connected to the acceptability issue. Applying psychological theories, Schade and Baum (2007) revealed that persons with a strong conviction about a definite introduction of road pricing exhibit much more positive acceptance than persons who are less convinced about the introduction. Continuous educating and informing travellers of the benefit of road pricing can improve public acceptance as suggested by Ieromonachou et al., (2006). Different tolling strategies including bridge tolls and cordon tolls were examined by Gupta et al. (2006) to evaluate impact of tolling on traffic redistribution, long-term location choices and traveller revenue for Austin, Texas.

2.3 ROAD PRICING IN PRACTICE

Other than looking only at the economic efficiency of transportation system from an economic viewpoint, practical road pricing scheme design has been explored in a wide range of applications. Lindberg (1995) identified three distinct objectives

for the implementation of road pricing schemes, namely, (a) to reduce congestion, (b) to improve environmental quality, and (c) to generate revenue. Various road pricing systems has been developed. According to Small and Gomez-Ibanez (1998), they can be generally categorized as follows:

2.3.1 City Centre Congestion Pricing

Area Licensing Scheme (ALS) is the first field implementation of congestion pricing scheme introduced by the Singapore government in 1975. The ALS covered the more congested parts of the Central Business District (CBD), called the Restricted Zone (RZ). An ALS license had to be purchased to gain access into the RZ. Initially, the license fee was S\$3 per day (US\$1 \cong S\$1.6) and it has crept to S\$5 per day in 1980. After a major review in 1989, the daily license fee was reduced back to S\$3.

The initial operating hours were from 7.30 to 9.30 am daily, except for Sundays and public holidays. However, this pricing scheme had triggered additional congestion immediately after 9.30 am in the RZ. Three weeks after the initial operation, the operating hours were extended by 45 minutes to 10.15 am. After 1989, to accommodate the increase in vehicle population, operating hours covering the evening peak time from 4.30 to 7 pm were added. On 3rd January 1994, the operating hours were further extended to cover the period from 10.15 am to 4.30 pm on weekdays.

According to Menon (2000), passenger car traffic entering the RZ during the morning peak period in 1992 was approximately half of the level in 1975 before the ALS was introduced. Public transport's share for work trips increased from 33% in 1973 to 67% in 1992. Seik (1998) considered the ALS successful and claimed that there was no significant impact on businesses inside the RZ.

Early ALS system was labour intensive and not flexible. In 1998, the electronic road pricing system (ERP) was implemented to replace the ALS. The ERP system makes it possible to charge differently for different vehicles entering the RZ at different times of the day. The operation of ERP system is reviewed quarterly to adjust the charge rate to maintain a targeted traffic speed on expressways and roads inside the RZ at 45-65 km/hr and 20-30 km/hr, respectively. After implementation, traffic volume entering the CBD decreased by 10-15% as compared to the ALS system (Chin, 2002).

2.3.2 Toll Ring

The road pricing implementation in Norway has taken the form of a toll ring for decades (May and Sumalee, 2003). The main objective is to raise funds for road projects. In 1986, the Bergen toll ring was introduced. Four years later, Oslo also introduced an urban toll ring to finance a new tunnel under the city centre. In 1991, the world's first toll ring with automated toll collection was introduced in Trondheim. A summary of these 3 toll ring schemes is given in Table 2.1.

As the original objective of the toll ring schemes was to generate revenue, traffic was slightly reduced (around 6-7% in Bergen, 3-4% in Oslo, and 10% in Trondheim during the charged periods). The public attitudes towards toll rings one year before and after implementation are summarized in Table 2.2. As shown in the table, attitudes towards Oslo and Trondheim were strongly negative at their introduction.

2.3.3 Single Facility Congestion Pricing

Under the single facility congestion pricing scheme, motorists are allowed to use a particular facility such as an expressway, a bridge, or a tunnel by paying a toll. For

instance, “State Route 91 Express Lanes” was the first field congestion pricing scheme in the United States. It is located in the freeway median between the State Route 91/55 junction in Anaheim and the Orange/Riverside County Line. The facility provides two extra lanes in each direction, isolated from the immediately adjacent freeway lanes. Along the 10-mile length, there are no intermediate entrances and exits. Heavy vehicles are prohibited from using the toll lanes. When it was opened in December, 1995, the same toll was charged throughout the four-hour peak period. High occupancy vehicles with three or more occupants (HOV3+) are free to use the 91 Express Lanes. Since September 1997, toll level has been re-evaluated hour by hour during the peak period. HOV3+ has to pay 50% of the published toll beginning January 1998.

Table 2.1 Overview of Norway’s Toll Ring (Source: May and Sumalee, 2003)

City	Bergen	Oslo	Trondheim
Population ('000)	213	456	138
Starting Date	Jan, 1986	Feb, 1990	Oct, 1991
Number of Toll Stations	7	19	22
Charging Regime	Uniform Charge	Uniform Charge	Peak and Off Peak Charge
Entry Charge for Small Vehicle ¹ (Norwegian Krone(NOK))	10	15	15 (for all period for manual payment ²)
Charging period and discount	Weekday 6am-10pm Discount for monthly subscriptions	All days, All hours Discount for prepaid tickets	Weekday 6am-6pm Discount for monthly subscriptions
Annual gross revenues (NOK, millions)	156	1046	168
Annual operating costs (NOK, millions)	30	103	17

¹Heavy vehicles are charged double the price ²For prepayment of 6000 NOK, 9 NOK between 6am-10am and 6 NOK between 10am-6pm; for prepayment of 3000 NOK, 10.5 NOK between 6am-10am and 7.5 NOK between 10am-6pm; for prepayment of 1000 NOK, 12 NOK between 6am-10am and 9 NOK between 10am-6pm.

Table 2.2 Public Attitudes towards Toll Rings (Source: Odeck and Brathen, 2002)

	Negative Percentage of Respondents	
	1 year before	1 year after
Bergen (1986)	54%	37%
Oslo (1990)	70%	64%
Trondheim (1991)	72%	48%

Sullivan (2000) did a study to evaluate the impacts of State Route 91 Express Lanes. It was generally concluded that an increasing number of motorists were willing to pay the toll charges to enjoy benefits such as reduced travel time, improved driving comfort and etc. The survey findings also showed that a majority of the motorists were from low and middle-income categories. The benefits of the express lanes were enjoyed widely at all income levels.

2.3.4 Area-wide Congestion Pricing

One of the proposals considering an area-wide congestion pricing scheme was at the Randstad region (including Amsterdam, Rotterdam, The Hague, and Utrecht, plus part of the province of Noord-Brabant) in the Netherlands during the late 1980s. It involved multiple cordon systems with 140 charging points and time-varying tolls. The main objectives were to reduce traffic during the peak hours and to finance new infrastructures. After failing to obtain enough support from the Parliament, the government in 1990 proposed a more conventional tolling scheme to solely finance the road infrastructures. In't Veld (1991) provided a detailed analysis on the reasons for the political failure.

In 1992, contemplating on the potential congestion caused by too much land being used for the toll plaza to implement the conventional tolling scheme, the proposal was rejected. The Ministry proposed a reduced scheme known as “peak charging” that involved a system of supplementary licensing for motorists using the main road network during the peak periods. The fee was set at about US\$2.85 per day in

1992 price during the morning rush hour from 6 to 10 am. The proposal was set aside during the 1994 government election and came back to the table in the name of “congestion charging” which would be a system of electronic toll cordons around the four main cities in the Randstad area. The “congestion charging” scheme was implemented in 2001.

In 2001, the “congestion charging” scheme became a political debate in the Netherlands. Objections came strongly from those who needed to travel by car during the proposed charging time period. The government is now considering a new system “Mobimeter” that charges motorists per kilometre of travel. In this system, every vehicle must be fitted with a device, known as mobimeter, which registers the number of kilometres driven and toll payable. The system is expected to be fully operational by 2006.

However, congestion charging in London has proven to be a popular success. After an 18-month public consultation, the congestion scheme came into operation in February 2003. The congestion charge is an area license which covers a 21 km² charging zone known as “Central London”. Driving or parking a vehicle within the charging zone is levied a £5.00 per day (increased to £8.00 in July 2005). Leape (2006) investigated the impact of congestion charging from different angles. Vehicles entering into central London dropped 27 percent after introducing the charging scheme. Revenues for congestion charging are channelled back on bus network and road safety improvements. This may also partly explain why net revenues have been far below than the predicted which was around £68 million for year 2003-2004 and £97 million for year 2004-2005 (Leape, 2006).

2.4 WARDROP'S PRINCIPLE AND STATIC TRAFFIC ASSIGNMENT

Wardrop's principle is the critical assumption in theoretical traffic assignment modelling. It describes route choice behaviour of all travellers.

2.4.1 Wardrop's Principle

Wardrop's first principle:

The journey times on all the routes actually used are equal, and less than those which would be experienced by a single vehicle on any unused route.

Wardrop's second principle:

The average journey time is a minimum.

The first principle which is often referred to as the user-equilibrium condition is most likely to happen in practice, especially for the long term. The proper explanation is that travellers tend to settle down into an equilibrium condition in which no one can reduce his travel time by choosing a new route. There are many different ways to mathematically express Wardrop's principle. Optimisation equivalence (Beckmann et al, 1956), variational inequality (Smith, 1979; Dafermos, 1980), nonlinear complementarity (Aashtiani and Magnanti, 1981), and gap function (Hearn et al, 1984) are such variations. A summary of such formulations can be found in Nagurney (1993) and Patriksson (1994, 2004).

Notations

A traffic network is represented by a graph $G(N_1, A)$ where N_1 is the set of nodes and A is the set of links. Each origin-destination (OD) pair is denoted by $(p, q) \in C$. R_{pq} is the set of simple routes connecting OD pair (p, q) . Let h_{pqr} denotes the volume of traffic on route $r \in R_{pq}$ between OD pair (p, q) , c_{pqr} is the travel cost on route r connecting OD pair (p, q) .

Thus, Wardrop's first principle can be expressed as:

$$h_r > 0 \Rightarrow c_r = \pi_{pq}, \quad r \in R_{pq}, \quad (p, q) \in C \quad (2.1a)$$

$$h_r = 0 \Rightarrow c_r \geq \pi_{pq}, \quad r \in R_{pq}, \quad (p, q) \in C \quad (2.1b)$$

where π_{pq} is the minimal route cost of OD pair (p, q) . Let \mathbf{h} be the vector of h_r with dimension of $|R|$ and $\mathbf{c} : R_+^{|R|} \mapsto R^{|R|}$ the vector-valued function of route cost vector \mathbf{h} . Let π be the vector of π_{pq} with dimension of $|C|$. Demand function is $\mathbf{g} : R^{|C|} \mapsto R_+^{|C|}$, a vector-valued function of π . Let $\Gamma \in R^{|R| \times |C|}$ denote the route-OD pair incidence matrix with the element γ_{rk} expressed as:

$$\gamma_{rk} = \begin{cases} 1 & \text{if route } r \text{ joins OD pair } k = (p, q) \in C \\ 0 & \text{otherwise} \end{cases}$$

Thus, the demand constraints can be written as:

$$\Gamma^T \mathbf{h} = \mathbf{g}(\pi) \quad \mathbf{h} \geq 0, \pi \geq 0 \quad (2.2)$$

The conditions for user equilibrium may be summarized by Equations (2.1) and (2.2).

2.4.2 Mathematical Programming Formulation

The first mathematical program formulation of user equilibrium may be due to Prager (1954) and Beckmann et al. (1956) as suggested by Patricksson (1994). The optimisation problem is based on link flows: $\mathbf{f}_{pq} = (f_{apq})$, which can be written as a function of h_{pqr} :

$$f_{apq} = \sum_{a \in A} \delta_{pqra} h_{pqr}, \quad \forall (p, q) \in C, a \in A \quad (2.3)$$

in which

$$\delta_{pqra} = \begin{cases} 1 & \text{if route } r \in R_{pq} \text{ uses link } a \\ 0 & \text{otherwise} \end{cases} \quad \forall a \in A, r \in R_{pq}, (p, q) \in C \quad (2.4)$$

Let Δ^T denotes the link-route incidence matrix with the element of δ_{pqra} , total link flow vector $\mathbf{f} = (f_a)$, and f_a is given by:

$$f_a = \sum_{(p,q) \in C} f_{apq} \quad \forall a \in A \quad (2.5)$$

The compact notation of Equations (2.3) to (2.5) can be expressed as:

$$\mathbf{f} = \Delta \mathbf{h} \quad (2.6)$$

Assume the route costs are additive of links costs on the corresponding route, i. e.:

$$c_{pqr}(\mathbf{h}) = \sum_{a \in A} \delta_{pqra} t_a(f_a) \quad \forall r \in R_{pq}, (p, q) \in C \quad (2.7)$$

$t_a(\cdot)$ is the link travel time which is usually denoted as the link cost. A more efficient form of Equation (2.7) is:

$$c(\mathbf{h}) = \Delta^T \mathbf{t}(\mathbf{f}) \quad (2.8)$$

The fixed demand case

When the demand between each OD pair (p, q) is fixed, $\mathbf{g}(\pi)$ in Equation (2.2) can be denoted as $\mathbf{g}(\pi) = \mathbf{d} \geq 0$.

【 Theorem 2.1 】 [Patriksson (1994)]. Assume the travel time function $t_a : R_+ \mapsto R_{++}$ is positive and continuous for each a and the network is strongly

connected, then the user equilibrium conditions (2.1) to (2.2) are equivalent to the first-order optimality conditions of the mathematical programming problem stated below:

$$\min T(\mathbf{f}) = \sum_{a \in A} \int_0^{f_a} t_a(s) ds \quad (2.9a)$$

Subject to

$$\Gamma^T \mathbf{h} = \mathbf{d} \quad (2.9b)$$

$$\mathbf{h} \geq 0 \quad (2.9c)$$

$$\mathbf{f} = \Delta \mathbf{h} \quad (2.9d)$$

The fixed demand model was first formulated by Dafermos and Sparrow (1969).

The elastic demand case

【 Theorem 2.2 】 [Patriksson (1994)]. Assume network is strongly connected, travel time function $t_a : R_+ \mapsto R_{++}$ is positive and continuous for each a , and demand function $g_{pq} : R_+ \mapsto R_+$ is nonnegative, continuous and strictly decreasing for each $(p, q) \in C$. The user equilibrium conditions (2.1) to (2.2) are equivalent to the first-order optimality conditions of the mathematical programming problem stated below:

$$\min T(\mathbf{f}, \mathbf{d}) = \sum_{a \in A} \int_0^{f_a} t_a(s) ds - \sum_{(p, q) \in C} \int_0^{d_{pq}} g_{pq}^{-1}(s) ds \quad (2.10a)$$

Subject to

$$\Gamma^T \mathbf{h} = \mathbf{d} \quad (2.10b)$$

$$\mathbf{h} \geq 0 \quad (2.10c)$$

$$\mathbf{f} = \Delta \mathbf{h} \quad (2.10d)$$

$$\mathbf{g}(\boldsymbol{\pi}) = \mathbf{d} \geq 0 \quad (2.10e)$$

The model with elastic demand was first formulated by Beckmann et al.(1956).

2.4.3 Variational Inequality Problems

Let $X \subseteq R^n$ be a nonempty, closed and convex set, and $F : X \mapsto R^n$ a continuous mapping on X . The Variational Inequality (VI) problem is to find an $\mathbf{x}^* \in X$ such that:

$$F(\mathbf{x}^*)^T (\mathbf{x} - \mathbf{x}^*) \geq 0, \quad \forall \mathbf{x} \in X \quad (2.11)$$

The fixed demand case

【Theorem 2.3】 [Patriksson (1994)]. Assume a network is strongly connected, demand d_{pq} is nonnegative for each $(p, q) \in C$ and route cost c_{pqr} is positive and continuous for each $r \in R_{pq}$ and $(p, q) \in C$, then user equilibrium conditions are equivalent to the variational inequality problem of finding an $\mathbf{h}^* \in H$ such that

$$\mathbf{c}(\mathbf{h}^*)^T (\mathbf{h} - \mathbf{h}^*) \geq 0 \quad \forall \mathbf{h} \in H \quad (2.12)$$

where $H = \{\mathbf{h} \mid \Gamma^T \mathbf{h} = \mathbf{d}\}$

If one assumes the route costs are additive, then the variational inequality problem can be expressed by link flows as:

$$\mathbf{t}(\mathbf{f}^*)^T (\mathbf{f} - \mathbf{f}^*) \geq 0 \quad \forall \mathbf{f} \in F \quad (2.13)$$

where $F = \{\mathbf{f} \mid \Gamma^T \mathbf{h} = \mathbf{d}, \mathbf{h} \geq 0, \mathbf{f} = \Delta \mathbf{h}\}$

The first model of variational inequality formulation with fixed demand was attributed to Smith (1979).

The elastic demand case

【Theorem 2.4】 [Patriksson (1994)]. Assume the network is strongly connected, demand function g_{pq} is nonnegative and continuous for each $(p, q) \in C$ and route cost c_{pqr} is positive and continuous for each $r \in R_{pq}$ and $(p, q) \in C$, then the user equilibrium conditions are equivalent to the variational inequality problem of finding an $(\mathbf{h}^*, \boldsymbol{\pi}^*) \in R_+^{|R|+|C|}$ such that

$$\begin{bmatrix} \mathbf{c}(\mathbf{h}^*) - \Gamma \boldsymbol{\pi}^* \\ \Gamma^T \mathbf{h}^* - \mathbf{g}(\boldsymbol{\pi}^*) \end{bmatrix}^T \left(\begin{pmatrix} \mathbf{h} \\ \boldsymbol{\pi} \end{pmatrix} - \begin{pmatrix} \mathbf{h}^* \\ \boldsymbol{\pi}^* \end{pmatrix} \right) \geq 0 \quad \forall (\mathbf{h}, \boldsymbol{\pi}) \in H_d \quad (2.14)$$

where $H = \{(\mathbf{h}, \boldsymbol{\pi}) \mid \Gamma^T \mathbf{h} = \mathbf{g}(\boldsymbol{\pi})\}$. If demand function $\mathbf{g}(\boldsymbol{\pi})$ is invertible, Equation (2.14) can be rewritten as:

$$\mathbf{c}(\mathbf{h}^*)^T (\mathbf{h} - \mathbf{h}^*) - \mathbf{g}^{-1}(\mathbf{d}^*)^T (\mathbf{d} - \mathbf{d}^*) \geq 0 \quad \forall (\mathbf{h}, \mathbf{d}) \in H_d \quad (2.15)$$

The variational inequality model with elastic demand was first formulated by Dafermos (1982).

2.4.4 Nonlinear Complementarity Problems

Let $F : R_+^n \mapsto R^n$ be continuous. The Nonlinear Complementarity Problem (NCP) is to find a $\mathbf{x}^* \in R^n$ such that:

$$F(\mathbf{x}^*)^T \mathbf{x}^* = 0 \quad (2.16a)$$

$$F(\mathbf{x}^*) \geq \mathbf{0} \quad (2.16b)$$

$$\mathbf{x}^* \geq \mathbf{0} \quad (2.16c)$$

Considering the variational inequality formulation (2.16), it is a nonlinear complementarity problem if we let $\mathbf{x} = (\mathbf{h}, \pi)$ and

$$F(\mathbf{x}) = \begin{pmatrix} \mathbf{c}(\mathbf{h}) - \Gamma \pi \\ \Gamma^T - \mathbf{g}(\pi) \end{pmatrix}$$

For a network with many OD pairs, the enumeration of all path flow \mathbf{h} is prohibitive. A multicommodity formulation can be adopted to remove the enumeration of all paths. Redefine the link flow variable as f_{ij} for the directed link (i, j) . Denote \mathbf{f}_k as the vector of (f_{ij}) for commodity $k \in C$, \mathbf{f} represents the vector of (\mathbf{f}_k) ; $\mathbf{t} = (\mathbf{t}_k)$, where $\mathbf{t}_k = (t_{ik})$, t_{ik} is the minimum cost (or time) to deliver commodity k from node i . For each pair $(k, i), k \in C, i \in N$, travel demand d_{ik} is a function of the minimum cost vector t and $\mathbf{d}_k = (d_{ik})$. There are two sets of equilibrium conditions. The first is the flow conservation of commodity k at node i which can be written as:

$$\sum_{j \in W_i} f_{ijk} - \sum_{j \in V_i} f_{jik} = d_{ik}, \quad \forall i \in N, k \in C \quad (2.17)$$

where $W_i = \{j \mid (i, j) \in A\}$ and $V_i = \{j \mid (j, i) \in A\}$ represent the set of nodes that heading and tailing node at i , respectively.

Define the node-link incidence matrix: $A = (a_{ib})$ with

$$a_{ib} = \begin{cases} 1 & \text{if } i \text{ is the origin node of link } b \\ -1 & \text{if } i \text{ is the destination node of link } b \\ 0 & \text{otherwise} \end{cases} \quad \forall b \in A, i \in N$$

(2.18)

Then, the compact form of Equation (2.17) can be obtained as:

$$A\mathbf{f}_k = \mathbf{d}_k(\mathbf{t}) \quad \forall k \in C \quad (2.19)$$

The second equilibrium condition is that positive link flow of commodity k along link (i, j) follows Wardrop's first principle.

$$0 \leq c_{ij}(\mathbf{f}) + t_{jk} - t_{ik} \perp f_{ijk} \geq 0 \quad \forall (i, j) \in A, k \in C \quad (2.20)$$

in which Equation (2.19) and (2.20) define a NCP problem. Additive cost function can be expressed as $c_a(\mathbf{f}) = c_a(\mathbf{f}_a)$ and $\mathbf{f}_a = \sum_{k \in C} f_{ak}$. Constant demand means

$\mathbf{d}_k(\mathbf{t}) = \mathbf{d}_k(\text{constant})$.

The NCP is equivalent to a variational inequality defined on $\mathbf{x} \geq 0$.

2.5 DYNAMIC WARDROP'S PRINCIPLE AND DYNAMIC TRAFFIC ASSIGNMENT

As static traffic assignment fails to represent the temporal traffic flow behaviour, dynamic traffic assignment tries to simulate all traffic phenomena by including the time dimension. Link or route travel time will not be treated as a constant value, but varies with time. Consequently, user equilibrium extended for dynamic situations should be established. By considering traffic propagation in more detail, some fundamental issues have to be raised, such as the first-in-first-out (FIFO) principle.

2.5.1 Dynamic User Equilibrium

When extending user equilibrium in a dynamic environment, route travel time calculation should be defined. Usually two types of route travel time are formulated: reactive (or instantaneous) and actual (or predictive); thus, contributing to two different dynamic user equilibrium conditions. The reactive route travel time is assessed as the sum of link travel times along the route at the time when travellers enter the route, whereas the actual route travel time is the sum of link travel times along the route estimated at the time when travellers enter each link. Thus, dynamic user equilibrium conditions can be outlined as follows:

For each OD pair at each instant of time, the actual (including reactive) travel times experienced by travellers departing at the same time are equal and minimal.

The definition extends Wardrop's first principle to dynamic situation. Under the assumption that travellers are aware of perfect dynamic traffic conditions, travellers are willing to use the routes with minimal actual (including reactive) times at each instant of time. Ran and Boyce (1996) named the DTA satisfying the dynamic user equilibrium conditions based on actual travel time as "ideal" DTA model and "instantaneous" DTA model based on reactive travel time.

2.5.2 FIFO Constraints

FIFO principle represents the proper traffic flow propagation. It means a flow that enters a link later cannot leave earlier by overtaking a flow which enters earlier. Mathematically, it can be expressed as:

$$t' > t'' \Rightarrow t' + \tau_a(t') > t'' + \tau_a(t'') \quad (2.21)$$

Where t' and t'' are the clock time. $\tau_a(t')$ is the travel time when the flow enters link a at time t' ; $\tau_a(t'')$ is similarly defined.

2.5.3 Mathematical Programming Approach

The earliest work on dynamic traffic assignment applying mathematical programming may be attributed to Merchant and Nemhauser (1978a). The authors considered a network with multiple origins but only to one destination. They formulated the DTA problem as a non-convex system-optimal problem. An existing flow function was adopted to represent congestion. The Kuhn-Tucker optimality conditions were shown to be consistent with Wardrop's second principle. In a follow-up paper by Merchant and Nemhauser (1978b), the authors discussed various ways of solving this mathematical problem noting that a globally optimal solution is rather difficult to obtain. Carey (1986) resolved an "open" question as to whether the Merchant and Nemhauser's model would satisfy a constraint qualification. It was shown that the models as developed by Carey would satisfy the linear independence constraint qualification which established the validity of the optimality analysis as presented in Merchant and Nemhauser's model. Later, Carey (1987) extended the Merchant and Nemhauser's model to a convex programming problem by introducing flow control which is defined as the difference between the actual exit flow and capacity flow. It was also suggested

that the piecewise linear method can be used to solve this convex problem. However, this model was still restricted to only one destination.

To deal with a network with multiple origins and destinations, Janson (1991) presented a mathematical programming formulation of the dynamic user-equilibrium problem. The O-D matrix is assumed to be fixed and congestion was not considered. This model can be seen as the basic model to represent the whole network. A heuristic dynamic traffic assignment procedure was proposed for applying to large networks to generate an approximate solution. But it was not a convergent solution algorithm.

After this method to cope with a general network was established, researchers focused on particular conditions to represent a more realistic world. Carey (1992) proposed the FIFO principle in a dynamic traffic assignment. The author formulated FIFO using discrete or integer conditions and continuous flow variables. The conclusion of these two ways was the result of non-convexity. Wu et al. (1998) considered the dynamic traffic assignment as a system of functional equations satisfying the FIFO principle. However, the FIFO conditions derived was only valid for a particular path, not ensured as well for the associated O-D pair. One restriction of this model is that the authors assumed that the path flow rate should be known in advance. In the real world, it seems impossible to get such information. Xu et al. (1999) extended their model as proposed earlier by Wu et al. (1998) to include some boundedness conditions of travel time function to ensure that the FIFO condition is applicable for all O-D pairs.

Beside the multiple O-D network and FIFO conditions, Kuwahara and Akamatsu (2001) considered DTA in more complicated networks with physical queuing. Kinematic wave theory (Newell, 1993a and 1993b) was applied to mimic the physical queues. However, the travellers were assumed to make route choice based on reactive travel time.

2.5.4 Optimal Control Theory Approach

Optimal control theory is an instrument to deal with a system including a dynamic process. Dynamic process is generally expressed by differential equations. Many natural phenomena follow this way. Friesz et al. (1989) were the first to apply the optimal control theory to dynamic traffic assignment problems. The authors assumed that the link exiting flow function is a function of traffic volume to represent traffic congestion. Link entering flows were taken as a control variable to determine the traffic volumes which were regarded as the state variables. The minimum principle as proposed by Pontryagin et al. (1962) was applied to derive the necessary conditions. However, their work was restricted to a single O-D pair. FIFO was satisfied by assuming the exiting flow function to be linear.

Wie et al. (1990) extended Friesz's model to a multiple O-D congestion network by applying a similar method. The link exit function was assumed to be linear and separable for different destinations. Their model was shown to be a proper dynamic extension of Beckmann's mathematical programming problem for a static user equilibrium traffic assignment.

Different from Wie and Friesz who only considered one control variable, Ran et al. (1993) formulated another class of instantaneous dynamic traffic assignment models in which the link exit flow was treated as another control variable. In this way, the dynamic user-optimal conditions were derived as defined in Section 2.5.1.

2.5.5 Variational Inequality Theory Approach

The static variational inequality theory has been introduced in Section 2.4.3. A general extension to the time-dependent situations can be shown as: denote a vector of control variables $\mathbf{u}(t) = [u_1(t), u_2(t), \dots, u_m(t)]$ and their dynamic process

$$\dot{\mathbf{x}}(t) = \mathbf{h}[\mathbf{x}(t), \mathbf{u}(t)]$$

where $\mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_n(t)]$ are defined as state variables and $\mathbf{h}(t) = [h_1(t), h_2(t), \dots, h_n(t)]$ defined as the state equations. Associated with the dynamic process, there is a vector of cost function $\mathbf{F}(t) = [F_1(t), F_2(t), \dots, F_m(t)]$. Each element of the cost function vector is a function of state and control variables, that is:

$$F_i(t) = F_i[\mathbf{x}(t), \mathbf{u}(t)] \quad \forall i = 1, 2, \dots, m$$

Since the state variables $\mathbf{x}(t)$ can be determined by the state equations when the control variables $\mathbf{u}(t)$ are given, the vector of cost functions can be simplified as $\mathbf{F} = \mathbf{F}[\mathbf{u}(t)]$. Define $G(t)$ as a given closed convex set of the control variables $\mathbf{u}(t)$, $\mathbf{F}(t)$ is assumed to be a set of continuous functions from $G(t)$ to $R^n(t)$. Then the definition of dynamic variational inequality problem is given as follows:

Determine a control vector $\mathbf{u}^*(t) \in G(t) \subset R^n(t)$, such that

$$\int_0^T \mathbf{F}[\mathbf{u}^*(t)][\mathbf{u}(t) - \mathbf{u}^*(t)] dt \geq 0 \quad \forall \mathbf{u}(t) \in G(t)$$

Friesz et al. (1993) were the first to propose the framework for formulating the dynamic user equilibrium problem as a variational inequality problem with simultaneous route choice and departure time decisions. Congestion was expressed by a delay function to ensure FIFO condition. Wie et al. (1995) formulated the dynamic user equilibrium problem as a variational inequality problem in discrete time. Ran and Boyce (1996) considered the case of a multiple O-D pair network. The dynamic user equilibrium conditions were established as in Section 2.5.1. A more complicated congestion network problem is to include queueing, multi-class motorists and dynamic stochastic assignment. Li et al. (2000) considered a reactive dynamic user equilibrium (DUE) model with queues. By constructing virtual links

for O-D pairs, the reactive DUE problem was formulated as a variational inequality problem over a polyhedral set. The problem is exactly the same as the fixed-demand static user equilibrium problem where the path enumeration is avoided. Huang and Lam (2002) modeled the point queueing problem as a discrete-time simultaneous VI problem. Departure time choice was also included. Bliemer and Bovy (2003) considered the multi-class dynamic traffic assignment problem for multiple O-D pairs and showed that even if link travel function was asymmetric, variational inequality theory was still efficient. However, it was a quasi-variational inequality model which implied that link travel time functions were identical for all classes of motorists. Han (2003) extended the DYNASTOCH (Ran and Boyce, 1996) algorithm to deal with the stochastic equilibrium problem on a general network.

Heretofore, the dynamics of traffic flow is involved with inflow rate, departure rate and so on. Based on these, the dynamic progress of flow and traffic assignment can be established. However, there is another way to describe the dynamics of traffic flow. That is the Cell Transmission Model (CTM) proposed by Daganzo (1994, 1995a and 1995b). Lo and Szeto (2002) formulated the user-optimal assignment problem as a variational inequality problem based on CTM. This method ensured that FIFO conditions are automatically satisfied and the cost function can be expressed by traffic flow explicitly, resulting in the direct use of the variational inequality algorithm.

2.5.6 Combined Approach

There exists a method to deal with dynamic traffic assignment by combining different approaches. Traffic simulator is usually employed to perform the traffic assignment work.

Srinivas (1994) incorporated the traffic simulator to study the system optimal dynamic traffic assignment. Tong and Wong (2000) employed a traffic simulator

to consider congested capacity-constrained road network. The traffic simulator incrementally loaded the traffic demand onto the network and updated the traffic conditions dynamically. However, the authors were not sure whether such results can reach convergence.

So far, many dynamic traffic assignment problems formulated by different analytical approaches have been discussed from a simple network representation to complicated modeling with queue, FIFO, stochastic and multi-class motorists. All of them are theoretical models and have not been calibrated and validated using field data. This is partially because some of the input parameters are difficult to get presently. For instance, time-variant demand, travel time functions under queuing phenomena and travellers' travel behaviour under dynamic situations. From modeling perspective, mathematical programming has its limitation in that it prevents a suitable description of traffic interactions and dynamics. The travel time function must be symmetric. In the model formulated by the optimal control theory, the exit flow function is restricted with a concave function and when the initial exit flow is zero, it causes unrealistic flow propagation. When strict link capacities are considered, there is no equivalent optimal control-based model. However, variational inequality approach can overcome the asymmetry of cost functions. It seems to be the most promising approach for formulating and solving the DTA model for road pricing.

2.6 SUMMARY

In this Chapter, we have discussed road pricing modeling from early simple economical principle approach to the latest complicated approach incorporating traffic assignment on the whole network. From the introduction of road pricing in practice, it is evident that different design objectives would require different road pricing schemes. To achieve short time traffic management (Ben-Akiva, et al., 2001) by adjusting travelers' route choice behavior, an entire network

consideration is obviously helpful. However, there is not much work dealing with dynamic road pricing incorporating dynamic traffic assignment. In reality, short term traffic management which intends to redistribute traffic within a short time period is usually required. Thus, developing a dynamic road pricing model and algorithm becomes a challenge in this research. In the next Chapter, static road pricing modeling and algorithm will be discussed in detail.

CHAPTER 3 OPTIMAL TOLL FORMULATION AND ALGORITHMS

3.1 INTRODUCTION

Static and dynamic traffic assignment models have been reviewed in the previous Chapter. In this Chapter, the optimal toll problem will be formulated as a MPEC (Mathematical Programs with Equilibrium Constraints) problem. Algorithms to solve MPEC problem such as PIPA (Penalty Interior Point Algorithm), implicit programming-based algorithm, smoothing method and nonlinear optimization programming-based algorithm will be reviewed. Two heuristics methods which are based on DE (Differential Evolutionary) global algorithm and PS (Pattern Search) will be proposed. A simple network will be tested to illustrate these algorithms.

3.2 MPEC FORMULATION OF OPTIMAL TOLL

Generally, travelers tend to choose departure time and paths based on user-equilibrium, while traffic managers are likely to achieve system efficiency by implementing pricing instrument. The objectives of both the travelers and the traffic managers are thus non-cooperative. Traffic managers usually have the ability to enforce their strategy onto the travelers. In fact, traffic managers always make known the tolling level to the travelers first and then the travelers do their choices with the knowledge of such information. In addition, traffic managers are aware of the possible choices made by the travelers when under tolling condition. This kind of decision-making problems is often referred to

as the Stackelberg (1952) game in which the one that holds the powerful position is called the leader, while the others the follower. In the optimal toll problem, traffic manager can be seen as the leader and all the travelers as the follower. If the decision problem of the follower can be expressed as a Mixed Complementarity Problem (MCP), then, such a Stackelberg game can be rewritten as a MPEC problem. The original MPEC model is given as follows:

$$\min_{\mathbf{x} \in R^n, \mathbf{y} \in R^m} f(\mathbf{x}, \mathbf{y}) \quad (3.1)$$

subject to:

$$g(\mathbf{x}, \mathbf{y}) \leq 0 \quad (3.2)$$

$$\mathbf{y} \text{ solves } MCP(\mathbf{h}(x, y), \mathbf{B}) \quad (3.3)$$

where \mathbf{x} is the decision variable of the leader, \mathbf{y} is the decision variable of the follower. f is the objective function of the leader; while g is the constraint involving decision variables from both sides. The follower would satisfy a MCP condition where \mathbf{B} is a box constraint on variable \mathbf{y} . For example, a point \mathbf{y} with $a_i \leq y_i \leq b_i$ and solving Equation (3.3) means that for each i , at least one of the following equations holds:

$$h_i(\mathbf{x}, \mathbf{y}) = 0 \quad (3.4)$$

$$h_i(\mathbf{x}, \mathbf{y}) > 0, y_i = a_i \quad (3.5)$$

$$h_i(\mathbf{x}, \mathbf{y}) < 0, y_i = b_i \quad (3.6)$$

A special case is that when $a_i = 0, b_i = +\infty$ for all i , Equation (3.4), (3.5) and (3.6) can be combined and re-written into a simple equation given as follows:

$$0 \leq h_i(\mathbf{x}, \mathbf{y}), y_i \geq 0, h_i(\mathbf{x}, \mathbf{y}) \cdot y_i = 0 \quad (3.7)$$

which is a Nonlinear Complementarity Problem (NCP) discussed in Chapter 2.

3.3 REVIEW OF ALGORITHMS FOR MPEC PROBLEMS

MPEC problems are extremely hard to solve. A special case of MPEC is a bi-level linear program and it belongs to the class of strongly NP-hard problems (Hansen et al., 1992). It has also been shown that MPEC problems violated the Mangasarian Fromovitz Constraint Qualification (MFCQ) (Chen and Florian, 1995) which resulted in difficulties in using the Non-Linear Programming (NLP) techniques to solve MPEC. By introducing slack variables (Dirkse et al., 2002), Equations (3.4), (3.5) and (3.6) can be reformulated in the form of Equation (3.7). When replacing Equation (3.3) in the original MPEC with Equation (3.7), MPEC can be written as follows (Fletcher and Leyffer, 2002a):

$$\min f(\mathbf{z}) \quad (3.8)$$

subject to

$$c_e(\mathbf{z}) = 0 \quad (3.9)$$

$$c_l(\mathbf{z}) \leq 0 \quad (3.10)$$

$$0 \leq \mathbf{z}_1 \perp \mathbf{z}_2 \geq 0 \quad (3.11)$$

where $\mathbf{z} = (\mathbf{z}_0, \mathbf{z}_1, \mathbf{z}_2)$, \mathbf{z}_0 is the control variable, $(\mathbf{z}_1, \mathbf{z}_2)$ are the state variables. $c_e(\mathbf{z}) = 0$ is the equality constraints while $c_l(\mathbf{z}) \leq 0$ is the inequality constraints.

3.3.1 NLP Techniques Based Method

By replacing $\mathbf{z}_1 \cdot \mathbf{z}_2 = 0$ in the complementarily Equation (3.11) with $\mathbf{z}_1 \cdot \mathbf{z}_2 \leq 0$, the MPEC is equivalent to a nonlinear programming (NLP) formulation. NLP packages have been widely used to search for a solution. Bard (1988) employed the gradient projection method on some bi-level problems and reported 50-70% failure on the test. Conn et al. (1996) and Ferris and Pang (1997) tested the NLP solver “Lancelot” and reported the same failure percentage. Fletcher and Leyffer (2002a) applied the filter SQP (Sequential Quadratic Programming) (Fletcher and Leyffer, 2002b) on a large collection of test problems and showed that SQP methods were very well suited to solving MPECs.

3.3.2 Smoothing Method

Smoothing method is to replace the complementarily equation by a well-behaved function. A simple smoothing is to replace $\mathbf{z}_1^T \cdot \mathbf{z}_2 \leq 0$ by $\mathbf{z}_1^T \cdot \mathbf{z}_2 \leq \tau_k$, and solve a sequence of NLPs for decreasing τ_k to zero as k increases (Ferris and Kanzow, 1999. Scheel and Scholtes, 2000). Facchinei et al. (1996) introduced a smoothed function:

$\psi_{\mu}(z_{1i}, z_{2i}) = \sqrt{(z_{1i} - z_{2i})^2 + 4\mu} - z_{1i} - z_{2i} = 0$ to replace the complementarily equation $0 \leq \mathbf{z}_1 \perp \mathbf{z}_2 \geq 0$. μ is a parameter. It can easily be seen that $\psi_{\mu \rightarrow 0}(z_{1i}, z_{2i}) = 0 \Leftrightarrow 0 \leq \mathbf{z}_1 \perp \mathbf{z}_2 \geq 0$ for all i . The resulting NLP satisfies MFCQ and differentiable with non-zero μ . Thus, algorithms can be designed to solve the sequence of NLPs for decreasing μ to zero. Similarly, Jiang and Ralph (1997) proposed the smoothed Fischer-Burmeister function: $\psi_{\mu}(z_{1i}, z_{2i}) = \sqrt{z_{1i}^2 + z_{2i}^2 + \mu} - z_{1i} - z_{2i} = 0$ to replace the complementarily equation. SQP methods can be applied to solve the resulting NLP where μ is reduced at each iteration.

3.3.3 Penalty and Interior Point Based Method

After using $\mathbf{z}_1^T \cdot \mathbf{z}_2 \leq 0$ to replace the complementarily constraints, another traditional way to deal with the resulting NLP is to penalize $\mathbf{z}_1^T \cdot \mathbf{z}_2 \leq 0$ into a minimising objective function $f(\mathbf{z}) + v_k \mathbf{z}_1^T \mathbf{z}_2$. v_k is the penalty parameters. Ferris and Tin Loi (1998) conducted some numerical experiments using this approach. The Penalty Interior Point Algorithm (PIPA) proposed by Luo et al. (1996) followed the same penalty mechanism and used SQP approach to solve the resulting NLP. However, Leyffer (2002) later found that PIPA can fail under certain conditions.

3.3.4 Implicit Programming Method

If under some conditions, for each fixed \mathbf{x} in the Equation (3.2), \mathbf{y} has a unique solution for the MCP (3.3), this would lead to the formulation of the MPEC as an “implicit program”. The objective of (3.1) can be written as $f(\mathbf{x}, \mathbf{y}(\mathbf{x}))$. Pang et al. (1991) proposed a descent direction search method for solving the resulting implicit program. This method was also discussed by Luo et al. (1996) and Lim (2002).

3.4 HEURISTICS ALGORITHMS

The traditional algorithms mentioned above are often seen as derivative-based methods because they always require the differential characteristics of the f with respect to the variable \mathbf{x} or \mathbf{y} . In fact, when problems are so complicated that they would usually not meet this requirement and heuristics algorithms become the potential candidates. A number of heuristics algorithms have been proposed to tackle the optimization problems. In this research study, two of them can be used to solve the MPEC problem. These two methods will also be applied to solve the dynamic road pricing problem illustrated in the next Chapter.

3.4.1 Evolutionary Algorithm Based Method (DE-Toll)

Evolutionary Algorithm, alternatively known as Generic Algorithm (GA) is usually used to deal with not well-behaved optimization problem. The objective function is not differentiable for instance. A special evolutionary algorithm is the Differential Evolution (DE) designed by Storn and Price (1995). It is a simple population-based, direct-search algorithm. *[DE won the 3rd prize*

at the First International Contest on Evolutionary Computation (1st ICEO) which was held in Nagoya, May 1996.] DE differs from the conventional GA in a number of points given below:

1. DE uses real number representation while conventional GA uses binary strings.
2. In DE, three parents are selected for crossover and the child is a perturbation of one of them. In GA, only two parents are chosen for crossover and the child is a recombination of them.
3. In conventional GA, children replace the parents with some probability, while in DE, the replacement happens only when child is of higher fitness.

The whole algorithm is illustrated in Figure 3.1. When DE is used to solve the MPEC, it assumes \mathbf{x} in Equation (3.1) as chromosomes and computes the fitness by solving the parametric MCP (3.3).

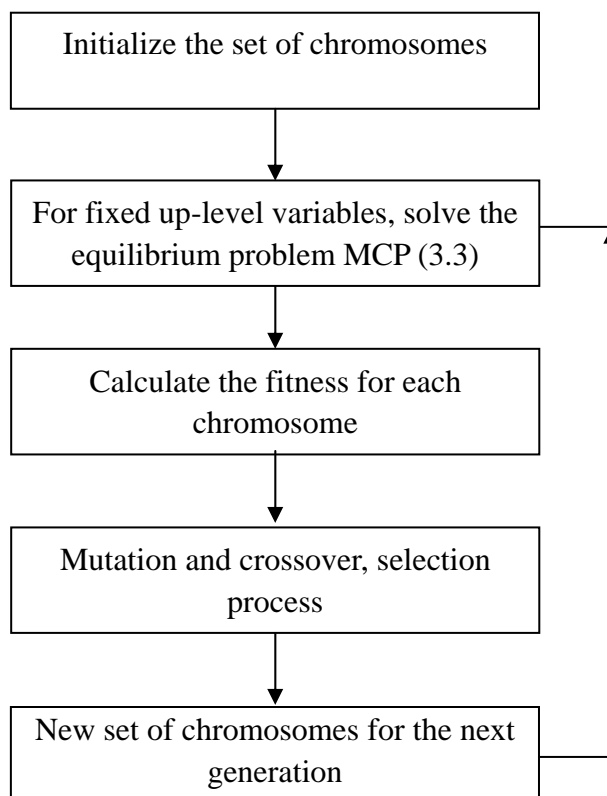


Figure 3.1 Differential Evolutionary Algorithm for MPEC

Initial Population

Suppose the population size is N_2 and will be maintained over the whole process. The initial population is randomly generated:

$$\mathbf{x}_{i,G=0} \quad i = 1, 2, \dots, N_2 \quad (3.12)$$

Mutation

For each individual vector $\mathbf{x}_{i,G}$ in the generation G , a mutant vector is generated as follows:

$$\mathbf{v}_{i,G+1} = \mathbf{x}_{r1,G} + F \cdot (\mathbf{x}_{r2,G} - \mathbf{x}_{r3,G}) \quad (3.13)$$

Integers $r1, r2, r3$ are randomly selected from the set $\{1, 2, \dots, N_2\}$. F is a real number parameter belonging to $[0, 2]$. It controls the amplification of the difference vector $\mathbf{x}_{r2,G} - \mathbf{x}_{r3,G}$. Figure 3.2 illustrates the mutation process in a two dimension setting.

Crossover

To keep the diversity of the children vectors in the next generation, a trial vector $\bar{\mathbf{x}}_{i,G+1} = (\bar{x}_{1i,G+1}, \bar{x}_{2i,G+1}, \dots, \bar{x}_{Di,G+1})$ is designed to represent the crossover, D which is the dimension of each vector $\mathbf{x}_{i,G}$. For each i from $[1, 2, \dots, N_2]$:

$$\bar{x}_{ji,G+1} = \begin{cases} v_{ji,G+1} & \text{if } (randb(j) \leq CR) \text{ or } j = rnbr(i) \\ x_{ji,G} & \text{otherwise} \end{cases} \quad j = 1, 2, \dots, D \quad (3.14)$$

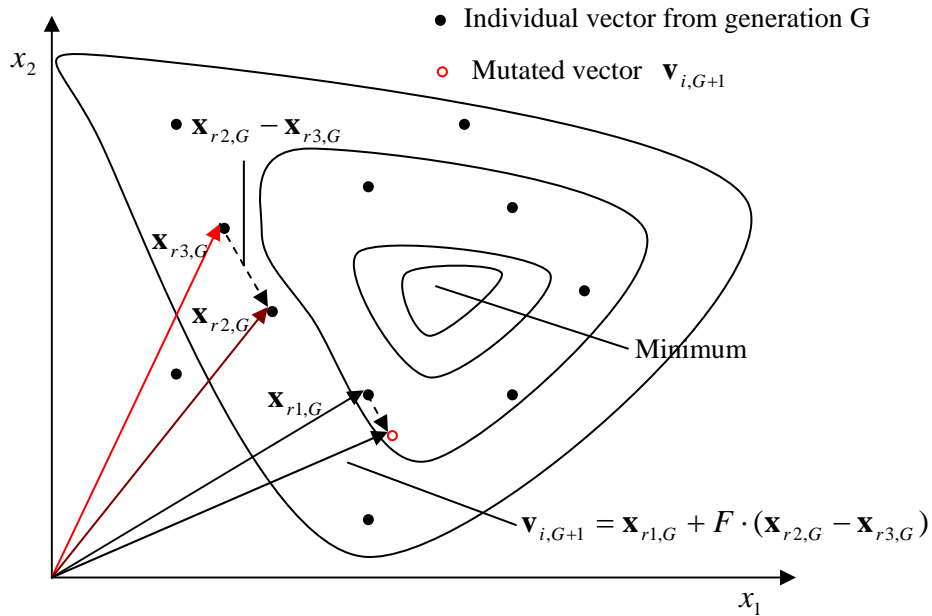


Figure 3.2 An Example of Mutation in Two Dimensional Setting

where $randb(j)$ is the j^{th} evaluation of a uniform random number generator between $[0,1]$. $rnbr(i)$ is an index randomly chosen from $[1,2,\dots,D]$. CR is the pre-determined crossover constant which belongs to $[0,1]$. An example of crossover is shown in Figure 3.3.

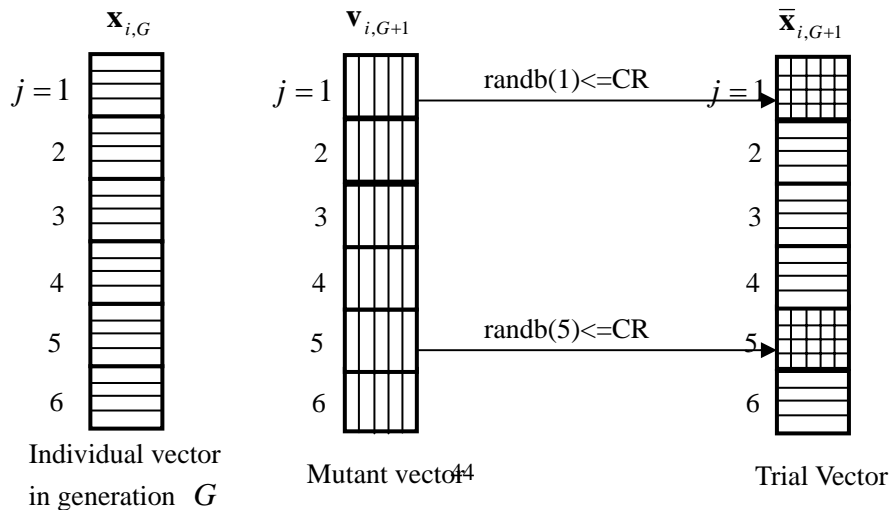


Figure 3.3 An Example of the Crossover Process for $D = 6$ Parameters

From the crossover process, it should be noted that at least one of the mutant vectors $\mathbf{v}_{i,G+1}$ should enter the trial vector $\bar{\mathbf{x}}_{i,G+1}$.

Selection

After getting the trial vector, one has to decide which one should be selected to enter the next generation. Fitness function is used as a criterion to make a decision. Usually, the objective function of (3.1) is used as a fitness function. Comparing the trial vector $\bar{\mathbf{x}}_{i,G+1}$ with the target vector $\bar{\mathbf{x}}_{i,G}$, one must make sure that those with smaller objective value will enter the next generation. It can be written as follows:

$$\mathbf{x}_{i,G+1} = \begin{cases} \bar{\mathbf{x}}_{i,G+1} & \text{if } f(\bar{\mathbf{x}}_{i,G+1}, \mathbf{y}(\bar{\mathbf{x}}_{i,G+1})) < f(\mathbf{x}_{i,G}, \mathbf{y}(\mathbf{x}_{i,G})) \\ \mathbf{x}_{i,G} & \text{otherwise} \end{cases} \quad (3.15)$$

Box Constraint

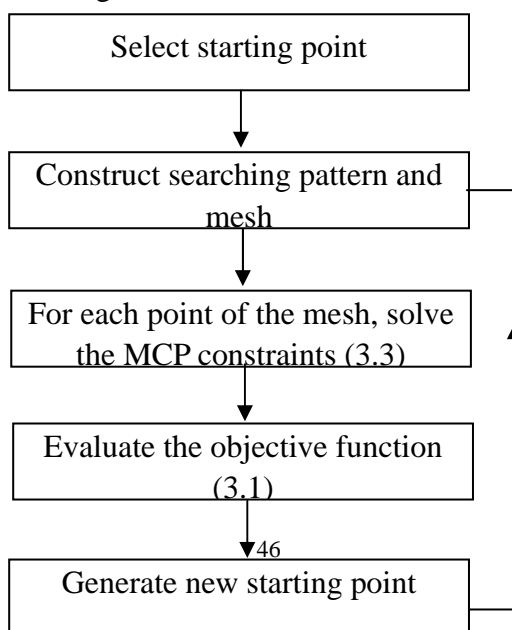
All practical problems generally have limits on their parameters. A simple variable constraint is a box constraint which means that every variable has its own upper bound and lower bound. It can be expressed as: $\mathbf{lb} \leq \mathbf{x} \leq \mathbf{ub}$, where \mathbf{lb} is the lower bound vector and \mathbf{ub} is the upper bound vector. To generate the in-bound offspring, one method is to adjust the offspring between the pre-mutation/crossover $\mathbf{x}_{i,G}$ and the box bounds as mentioned in Corne et al. (1999).

$$\bar{x}_{j,i,G+1} = \begin{cases} (x_{j,i,G} + lb_j) / 2 & \text{if } \bar{x}_{j,i,G+1} < lb_j \\ (x_{j,i,G} + ub_j) / 2 & \text{if } \bar{x}_{j,i,G+1} > ub_j \\ x_{j,i,G+1} & \text{otherwise} \end{cases} \quad (3.16)$$

Linear constraints are not discussed in that only box constraints are incurred in this research.

3.4.2 Pattern Search Algorithm (PS-Toll)

The pattern search algorithm is originated from Hooke and Jeeves (1961). It is an iterative searching process that repeatedly searching a set of points (called mesh) around the current point. If a point is found closer to the optimal point than the current point, the current point will be replaced by this new point and the next round of searching starts again. The different ways to construct a new mesh around the new current points are called the patterns. The MathWorks has incorporated two pattern search algorithms in the Genetic Algorithm and Direct Search (GADS) Toolbox. One is the Generalized Pattern Search (GPS) algorithm which originates from Booker et al. (1999). The other is the Mesh-Adaptive Direct Search (MADS) algorithm (Audet and Dennis, 2004) for solving the optimization problems with general nonlinear constraints. The procedure is depicted in Figure 3.4:



The following example in a two dimension space shows the default pattern search algorithm in GADS Toolbox. The problem is set to find some \mathbf{x} to minimize f .

First, the starting point is set as \mathbf{x}_0 , the searching pattern consists of four directions as $\mathbf{d}_1 = (1, 0)$, $\mathbf{d}_2 = (-1, 0)$, $\mathbf{d}_3 = (0, 1)$ and $\mathbf{d}_4 = (0, -1)$. Then, a mesh M_1 can be constructed by adding the four directions to the starting point \mathbf{x}_0 . $M_1 = \{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4\}$, where \mathbf{x}_i ($i = 1, 2, 3, 4$) is determined as:

$$\mathbf{x}_i = \mathbf{x}_0 + \lambda \cdot \mathbf{d}_i \quad i = 1, 2, 3, 4 \quad (3.17)$$

in which λ is the expand factor. For the first iteration, $\lambda = 1$. This process is shown in Figure 3.5.

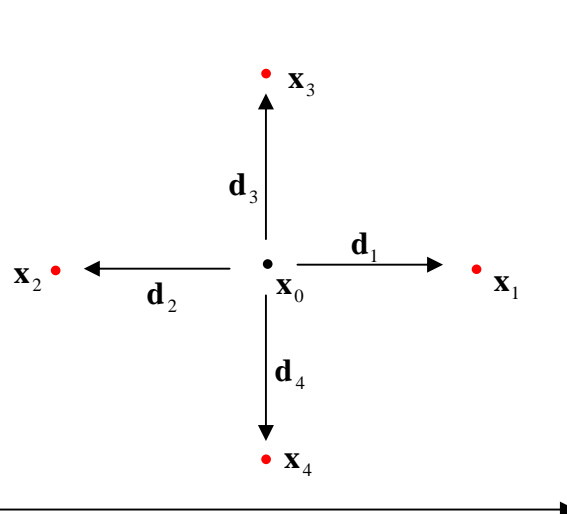


Figure 3.5 Pattern and Mesh Illustration

Second, the objective function is evaluated at each of the point in the mesh. These four objective function values $f(\mathbf{x}_i)$ are compared with the objective function value at \mathbf{x}_0 , i.e., $f(\mathbf{x}_0)$. If one of the \mathbf{x}_i is found such that its objective function value is less than $f(\mathbf{x}_0)$, for example $f(\mathbf{x}_2) < f(\mathbf{x}_0)$, then, the point (\mathbf{x}_2 in this example) will be set as the new starting point in the next iteration. At the same time, the expand factor in Equation (3.17) is doubled for the next iteration. If none of the points in the mesh satisfy $f(\mathbf{x}_i) < f(\mathbf{x}_0)$, then, \mathbf{x}_0 is taken as the starting point in the next iteration. However, the expand factor will be adjusted as half of its value in the previous iteration.

3.5 NUMERICAL EXAMPLE

In this section, a nine-node network (Hearn and Yildirim, 2002) as shown in Figure 3.6 will be used to demonstrate the Evolutionary Algorithm for the MPEC and the Pattern Search Algorithm for MPEC. At the same time, this MPEC will be solved by a solver named NLPEC (Non-linear Problem with Equilibrium Constraints) from the General Algebraic Modeling System (GAMS) which uses the NLP method as discussed above. The Pattern Search Algorithm is applied by the Generic Algorithm and the Direct Search Toolbox.

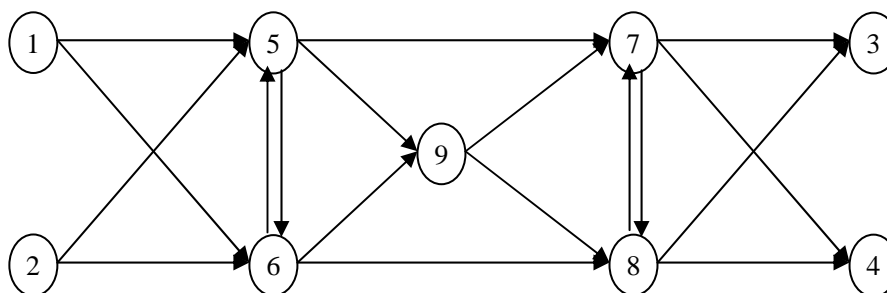


Figure 3.6 The Nine Node Network

Link $8 \rightarrow 4$ is subjected to toll to minimize the total travel delay $\sum f_a t_a$, where f_a is the link flow on link a and t_a is the link travel time. The user-equilibrium traffic assignment problem with elastic demand was considered in Equations (2.20) and (2.21) in Chapter 2 except the link cost that should include the tolling effect. The problem formulation is given as follows:

$$\min \sum_{a \in A} f_a t_a \quad (3.18)$$

Subject to

$$A\mathbf{f}_k = \mathbf{d}_k(\mathbf{t}) \quad \forall k \in C \quad (3.19)$$

$$0 \leq c_{ij}(\mathbf{f}) + t_{jk} - t_{ik} \perp f_{ijk} \geq 0 \quad \forall (i, j) \in A, k \in C \quad (3.20)$$

$$c_{ij}(\mathbf{f}) = \begin{cases} t_{ij}(\mathbf{f}) + tol & \text{if } (i, j) = (8, 4) \\ t_{ij}(\mathbf{f}) & \text{otherwise} \end{cases} \quad (3.21)$$

Note that in Equation (3.21), the toll level is directly added to link cost indicating that parameter “value of time” is set to \$ 1 /hour.

The matrix A is a node-link incidence matrix as explained in Equation 2.19 in Chapter 2. The network has 4 OD pairs and the demand function of each OD pair with respect to its travel time is given in Table 3.1.

Table 3.1 Elastic Demand Function of Each OD Pair

OD pair	Demand Function
1→3	$d_{(1,3)} = 10 - 0.5t_{(1,3)}$
1→4	$d_{(1,4)} = 20 - 0.5t_{(1,4)}$
2→3	$d_{(2,3)} = 30 - 0.5t_{(2,3)}$
2→4	$d_{(2,4)} = 40 - 0.5t_{(2,4)}$

The link travel time used is $t_a(f_a) = A_a(1 + 0.15(f_a / B_a)^4)$ where parameters

A_a and B_a are tabulated as shown in Table 3.2.

Table 3.2 Parameters of Link Travel Time Function

Link a	A_a	B_a
1→5	5	12
1→6	6	18
2→5	3	35
2→6	9	35
5→6	4	11
6→5	9	20
5→7	2	11
6→8	6	33
5→9	8	26
6→9	7	32
9→7	4	26
9→8	8	30
7→8	4	36
8→7	2	19
7→3	3	25
7→4	6	24
8→3	8	39

8→4	6	43
-----	---	----

Results

For the DE-Toll algorithm, the parameters used in this example is shown in Table 3.3:

Table 3.3 Parameters for DE-Toll

<i>N</i>	<i>F</i>	<i>CR</i>
10	0.8	0.5

The results computed from these 3 different algorithms were quite similar and are tabulated in Table 3.4.

Table 3.4 Comparison of NLPEC, DE-Toll and PS-Toll

	NLPEC	DE-Toll	PS-Toll
The optimal toll level (\$)	1.08	1.08	1.08
The total travel delay (veh. hour)	1236.74	1236.74	1236.74
CPU time (second)	1	660	40

Figure 3.7 shows the total travel time convergence using the DE-Toll method. Figure 3.8 and Figure 3.9 show the convergence using the PS-Toll method with different starting points. The CPU time using PS-Toll method in Table 3.4 is the average value of those with different starting points. From Table 3.4, we found that the computation results were almost identical despite the difference in computation time. We note that in each iteration of the DE-Toll and PS-Toll methods, there is a need to formulate a static traffic assignment problem in order to solve the assignment under a given toll level. These static traffic

assignment problems can be formulated as a MCP as discussed in the last Chapter. For this particular example, where the travel time is a smooth function of link flow, a strong PATH solver (Ferris and Munson, 2000) can be used to tackle the MCP. The solver PATH is integrated in GAMS. Both the DE-Toll and PS-Toll methods are under the Matlab environment. However, PATH is needed to be called from Matlab in each iteration. The running time is consequently longer than that for the NLPEC.

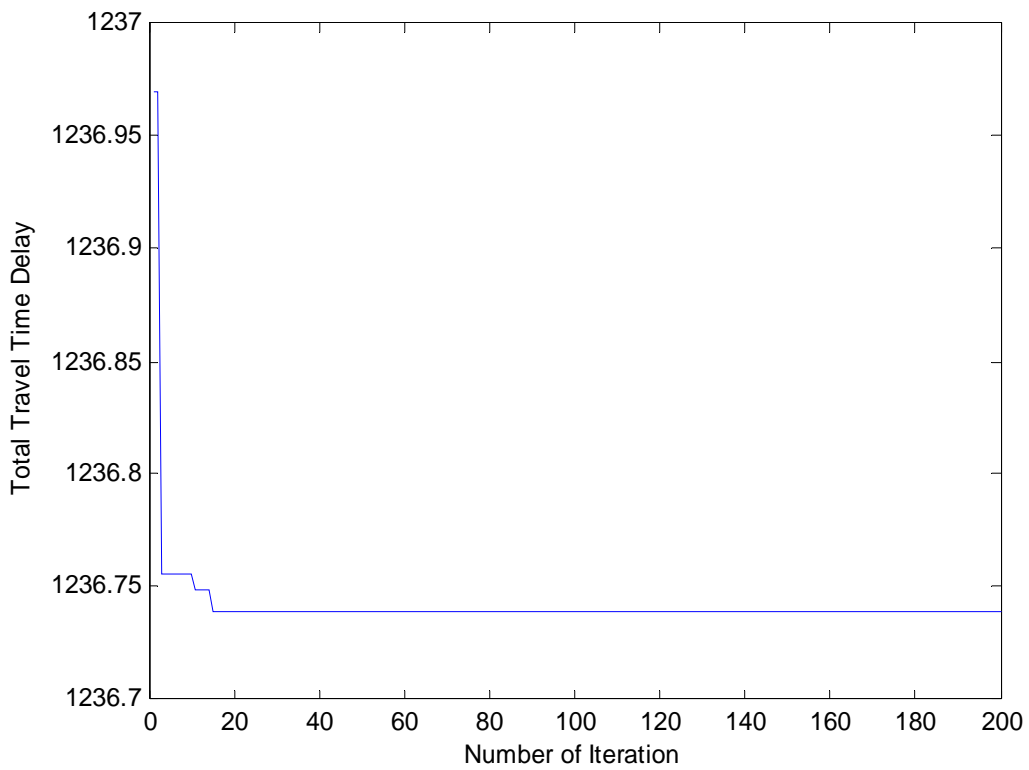


Figure 3.7 Total Travel Time Delay Using DE-Toll

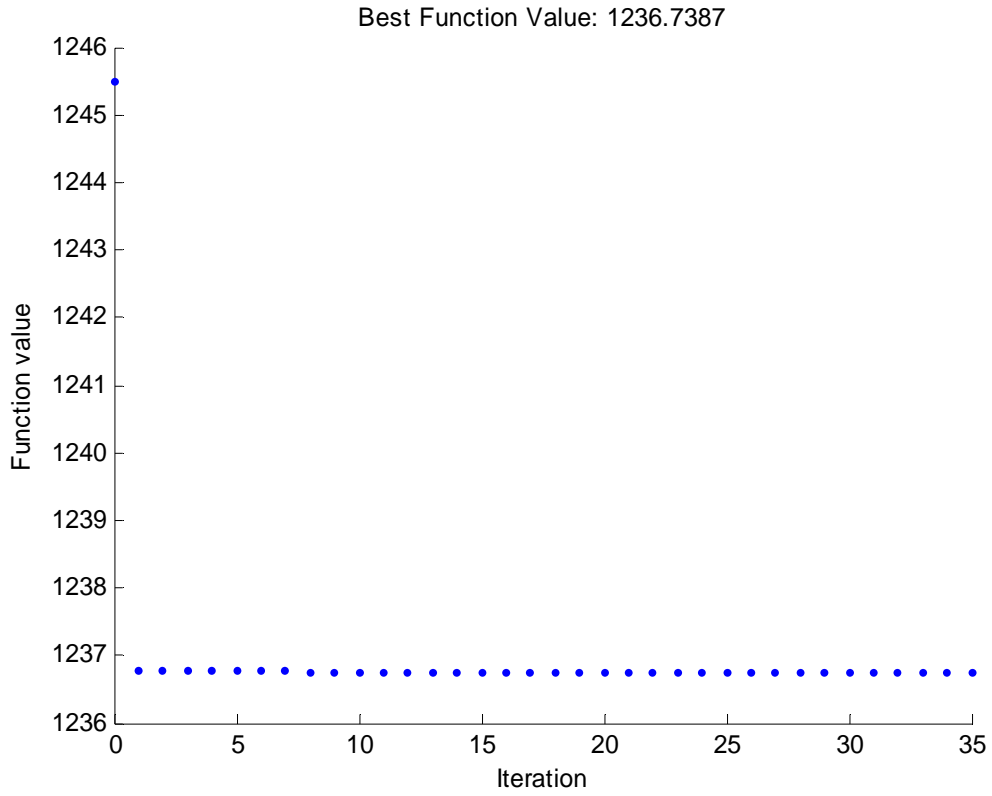


Figure 3.8 Total Travel Time Delay Using PS-Toll with Starting Point 0

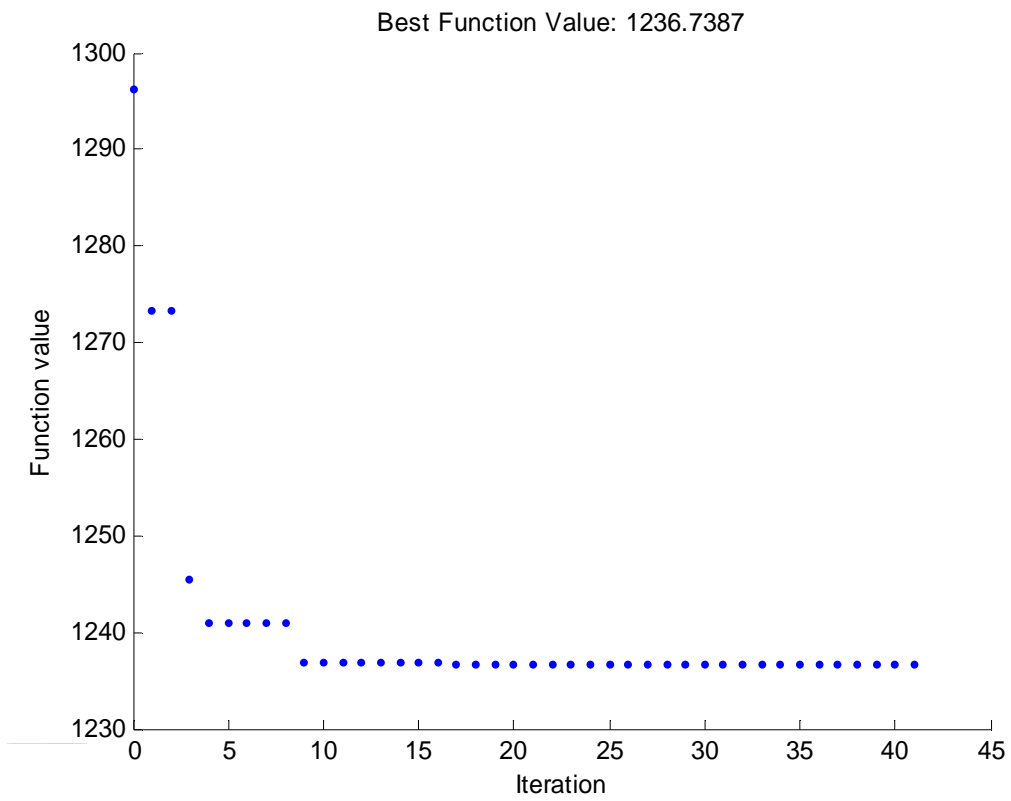


Figure 3.9 Total Travel Time Delay Using PS-Toll with Starting Point 5

3.6 SUMMARY

In this chapter, we formulated the optimal road pricing problem as a MPEC problem while considering traffic assignment over the entire network. Algorithms for MPEC were reviewed and heuristics algorithms such as DE-Toll and PS-Toll were proposed. A nine-node network was used to demonstrate the capability and efficiency of the algorithm. The static traffic assignment problem is formulated as a MCP and solved using the PATH solver. However, in dynamic situations, due to the complexity of dynamic travel time, a “well-behaved” MCP usually cannot be obtained; thus, resulting in PATH losing its problem solving capability and efficiency. In the next two chapters, dynamic optimal road pricing problem incorporating dynamic traffic assignment will be discussed.

CHAPTER 4 DYNAMIC TRAFFIC ASSIGNMENT MODEL AND ALGORITHM

4.1 INTRODUCTION

It was discussed in Chapter 2 that there is no general formulation of DTA. In most situations, the formulation of DTA is problem-specific. Since the study focuses on dynamic optimal toll on a network, a suitable network formulation is therefore needed. A network which can reflect the tolling effect on traffic distribution in the extent of space and time is preferred. In this chapter, the day-to-day DTA model from Huang and Lam (2002) was modified and applied to our proposed model, which satisfies the dynamic user equilibrium and determines the right departure time choice for travelers. It also satisfies FIFO as defined in Chapter 2. However, a problem in the formulation of the DTA model as proposed by Huang and Lam (2002) will be highlighted and a new procedure to compute some inter-variables will be discussed. At last, a different algorithm which is based on projection method will be proposed and the results obtained will be compared with those obtained from the “route/time swapping” algorithm proposed by Huang and Lam.

4.2 DTA MODEL DESCRIPTION

In this section, the detailed formulation of a DTA model on a network with queue will be introduced. The basic model is from Huang and Lam (2002) to represent dynamic traffic assignment using variational inequality theory. The idea is to represent queues, travel time and path costs as continuous functions of

inflow. Thus, solution existence conditions are satisfied which is discussed in Section 4.3.

4.2.1 Problem formulation

The network is represented as a graph $G(N_1, A)$, in which N_1 is the set of all nodes and A is the set of all links. Let P_{rs} denotes the set of all feasible paths connecting origin r and destination s . The total time period of interest is T which is discretized into a finite set of time intervals. The total number of time intervals is K . δ is the time interval length. Thus, $\delta \cdot K = T$. The time period $[0, T]$ is large enough so that all traffic can leave the network before the end of T . The total traffic demands of the respective origin-destination during the total time period $[0, T]$, *i.e.* F^{rs} , are assumed to be known. The model satisfies the dynamic user-equilibrium and FIFO requirements, and the user-equilibrium is based on the “actual travel time” or the predictive travel time.

The mathematical expression of the model is expressed as follows:

$$c_p^{rs}(k, \mathbf{f}^*) \begin{cases} = c_{\min}^{rs}(\mathbf{f}^*) & \text{if } f_p^{rs^*}(k) > 0 \\ \geq c_{\min}^{rs}(\mathbf{f}^*) & \text{if } f_p^{rs^*}(k) = 0 \end{cases} \quad \forall p \in P_{rs}, k \in K \quad (4.1)$$

$$\sum_{p \in P_{rs}} \sum_{k \in K} \delta \cdot f_p^{rs^*}(k) = F^{rs} \quad (4.2)$$

$$f_p^{rs^*}(k) \geq 0 \quad (4.3)$$

in which $f_p^{rs}(k)$ is the flow rate of path p that enters the network from origin r to destination s during time interval k . $c_p^{rs}(k, \mathbf{f})$ is the predictive path cost departing from r to s and selecting path p during time interval k . $c_{\min}^{rs}(\mathbf{f}) = \min\{c_p^{rs}(k, \mathbf{f}) : p \in P_{rs}, k \in K\}$. \mathbf{f} is the vector expression of $f_p^{rs}(k)$. The notation with asterisk represents the optimal value.

The formulation above can also be rewritten as a finite-dimensional VI problem given as follows:

$$\sum_k \sum_{rs} \sum_p c_p^{rs}(k, \mathbf{f}^*) \cdot [f_p^{rs}(k) - f_p^{rs*}(k)] \geq 0 \quad (4.4)$$

Subject to a closed convex set $\Omega \subset R^n$:

$$\Omega = \left\{ \mathbf{f} \geq 0 : \sum_p \sum_k \delta \cdot f_p^{rs}(k) = F^{rs} \right\} \quad (4.5)$$

From condition (4.1), it leads to the following equation:

$$\sum_k \sum_{rs} \sum_p f_p^{rs*}(k) [c_p^{rs}(k, \mathbf{f}^*) - c_{\min}^{rs}(\mathbf{f}^*)] = 0 \quad (4.6)$$

Thus, the VI problem is equivalent to finding \mathbf{f}^* such that $W(\mathbf{f}^*) = 0$ with the constraints (4.1)-(4.3) satisfied. Here, $W(\mathbf{f})$ is written as follows:

$$W(\mathbf{f}) = \sum_k \sum_{rs} \sum_p f_p^{rs}(k) [c_p^{rs}(k, \mathbf{f}) - c_{\min}^{rs}(\mathbf{f})] \quad (4.7)$$

The DTA model incorporates queues on the network and satisfies FIFO conditions. Section 4.2.2 and section 4.2.3 deduce the link travel time and

queue representation under FIFO conditions. Section 4.2.4 and section 4.2.5 describe how to write $c_p^{rs}(k, \mathbf{f})$ as a continuous function of \mathbf{f} .

4.2.2 Link Travel Time

The link travel time design is critical as it reflects different model formulation and computational complexity. Since the network is simulating queuing phenomenon, queuing is assumed to have an effect on link travel time. The point queuing concept (Kuwahara and Akamatsu, 1993, 1997) is employed. The relationship can be expressed as follows:

$$t_a(k) = t_a^0 + \frac{q_a(k)}{\delta \cdot s_a} \quad (4.8)$$

t_a^0 is the free-flow travel time which is a given constant value. $q_a(k)$ is the queue size experienced by vehicles entering link a during interval k . s_a is the maximum exit rate which is a constant value.

4.2.3 Representation of Queuing

Let $v_a(k)$ denotes the departure rate from link a during interval k . The flows entering link a in interval $k-1$ leave the link before the end of the interval $k-1+t_a(k-1)$ at the departure rate $v_a(k-1+t_a(k-1))$. One assumes the flows entering in the next interval k leave the link during $[k-1+t_a(k-1), k+t_a(k)]$ at the departure rate $v_a(k+t_a(k))$ and the value will

not vary during this time period. Thus:

$$V_a(k + t_a(k)) = V_a(k - 1 + t_a(k - 1)) + \delta[k + t_a(k) - (k - 1 + t_a(k - 1))]v_a(k + t_a(k)) \quad (4.9)$$

where $V_a(k)$ is the cumulative departure from link a during interval k .

Note that the FIFO conditions: $U_a(k) = V_a(k + t_a(k))$ in which $U_a(k)$ is the cumulative arrivals at link a during interval k and the flow equation is:

$$U_a(k) = U_{a-1}(k) + \delta u_a(k) \quad (4.10)$$

With (4.9) and (4.10), we have:

$$u_a(k) = v_a(k + t_a(k))[1 + t_a(k) - t_a(k - 1)] \quad (4.11)$$

The deterministic queuing theory requires the departure rate be evaluated as follows:

$$v_a(k + t_a(k)) = \begin{cases} s_a & \text{if } t_a(k) > t_a^0 \text{ or } u_a(k) > s_a \\ u_a(k) & \text{otherwise} \end{cases} \quad (4.12)$$

With (4.8), (4.11) and (4.12), we can derive the following expression of queue:

$$q_a(k) = \max \left\{ \left[q_a(k - 1) + \delta (u_a(k) - s_a) \right], 0 \right\} \quad (4.13)$$

4.2.4 Link Flow Conservation

Link inflow rates are given by the following relations:

$$u_a(k) = \sum_{rs} \sum_p u_a^{rsp}(k) \quad (4.14)$$

$$u_a^{rsp}(k) = f_p^{rs}(k) \xi_a^{rsp} + v_b^{rsp}(k) \xi_{ba}^{rsp} \quad (4.15)$$

$$v_b^{rsp}(k) = \begin{cases} u_b^{rsp}(k - t_b^0) & \text{if the queue is null at interval } k - t_b^0 \\ s_b \frac{u_b^{rsp}(i)}{u_b(i)} & \text{otherwise, where } i + t_b(i) = k \end{cases} \quad (4.16)$$

$u_a^{rsp}(k)$ is the flow rate that enters link a of path p during interval k .

$\xi_a^{rsp} = 1$ if link a is the first link of path p , $\xi_a^{rsp} = 0$ otherwise. $v_b^{rsp}(k)$

is the flow rate that exits from link b of path p during interval k .

$\xi_{ba}^{rsp} = 1$ if link b is the predecessor of link a on path p and equals zero otherwise.

In this way, all the inflow rate, exit flow rate and queue size can be expressed as functions of $f_p^{rs}(k)$.

4.2.5 Path Cost

Since the model is based on “actual” travel time. The route travel time during each interval k is calculated using the following nested function:

$$t_p^{rs}(k) = t_{a_1}(k) + t_{a_2}(k + t_{a_1}(k)) + \dots + t_{a_m}(k + t_{a_1} + t_{a_2} + \dots + t_{a_{m-1}}) \quad (4.17)$$

in which, $t_{a_1} = t_{a_1}(k)$, $t_{a_2} = t_{a_2}(k + t_{a_1}(k))$, ..., for short.

In this model, scheduled arrival time at destination is specified. If actual arrival time is greater than the scheduled arrival time window, cost of schedule delay time-late should be considered. On the counterpart, cost of schedule delay time-early should be considered when actual arrival time is less than the scheduled arrival time window. Smith (1982) discussed the modeling of schedule delay cost. In general, path cost function can be defined as follows:

$$c_p^{rs}(k, \mathbf{f}) = \alpha \cdot t_p^{rs}(k) + \begin{cases} \beta \cdot [k_{rs}^* - \Delta_{rs} - k - t_p^{rs}(k)] & \text{if } k + t_p^{rs}(k) < k_{rs}^* - \Delta_{rs} \\ \gamma \cdot [k + t_p^{rs}(k) - k_{rs}^* - \Delta_{rs}] & \text{if } k + t_p^{rs}(k) > k_{rs}^* + \Delta_{rs} \\ 0 & \text{otherwise} \end{cases} \quad (4.18)$$

where α is the unit cost of travel time, β is the unit cost of schedule delay time-early, γ is the unit cost of schedule delay time-late, and $[k_{rs}^* - \Delta_{rs}, k_{rs}^* + \Delta_{rs}]$ is the window of arrival times at destination s without the schedule delay penalty from origin r . Lemma 2 in Huang and Lam (2002) established that cost function $c_p^{rs}(k, \mathbf{f})$ is a continuous function of \mathbf{f} .

4.3 EXISTENCE OF EQUILIBRIUM

The existence of the dynamic user-equilibrium is established by Brouwer's Fixed Point Theorem stated below.

【Theorem 4.1】 (Brouwer's Fixed Point Theorem) Let C be a nonempty convex compact subset of Ω and let $\mathbf{T}:\Omega \rightarrow \Omega$ be a continuous function. Then \mathbf{T} has a fixed point, *i.e.*, there is a point $\mathbf{f} \in \Omega$ such that $\mathbf{T}(\mathbf{f}) = \mathbf{f}$.

Smith (1979, 1993) proved the equivalence between the variational inequality problem and fixed point problem under both static and dynamic situations. More details on the relationship between variational inequality problem and fixed point problem can be found in Harker and Pang (1990) and Pang (1994). This relationship indicates that the solution of the variational inequality problem (4.4), \mathbf{f}^* is the same as the solution of the fixed point problem, *i.e.*, $\mathbf{T}(\mathbf{f}^*) = \mathbf{f}^*$. Here, the map \mathbf{T} is defined as follows:

$$\mathbf{T} = Proj_{\Omega}(I - \rho\mathbf{c}) \quad (4.19)$$

where I is the identity map on Ω , ρ is any positive constant. $Proj_{\Omega}(\mathbf{x})$ is the optimal solution to the problem $\min_{\mathbf{z} \in \Omega} \|\mathbf{z} - \mathbf{x}\|$.

Thus, to establish the existence condition of the variational inequality problem, one only needs to satisfy the condition in the fixed point problem, that is, let $\mathbf{T} = Proj_{\Omega}(I - \rho\mathbf{c})$ be a continuous map. Specifically, we need to express $c_p^{rs}(k, \mathbf{f})$ as a continuous function of \mathbf{f} .

The technique to express $c_p^{rs}(k, \mathbf{f})$ as a continuous function of \mathbf{f} is a linear interpolation as shown in Huang and Lam (2002). Here, one introduces the procedure below.

$c_p^{rs}(k, \mathbf{f})$ is a continuous function of $t_p^{rs}(k)$ as shown in (4.18). From Equation (4.17), we find that the path travel time $t_p^{rs}(k)$ is a continuous function of each related link travel time $t_a(w)$. w is the clock time and the computation is stated in (4.17). We note that w may not be an integer number.

For any integer w , $t_a(w)$ is a linear function of $q_a(w)$ as shown in (4.8). $q_a(w)$ is a piecewise-linear function of $q_a(w-1)$ and $u_a(w)$ is as shown in (4.13). Recursively applying (4.13), we find $q_a(w)$ is a continuous function of all $u_a(k)$ ($k \leq w$). From the link conservation Equations (4.14) and (4.15), $u_a(k)$ is a continuous function of $f_p^{rs}(k)$ and $v_b^{rsp}(k)$. Then, we need to express $v_b^{rsp}(k)$ as a continuous function of $f_p^{rs}(k)$. From Equation (4.16), $v_b^{rsp}(k)$ is expressed as a function of $u_b^{rsp}(i)$, where i satisfies $i + t_b(i) = k$. If such i is not an integer, we can let $u_b^{rsp}(i)$ be the linear interpolation of two adjacent flow rates associated with the nearest integer of i . Denote $\text{floor}(i)$ as the largest integer smaller than i and $\text{ceil}(i)$ the smallest integer larger than i . For example, if $i = 1.4$, then $\text{floor}(i) = 1$ and $\text{ceil}(i) = 2$. $u_b^{rsp}(i)$ is then given by the following equation:

$$u_b^{rsp}(i) = u_b^{rsp}(\text{floor}(i)) + [i - \text{floor}(i)] \cdot [u_b^{rsp}(\text{ceil}(i)) - u_b^{rsp}(\text{floor}(i))] \quad (4.20)$$

This procedure is shown in Figure 4.1:

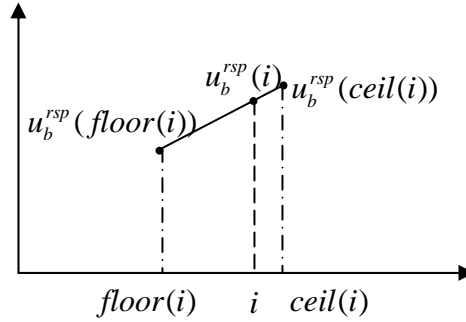


Figure 4.1 An Example of Linear Interpolation

Recursively using Equations (4.14), (4.15) and (4.16), finally, all link flows are expressed as a continuous function of path flow $f_p^{rs}(k)$.

For any non-integer w , $t_a(w)$ is still a linear function of $q_a(w)$ from Equation (4.8). The linear interpolation technique still can be used to express $q_a(w)$ as a continuous function of $u_b^{rsp}(k)$.

4.4 MODEL MODIFICATION

This section provides an improvement on the calculation of link exit rate which is different from the approximating method adopted by Huang and Lam (2002).

Link exit flow $v_b^{rsp}(k)$ as shown in Equation (4.16) equals to $s_b \frac{u_b^{rsp}(i)}{u_b(i)}$ when

there is a queue at interval $k - t_b^0$ and $i + t_b(i) = k$. In Huang and Lam (2002), when $u_b(i) = 0$, $v_b^{rsp}(k)$ is estimated by taking an average of s_b on all paths that use link b at interval i . In this research study, a method to compute $v_b^{rsp}(k)$ without such approximation is proposed as follows.

Step 1. Set $k = K$ and $v_b^{rsp}(K) = 0$, then eventually, all traffic leave the network

Step 2. Set $k = k - 1$. If the queue is null at interval $k - t_b^0$,

$$v_b^{rsp}(k) = u_b^{rsp}(k - t_b^0)$$

If there is a queue at interval $k - t_b^0$, search the first i from $i = 1$ to

k such that $i + t_b(i) = k$. Let $v_b^{rsp}(k) = s_b \frac{u_b^{rsp}(i)}{u_b(i)}$, if such i cannot

be found, let $v_b^{rsp}(k) = v_b^{rsp}(k + 1)$

Step 3. Repeat step 2 until k reaches t_b^0

This method works due to two reasons. First, for the first i satisfying step 2, $u_b(i)$ would not be zero. Second, the equation $v_b^{rsp}(k) = v_b^{rsp}(k + 1)$ in the step 2 is valid because of the constant exit rate assumption made at the beginning of Section 4.2.2.

We use an example to show this method.

Consider only one link a with a given inflow rate $u_a(k)$ as shown in Table 4.1:

Table 4.1 Link Inflow Rate

Time interval k	1	2	3	4	5	6	7
Inflow rate	$0.5 s_a$	0	$2 s_a$	$2 s_a$	0	0	0

$u_a(k)$							
----------	--	--	--	--	--	--	--

The free-flow travel time is $t_a^0 = 3$. Applying Equation (4.13), we have the following results tabulated in Table 4.2.

Table 4.2 The Resulting Queue and Travel Time

Time interval k	1	2	3	4	5	6	7
Queue $q_a(k)$	0	0	δs_a	$2\delta s_a$	δs_a	0	0
Travel time $t_a(k)$	3	3	4	5	4	3	3

As queue is already computed in Table 4.2 using the proposed method, one can get the following link exit flow rate given in Table 4.3.

Table 4.3 Link Exit Flow

k	...	4	5	6	7	8	9	...
$v_b(k)$	0	$\frac{1}{2}s_a$	0	s_a	s_a	s_a	s_a	0

4.5 ROUTE/TIME SWAPPING ALGORITHM

The “route/time swapping” process used to solve the above VI problem is based on the projected dynamical systems introduced by Nagurney and Zhang (1997a, 1997b). The essence of this process is to adjust the flows on the non-cheapest time-dependent paths to the cheapest time-dependent paths. The volume moved is proportional to $f_p^{rs}(k) \times [c_p^{rs}(k, f) - c_{\min}^{rs}(f)]$.

The step-by-step algorithm suggested by Huang and Lam (2002) is as follows:

Step 1. Set the iteration index $\tau=1$, choose the initial path inflow rate

$$f_p^{rs}(k)_\tau \text{ in } \Omega.$$

Step 2. Calculate the link inflow rate $u_a(l)_\tau$ by (4.14)-(4.16).

Step 3. Compute the path travel time $t_p^{rs}(k)_\tau$ by (4.17), and the path cost

$$c_p^{rs}(k, \mathbf{f})_\tau \text{ by (4.18). Find } c_{\min}^{rs}(\mathbf{f})_\tau = \min\{c_p^{rs}(k, \mathbf{f})_\tau : p \in P_{rs}, k \in K\}$$

$$\text{and let } \hat{P}_\tau = \{(p, k) : c_p^{rs}(k, \mathbf{f})_\tau = c_{\min}^{rs}(\mathbf{f})_\tau, p \in P_{rs}, k \in K\}.$$

Step 4. Update the path inflow rates as below:

$$f_p^{rs}(k)_{\tau+1} = f_p^{rs}(k)_\tau - \rho_\tau f_p^{rs}(k)_\tau [c_p^{rs}(k, \mathbf{f})_\tau - c_{\min}^{rs}(\mathbf{f})_\tau], \quad p \in P_{rs}, k \in K, \quad (4.21)$$

$$f_p^{rs}(k)_{\tau+1} = f_p^{rs}(k)_\tau + \frac{\psi_\tau}{|\hat{P}_\tau|} \text{ for } (p, k) \in \hat{P}_\tau, \quad (4.22)$$

$$\text{where } \psi_\tau = \sum_{p \in P_{rs}, k \in K} \rho_\tau f_p^{rs}(k)_\tau [c_p^{rs}(k, \mathbf{f})_\tau - c_{\min}^{rs}(\mathbf{f})_\tau].$$

Step 5. The iteration terminates if

$$I = \frac{\sum_k \sum_{rs} \sum_p f_p^{rs}(k) [c_p^{rs}(k, \mathbf{f}) - c_{\min}^{rs}(\mathbf{f})]}{\sum_k \sum_{rs} \sum_p f_p^{rs}(k) c_{\min}^{rs}(\mathbf{f})} < \varepsilon, \quad (4.23)$$

where ε is pre-determined tolerance level. Otherwise, go to step 2.

I is called the convergence indicator.

4.6 NUMERICAL EXAMPLE FOR “ROUTE/TIME SWAPPING” ALGORITHM

To show the application of the “route/time swapping” algorithm, the original nine-node grid network from Huang and Lam (2002) as shown in Figure 4.2 is used. Link parameters are given in Table 4.4.

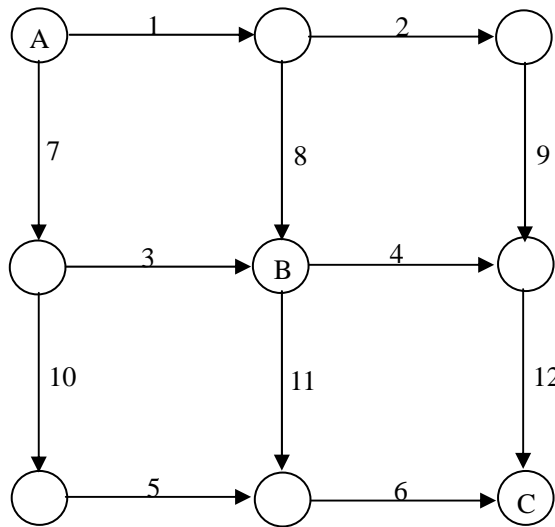


Figure 4.2 Grid Network

Table 4.4 Link Parameters

Link	Free-flow Time (hr)	Maximum Exit Rate (veh/hr)
1	0.2	8000
2	0.3	4000
3	0.2	7000
4	0.2	8000
5	0.3	3000
6	0.2	7000
7	0.2	8000
8	0.2	7000
9	0.3	3000
10	0.3	4000
11	0.2	8000
12	0.2	7000

There are two O-D pairs: $A \rightarrow C$ and $B \rightarrow C$. The eight paths are listed in Table 4.5. The demand over the total time period is: $F^{AC} = 20,000$ veh, $F^{BC} = 10,000$ veh. Let the total time period be $T = 4$ hours representing from 6:00 am to 10:00 am. $K = 400$. So, for each time interval, $\delta = 0.6$ minutes. Other parameters are set as $\alpha = \$6.4$ /hr, $\beta = \$3.9$ /hr, $\gamma = \$15.21$ /hr, $k^* = 9.00$ am

Table 4.5 Path List

Path Number	OD pair	Path
1	$A \rightarrow C$	(1, 2, 9, 12)
2		(1, 8, 4, 12)
3		(1, 8, 11, 6)
4		(7, 3, 4, 12)
5		(7, 3, 11, 6)
6		(7, 10, 5, 6)
7	$B \rightarrow C$	(4, 12)
8		(11, 6)

Figure 4.3 shows the convergence indicator as defined in Equation 4.23. The convergence indicator descends to 0.15 after 50,000 iterations. The computing time is around 15 minutes.

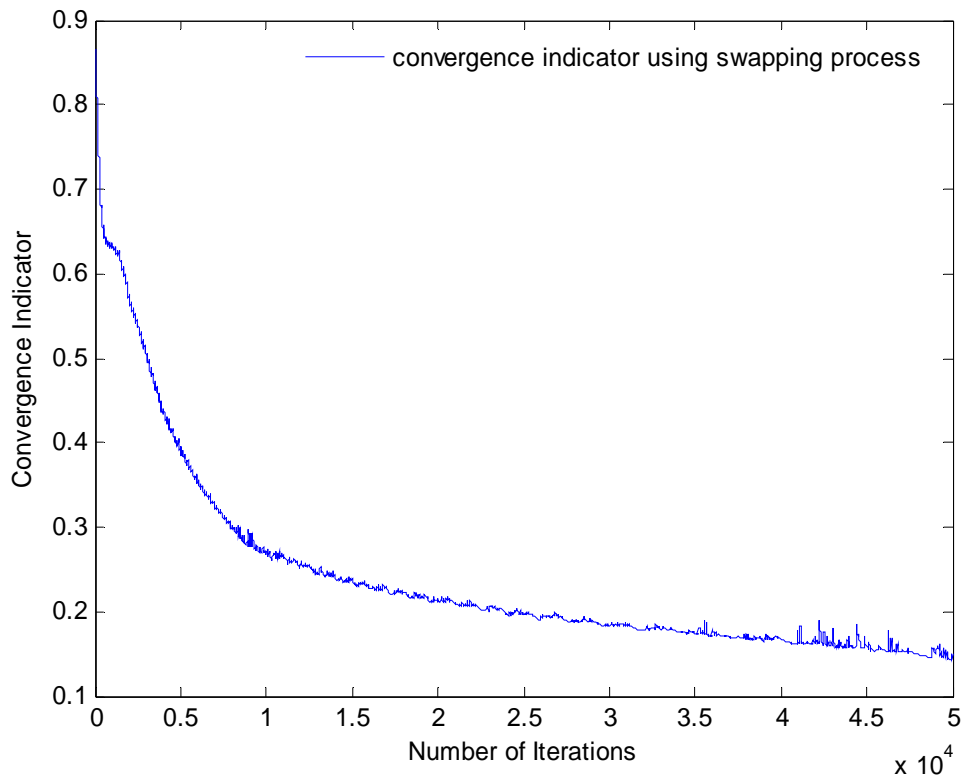


Figure 4.3 Convergence of Route/time Swapping Algorithm

Figure 4.4 and Figure 4.5 show the inflow rate and path cost on paths 1, 2, and 7. As this is a symmetric network, symmetrical results should be obtained, i.e., the inflow rate and travel cost on path 1 are the same as those on path 6; path 2 the same as paths 3, 4, 5, and lastly, path 7 the same as path 8. Figure 4.6 shows the queue length on the network.

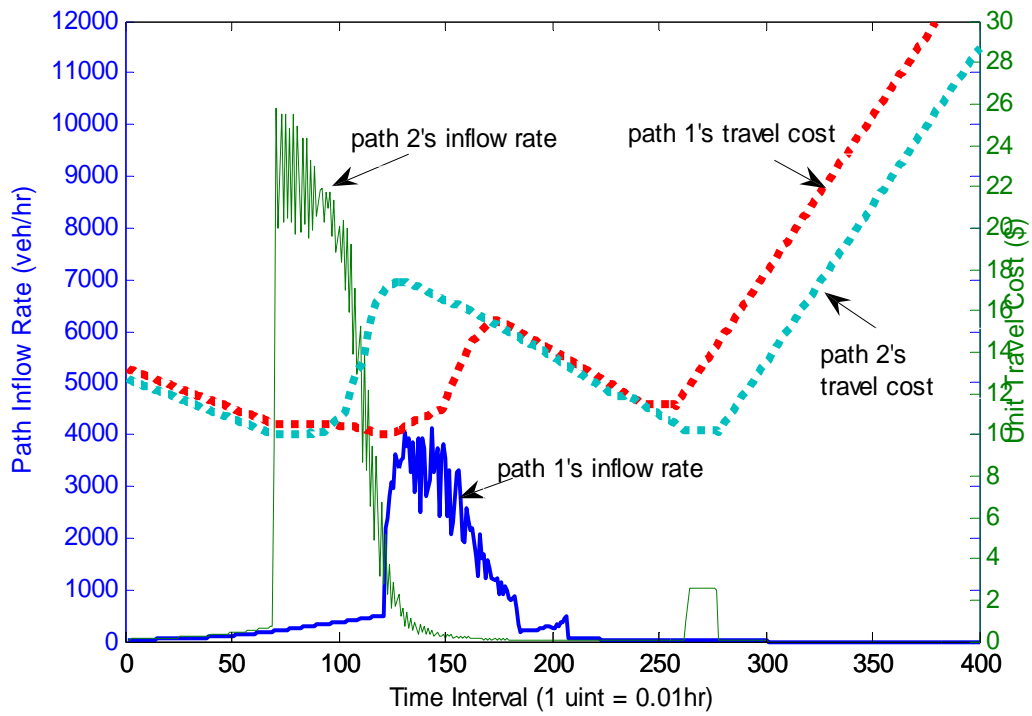


Figure 4.4 Path Inflow Rate and Travel Cost by Using Route/time Swapping Algorithm (Path 1 and 2)

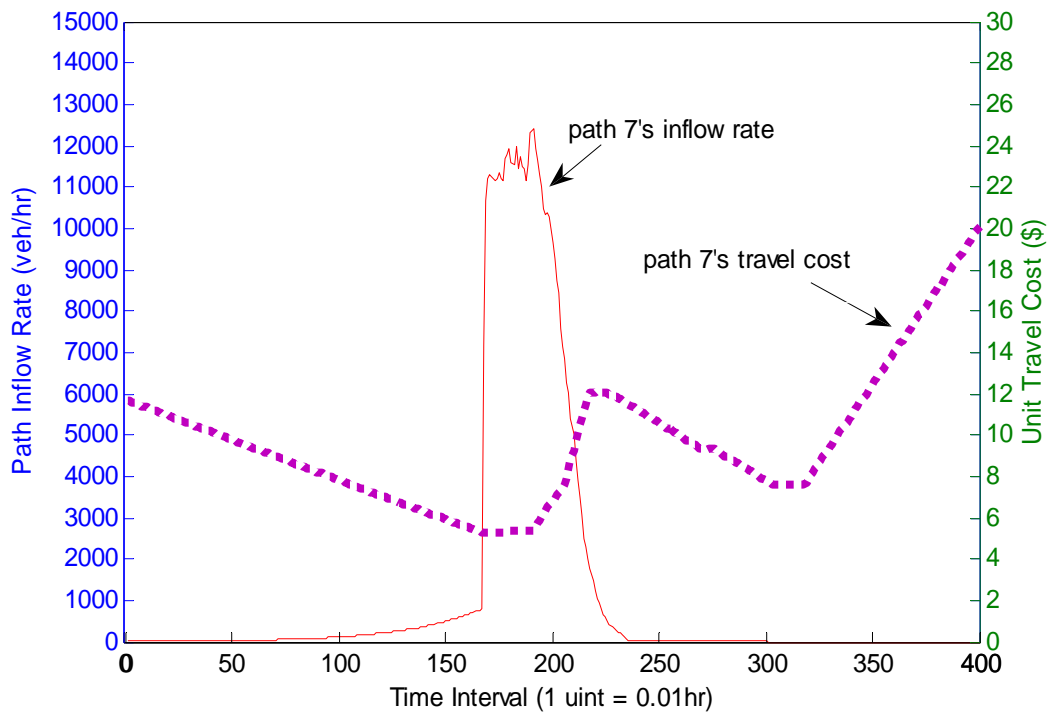


Figure 4.5 Path Inflow Rate and Travel Cost by Using Route/time Swapping Algorithm (Path 7)

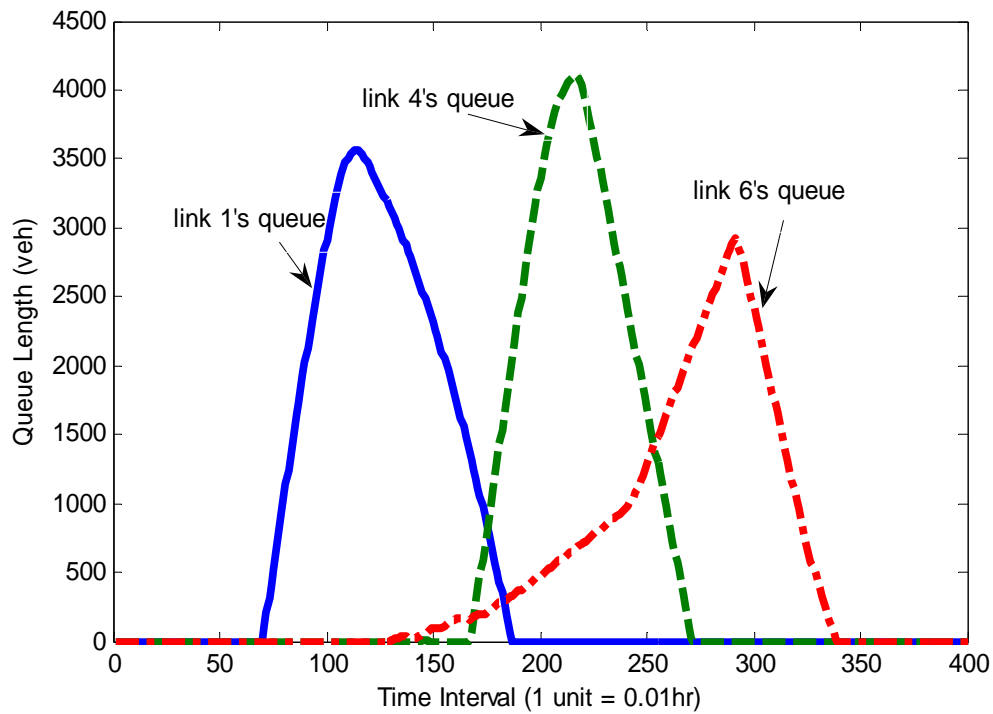


Figure 4.6 Queue Length by Using Route/time Swapping Algorithm

4.7 A PRIME PROJECTION-BASED ALGORITHM

In this section, a new algorithm based on a projection method is proposed. A few issues related to this algorithm are also discussed. We will show that this algorithm does not guarantee convergence.

4.7.1 Algorithm Description

We have shown that DTA can be formulated as a VI problem and note that the constraint of (4.5) is a simple convex set. Quadratic programming on convex set is relatively simple to obtain solutions. Bearing this in mind, we try the

conventional projection method for the monotone VI problem.

For the monotone $VI(F, C)$, starting from an initial \mathbf{x}_0 , the projection method is to find a new point iteratively following the equation:

$$\mathbf{x}^{k+1} = P_C(\mathbf{x}^k - \tau F(\mathbf{x}^k)) \quad (4.24)$$

where $P_C(\cdot)$ denotes the orthogonal projection map onto C and τ is a judiciously chosen positive step length. Here, $P_C(\mathbf{x}^k - \tau F(\mathbf{x}^k))$ is the solution of the following quadratic programming problem:

$$\min_{\mathbf{x} \in C} \frac{1}{2} \mathbf{x}^T \mathbf{x} - (\mathbf{x}^k - \tau \mathbf{F}(\mathbf{x}^k))^T \mathbf{x} \quad (4.25)$$

However, the projection method requires a strong assumption on F for convergence. \mathbf{F} is needed to be strong monotone and Lipschitz continuous for a theoretical convergence. In case of no convergence, the convergence indicator suggested by Huang and Lam (2002) as in (4.23) is used to observe the convergence status. The indicator is defined as:

$$I = \frac{\sum_k \sum_{rs} \sum_p f_p^{rs}(k) [c_p^{rs}(k, \mathbf{f}) - c_{\min}^{rs}(\mathbf{f})]}{\sum_k \sum_{rs} \sum_p f_p^{rs}(k) c_{\min}^{rs}(\mathbf{f})} \quad (4.26)$$

The step-by-step procedure of the algorithm is given below:

Step 1. Set the iteration index $k = 0$ and choose the initial path inflow rates \mathbf{f}^k from Ω as defined by Equation (4.5).

Step 2. Calculate the path cost $\mathbf{c}(\mathbf{f}^k)$ and the resulting convergence indicator

I^k . If $I^k \leq \epsilon$, stop. ϵ is a predetermined tolerance value. Otherwise, go to step 3.

Step 3. Calculate the projection of $\mathbf{f}^k - \tau\mathbf{c}(\mathbf{f}^k)$ onto Ω . *i.e.* solve the

quadratic programming problem $\min_{\mathbf{x} \in \Omega} \frac{1}{2} \mathbf{x}^T \mathbf{x} - (\mathbf{f}^k - \tau\mathbf{c}(\mathbf{f}^k))^T \mathbf{x}$. Denote the solution as $\bar{\mathbf{f}}$.

Step 4. Set $\mathbf{f}^{k+1} = \bar{\mathbf{f}}$ and $k \leftarrow k + 1$; go to step 2

Note that the tolerance value ϵ should be selected carefully. As convergence is not guaranteed, avoid using small value of ϵ that will not lead to a stable solution. We will illustrate this point in the example in the following section. The network example and parameters setting are same as in Section 4.6.

4.7.2 The Selection of Tolerance ϵ

As pointed out in the last section, \mathbf{F} needs to be strong monotone and Lipschitz continuous for theoretical convergence. The absolute convergence conditions can not be meet in this example if too small a value of ϵ was chosen as illustrated in the computed results shown in Figure 4.7. In fact, the convergence indicator defined in Equation (4.26) oscillates nearly between 0.1 and 0.6 when the step length τ is set as 10. Also, during the repeated up and down process, it can reach to a minimum value of 0.1 at some iteration. Thus, it is reasonable to set ϵ to be 0.1 in this example. Thus, the value of $W(\mathbf{f})$ in Equation (4.7) never reach zero. Nevertheless, it oscillates the same way as

the convergence indicator so that when the convergence indicator reaches to 0.1, $W(\mathbf{f})$ achieves its minimum.

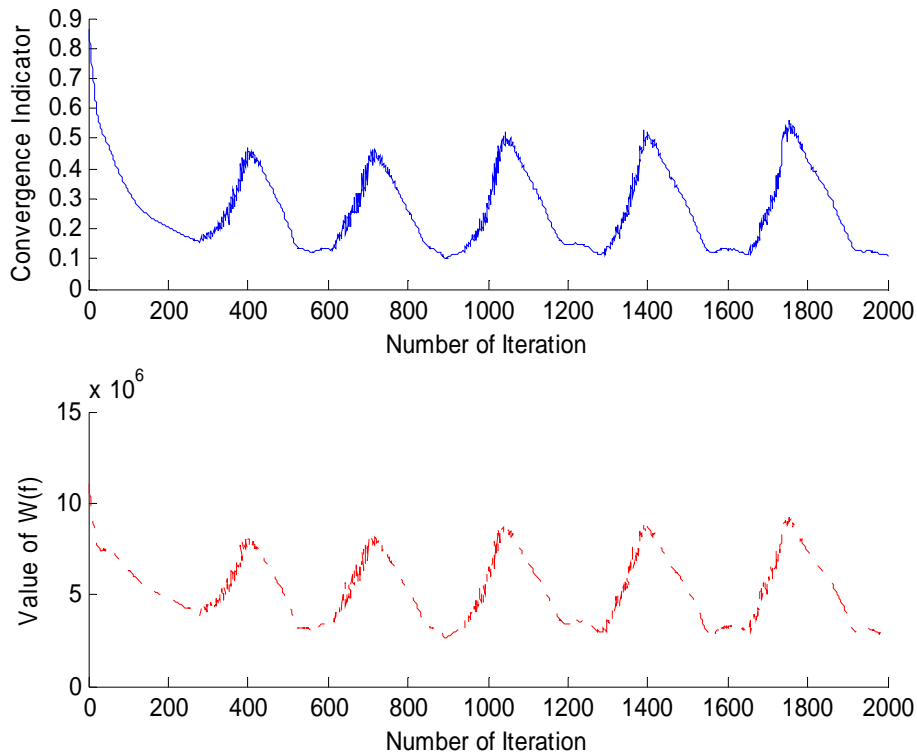


Figure 4.7 Value of Convergence Indicator and $W(\mathbf{f})$ under too Small Tolerance ϵ

4.7.3 The Selection of Step Length τ

In the conventional projection method, the selection of step length τ is a pre-determined parameter. Figure 4.8 shows the results of the convergence indicator when different step lengths of 1, 10 and 50 were used. For a step length of 1, the convergence indicator descends very smoothly, but at a very slow speed. After 2000 iterations, it just descended to 0.2. When step length is set larger to 10 or 50, it descends faster at the expense of frequent oscillation. But too large a step length will result in an increase of the minimal setting of convergence indicator.

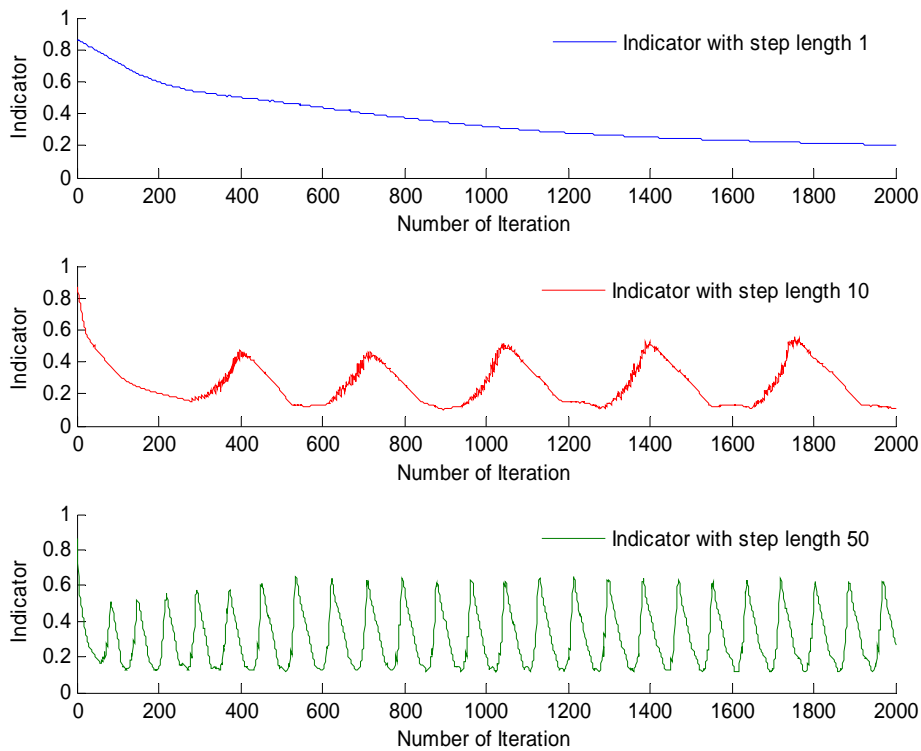


Figure 4.8 Convergence Indicator with Different Step Lengths (1)

As Figure 4.9 shows, the larger the step length, the bigger the minimal value of the convergence indicator achieved. For step length 200, the minimal value of convergence indicator is below 0.2. The convergence indicator is larger than 0.2 for step length 400. So we are facing a dilemma whereby on one side we want to get the minimal iteration by increasing step length and on the other side we have to keep the precision by using shorter step lengths.

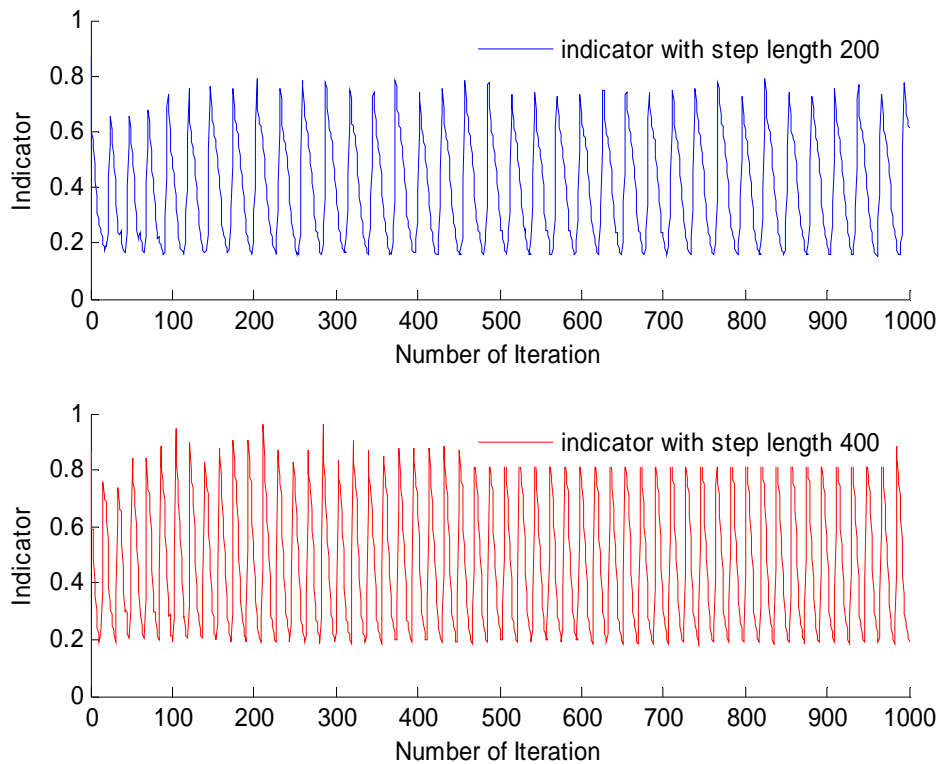


Figure 4.9 Convergence Indicator with Different Step Lengths (2)

4.8 PROJECTION-BASED ALGORITHM WITH MSA AND EXAMPLE

As discussed in the previous section, the pure projection method can not guarantee convergence of the algorithm because of the ill-behaved objective function. To make the projection method run well, we have to choose the tolerance ϵ appropriately. If the tolerance chosen is too large, the results are not accurate, i.e. there is still a potential to make $W(\mathbf{f})$ smaller. If tolerance is chosen to be too small, there is no ending for the algorithm. We propose another mechanism to achieve convergence by employing the Method of Successive Averages (MSA) to the projection method. The MSA is applied in the last step of the algorithm.

The step-by-step procedure of the algorithm is given below:

Step 1. Set the iteration index $k=0$ and choose the initial path inflow rates \mathbf{f}^k from Ω as defined by Equation (4.5).

Step 2. Calculate the path cost $\mathbf{c}(\mathbf{f}^k)$ and the resulting convergence indicator I^k . If $I^k \leq \epsilon$, stop. ϵ is a predetermined tolerance value. Otherwise, go to step 3.

Step 3. Calculate the projection of $\mathbf{f}^k - \tau \mathbf{c}(\mathbf{f}^k)$ onto Ω . *i.e.* solve the quadratic programming problem $\min_{\mathbf{x} \in \Omega} \frac{1}{2} \mathbf{x}^T \mathbf{x} - (\mathbf{f}^k - \tau \mathbf{c}(\mathbf{f}^k))^T \mathbf{x}$. Denote the solution as $\bar{\mathbf{f}}$.

Step 4. Set $\mathbf{f}^{k+1} = \mathbf{f}^k + \alpha^n (\bar{\mathbf{f}} - \mathbf{f}^k)$ and $k \leftarrow k + 1$; go to step 2.

The network example and parameters setting are the same as those in Section 4.6.

$$\alpha^n \text{ satisfy } \sum_{n=1}^{\infty} \alpha^n = \infty,$$

$$\alpha^n = \frac{1}{p} (p=1, n=1 \rightarrow 1000; p=2, n=1001 \rightarrow 2000; \dots), \quad \tau = 1.$$

After 8000 iterations the convergence indicator descends to 0.1534. The computing time is around 35 minutes. The convergence indicator is shown in Figure 4.10.

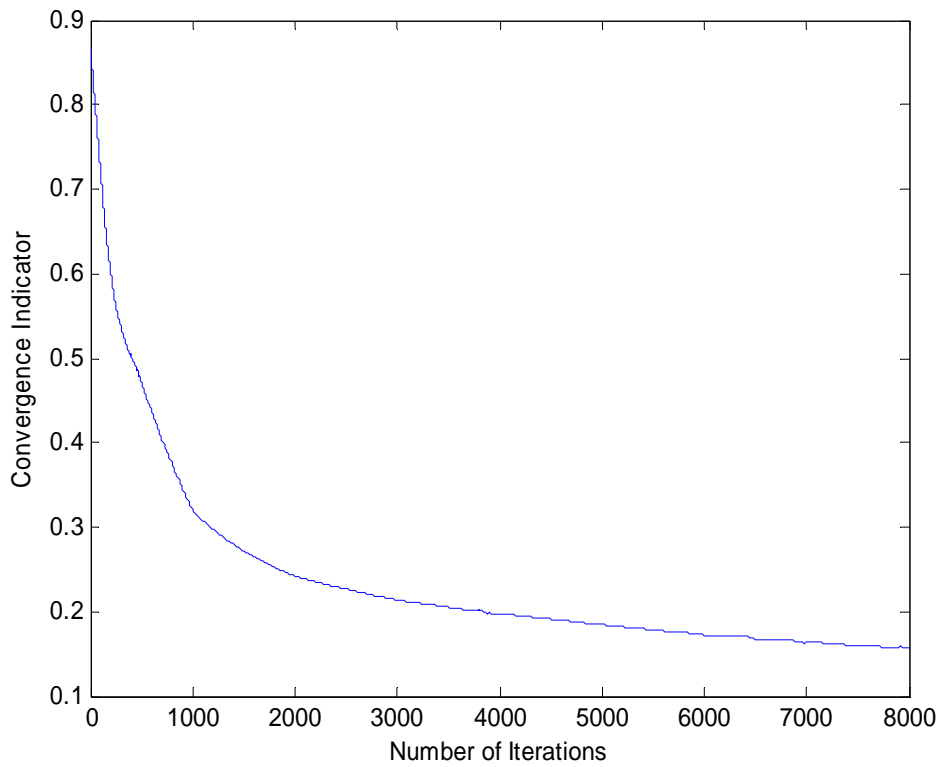


Figure 4.10 Convergence Indicator by Using Projection-based Algorithm with MSA

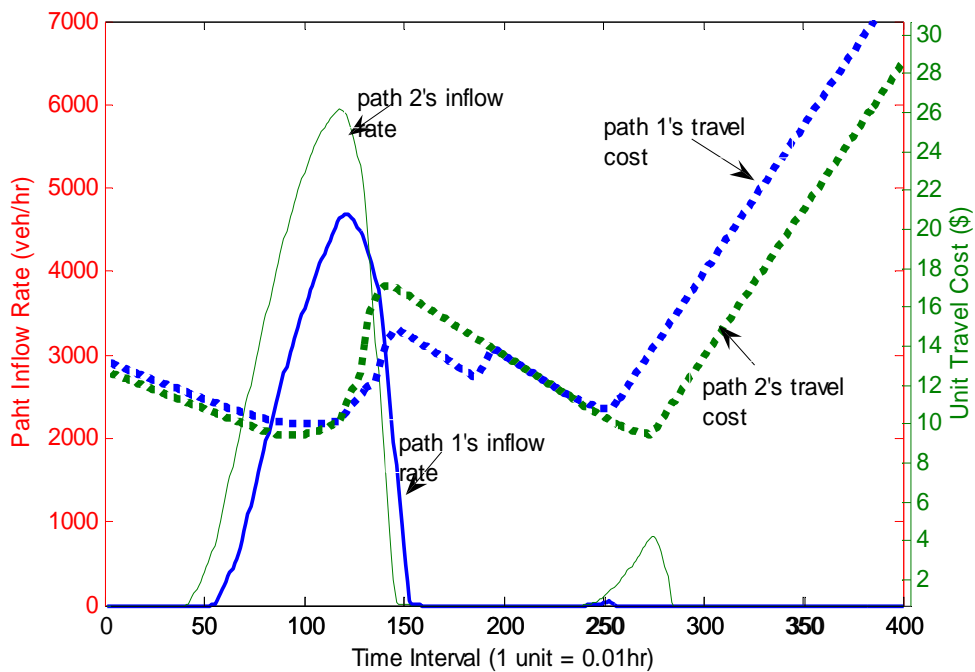


Figure 4.11 Path Inflow Rate and Travel Cost by Using Projection-based Algorithm with MSA (Path 1 and 2)

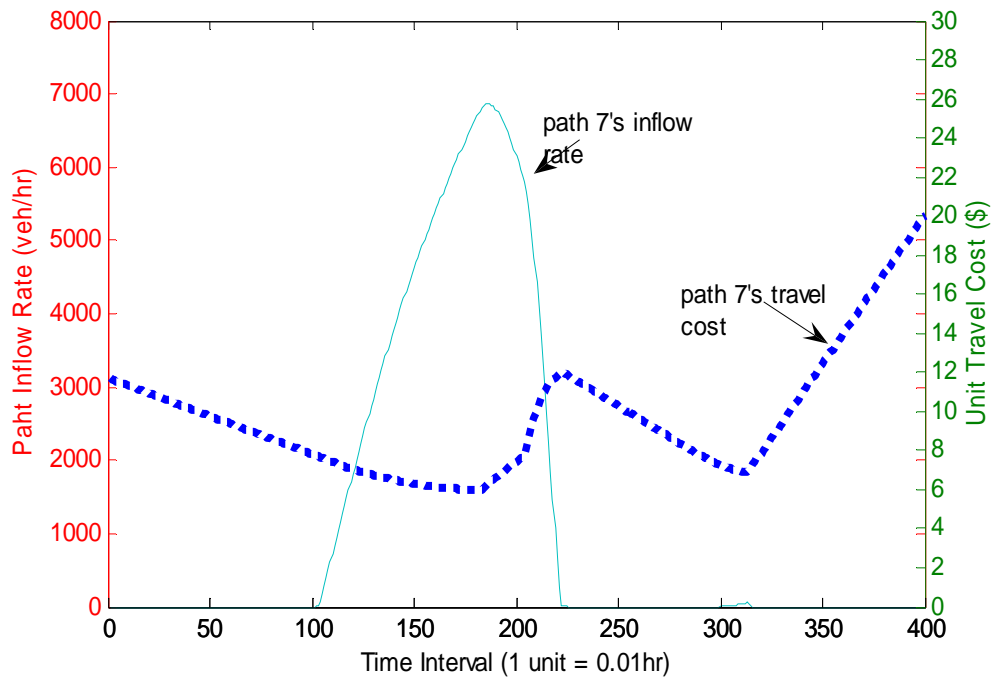


Figure 4.12 Path Inflow Rate and Travel Cost by Using Projection-based Algorithm with MSA (Path 7)

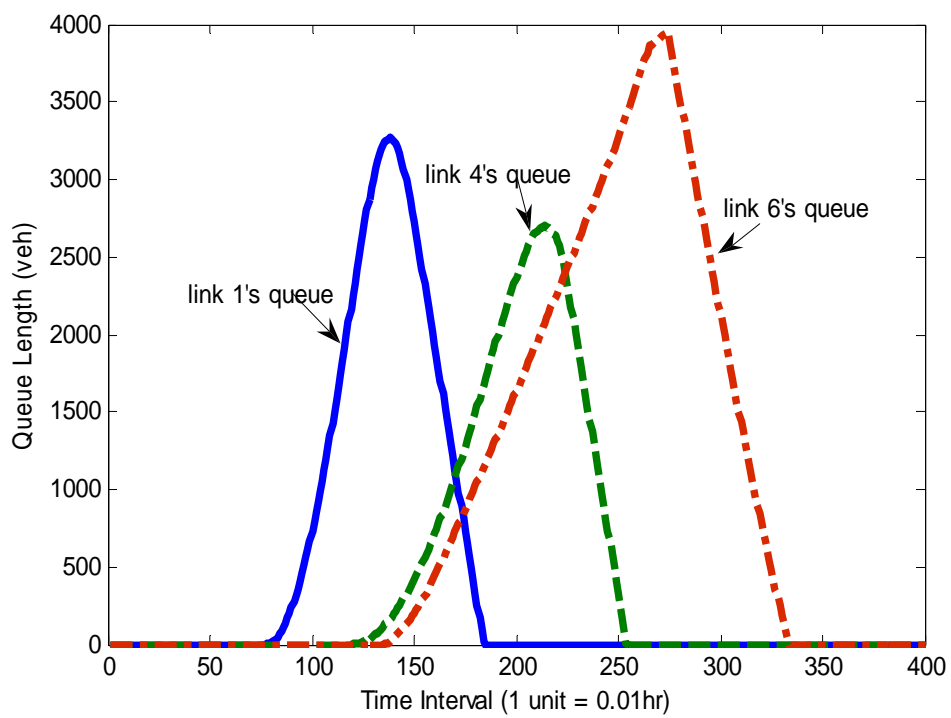


Figure 4.13 Queue Length by Using Projection-based Algorithm with MSA

The selection of parameters α^n can be given with a certain level firstly. Adjusting of parameters can be conducted after observing the behavior of Convergence Indicator I . If I goes like Figure 4.7, then a smaller pace of α^n is needed. If I converges too slowly, then a larger pace of α^n is needed.

4.9 ALGORITHMS COMPARISON

Comparing Figure 4.4 with Figure 4.11, we note that both algorithms verify the dynamic user equilibrium discipline. Travelers always choose the departure time when the travel cost is the lowest. However, the results using projection-based algorithm are smoother than those obtained by the “route/time swapping” algorithm; that is, the path inflow rate oscillates too rapidly within very short time interval using the route/time swapping algorithm (see Figure 4.4 and Figure 4.5). For realistic computation of results, we choose the projection-based algorithm with MSA to solve the parametric VI problems encountered in the dynamic road pricing problem formulation which is discussed in the next chapter.

The quadratic programming problem (Equation 4.25) is solved by a well-known nonlinear optimization solver CONOPT (Drud, 1992) as a sub-system in GAMS. The computation is coded by MATLAB language integrated with GAMS (Ferris, 1998). About an hour of processing time is needed to complete the computation using P4 personal computer.

4.10 SUMMARY

In order to represent a dynamic traffic evolution on a network for the design of dynamic road pricing, a basic DTA model with queue from Huang and Lam (2002) has been introduced. This basic DTA model satisfies the FIFO condition and travelers are allowed to choose their departure times. The apparent weaknesses of the basic model are pointed out and corrected. A new procedure to compute some inter-variables is proposed. These modifications can represent traffic evolution more appropriately which is demonstrated through a counter example. Under the original “route/time” swapping algorithm (Huang and Lam, 2002), the traffic flow can oscillate rapidly within a very short time. As the constraint of the DTA model is a simple convex set and quadratic programming on a convex set is relatively simple to obtain solutions, a new algorithm based on projection method has been designed. It is found that this projection based algorithm can generate a smoother traffic flow than the “route/time” swapping algorithm. The primary projection-based algorithm does not guarantee convergence. However, the other projection-based algorithm uses MSA to achieve stable solution. Note that if the tolerance is selected appropriately, the primary projection-based algorithm can achieve solution using less computation time. In the next Chapter, the dynamic road pricing problem will be constructed incorporating DTA. The projection-based algorithm with MSA will be used to solve the parametric VI problem.

CHAPTER 5 DYNAMIC ROAD PRICING MODEL

5.1 INTRODUCTION

In this Chapter, the dynamic road pricing model is formulated as a MPEC problem. The only road pricing objective is to achieve a minimal total queue size on a network. In this model, toll is applied per unit time. Thus, the amount of toll depends on the length of stay on the tolled link. The evaluation of the toll impact is firstly introduced in this chapter. The traffic equilibrium conditions in MPEC are represented by the DTA model which has been discussed in Chapter 4. DE-Toll and PS-Toll introduced in Chapter 3 for MPEC problem are used as the solving algorithm. A numerical example is used to show the difference of these two algorithms. Sensitivity analysis is carried for pre-determined parameters. The network used in the example contains only one single class of roads. One single class of travelers is also assumed.

5.2 MPEC FORMULATION

The dynamic pricing problem aims to satisfy some system performance. For example, it can be a reduction of total queues on the network, a decrease in total travel time and so on. Henceforth, the minimization of total queues on the network is taken as the main objective. Before discussing on the setup of the MPEC, how the toll levels interfere with travel cost would be clarified. One can assume that the toll level increases directly with the path cost incurred by the travelers when using the tolling link during the tolling time period. It is

expressed below.

$c_p^{rs}(k, \mathbf{f})$ is the predictive path cost entering path p during the time interval k . Suppose toll is applied on link $b = (i_1, i_2)$. i_1 is the head of link b and i_2 is the tail. The tolling time period is $[m, n]$. The toll level is denoted as $\$ tol(j)$ per unit time for $j \in [m, n]$. m is the time interval when tolling begins and n is the time interval when tolling ends. $Tc_p^{rs}(k, \mathbf{f})$ denotes the total travel cost after considering the tolling effect. For those path not using the tolling link b , $Tc_p^{rs}(k, \mathbf{f}) = c_p^{rs}(k, \mathbf{f})$. For those path using the tolling link b , $Tc_p^{rs}(k, \mathbf{f})$ is expected to include the tolling effect. Similar to Equation (4.17), one can calculate the sub-path travel time from an origin r to node i_1 when entering during the interval k . Denote it as $t_p^{r i_1}(k, \mathbf{f})$. Thus, the clock time arriving node i_1 can be obtained as $ct_p^{r i_1}(k, \mathbf{f}) = k + t_p^{r i_1}(k, \mathbf{f})$. The clock time when passing the link b can also be obtained as $ct_p^{r i_2}(k, \mathbf{f}) = ct_p^{r i_1}(k, \mathbf{f}) + t_b(ct_p^{r i_1}(k, \mathbf{f}))$. $ct_p^{r i_1}(k, \mathbf{f})$ can be before, between or after the tolling time window $[m, n]$. The total travel cost in each situation will be described below.

If $ct_p^{r i_1}(k, \mathbf{f}) < m$, the total travel cost can be written as:

$$Tc_p^{rs}(k, \mathbf{f}) = \begin{cases} c_p^{rs}(k, \mathbf{f}) & \text{if } ct_p^{r i_2}(k, \mathbf{f}) < m \\ c_p^{rs}(k, \mathbf{f}) + \mu \cdot \sum_{j=m}^{ct_p^{r i_2}(k, \mathbf{f})} tol(j) & \text{if } m \leq ct_p^{r i_2}(k, \mathbf{f}) \leq n \\ c_p^{rs}(k, \mathbf{f}) + \mu \cdot \sum_{j=m}^n tol(j) & \text{if } ct_p^{r i_2}(k, \mathbf{f}) > n \end{cases} \quad (5.1)$$

If $m \leq ct_p^{r_1}(k, \mathbf{f}) \leq n$, the total travel cost can be written as:

$$Tc_p^{rs}(k, \mathbf{f}) = \begin{cases} c_p^{rs}(k, \mathbf{f}) + \mu \cdot \sum_{j=ct_p^{r_1}(k, \mathbf{f})}^{ct_p^{r_2}(k, \mathbf{f})} tol(j) & \text{if } ct_p^{r_2}(k, \mathbf{f}) \leq n \\ c_p^{rs}(k, \mathbf{f}) + \mu \cdot \sum_{j=ct_p^{r_1}(k, \mathbf{f})}^n tol(j) & \text{if } ct_p^{r_2}(k, \mathbf{f}) > n \end{cases}$$

(5.2)

If $n \leq ct_p^{r_1}(k, \mathbf{f})$, obviously, $Tc_p^{rs}(k, \mathbf{f}) = c_p^{rs}(k, \mathbf{f})$. The parameter μ mentioned above is a parameter denoting the impact of the pricing level on travel cost. This parameter describes the sensitivity to road pricing for different classes of travelers.

Thus, the MPEC can be expressed as:

$$\text{Min} \sum_k \sum_a q_a(k) \quad (5.3)$$

\mathbf{f} solves the VI problem

$$\sum_k \sum_{rs} \sum_p c_p^{rs}(k, \mathbf{f}^*) \cdot [f_p^{rs}(k) - f_p^{rs*}(k)] \geq 0 \quad (5.4)$$

Satisfying constraints:

$$\Omega = \left\{ \mathbf{f} \geq 0 : \sum_p \sum_k \delta \cdot f_p^{rs}(k) = F^{rs} \right\} \quad (5.5)$$

Here, $c_p^{rs}(k, \mathbf{f})$ in the VI problem (5.4) should be replaced by $Tc_p^{rs}(k, \mathbf{f})$. Λ

is the set of constraints for the tolling variables.

5.3 NUMERICAL EXAMPLE

In this section, we apply two algorithms DE-Toll and PS-Toll as introduced in Chapter 3 to solve the discrete dynamic road pricing problem. The NLPEC loses its power under DTA environment because travel cost $Tc_p^{rs}(k, \mathbf{f})$ can not be expressed as an analytical function of $f_p^{rs}(k)$.

5.3.1 Application of DE-Toll

As discussed earlier, the DE method is population based and it needs to compute a large amount of traffic distribution under the fixed tolling level. To alleviate the burdensome computation, we set total time interval $K=96$ instead of 400 used in the section 4.6 and $F^{AC}=10,000$ veh, $F^{BC}=5,000$ veh. The parameter $\mu=0.1$. A non-symmetric network is used. The topology of the network is the same as Figure 4.2. However, to represent a non-symmetric property, we simply change the link parameters for the 12 links as shown in Table 5.1.

First, the network is not subject to any toll level. Tolerance ϵ is set to 0.08. One can use the project-based algorithm with MSA to get a heuristics solution. The queue distribution on the network is shown in Figure 5.1.

Table 5.1 Link Parameters for MPEC

Link	Free-flow time (hr)	Maximum Exit Rate (veh/hr)
1	0.2	7000
2	0.3	8000
3	0.2	6000
4	0.2	4000
5	0.3	8000
6	0.3	8000
7	0.3	8000
8	0.2	8000
9	0.2	6000
10	0.3	8000
11	0.3	8000
12	0.2	8000

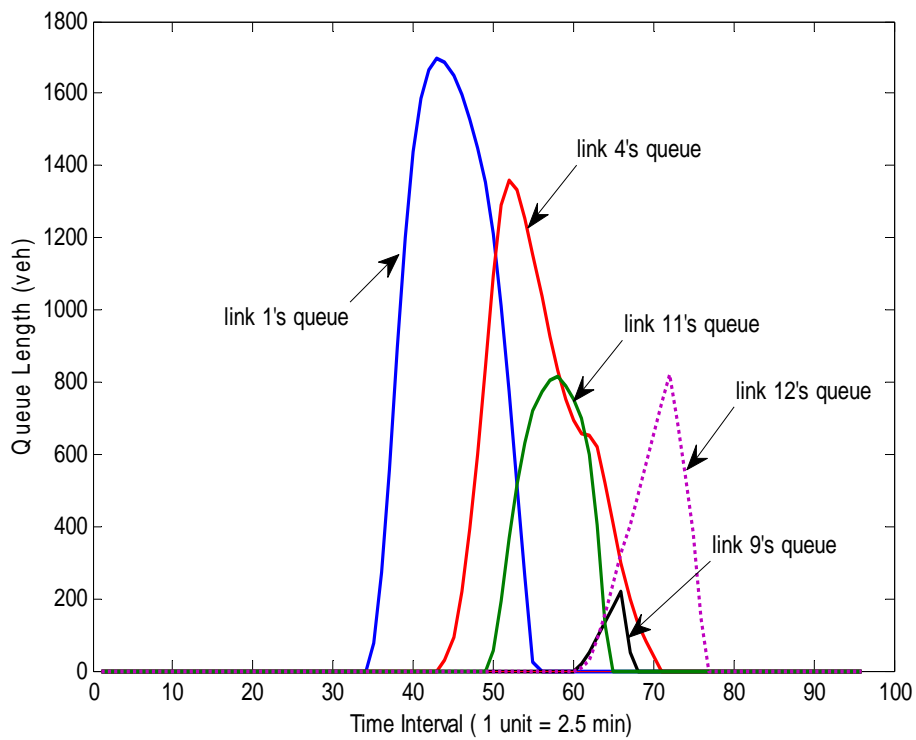


Figure 5.1 Queue Length on the Network without Tolling

The total queue volume $TQ = 55236$ (Veh). From Figure 5.1, it can be seen that link 1 is under the most serious congestion during the time interval [35, 55]. Thus, one possible strategy to alleviate congestion is to toll every 15 minutes beginning from the start of queue. In our setting, 1 unit equals to 2.5 minutes. Thus, 15 minutes equals to 6 units. We choose 3 different time windows; that is, intervals from [36, 41], [42, 47] and [48, 53]. Denote the toll variables $tol(j)$ $j = 1, 2, 3$ for these different time windows. Suppose the toll range is from 0 to 2; that is, $\Lambda = \{tol(j) \in [0, 2] \quad j = 1, 2, 3\}$.

In DE settings, the population size is set as 15, crossover probability 0.8, and DE step-size 0.8. Applying the DE based algorithm, the total queue volume decreased to 30930 after 200 iterations as shown in Figure 5.2.

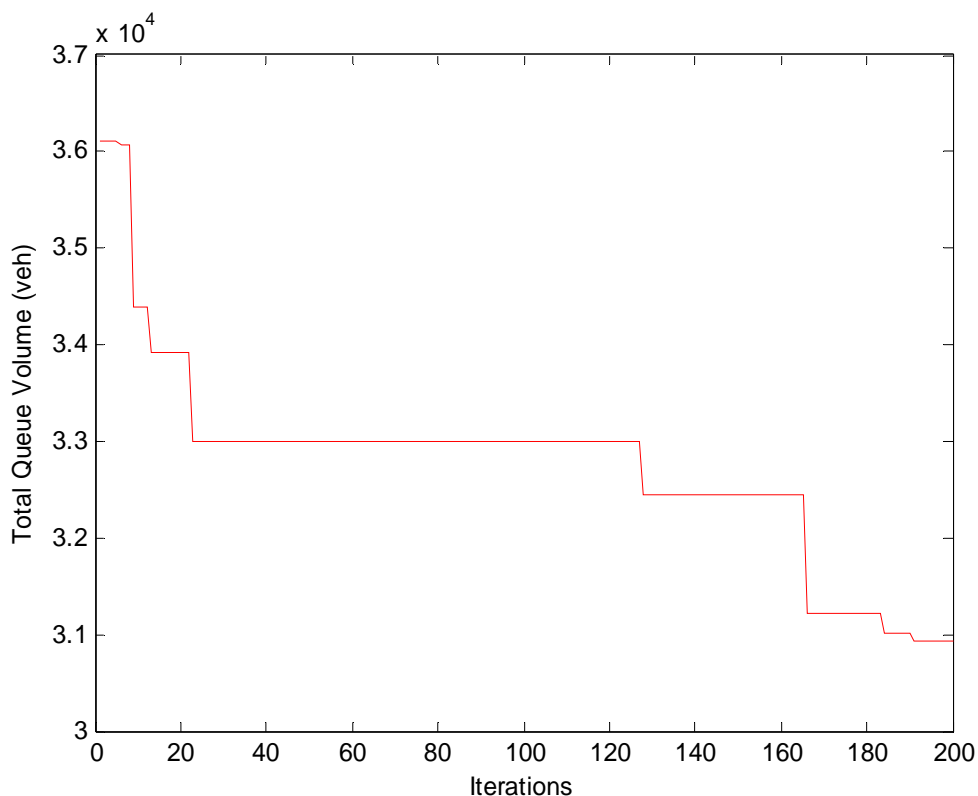


Figure 5.2 Convergence of Total Queue Volume by Using DE-Toll

The optimal toll level achieved is: $tol(1) = \$0.67$, $tol(2) = \$1.99$, $tol(3) = \$1.83$.

Under such toll level, the resulting queue distribution on each link is shown in Figure 5.3.

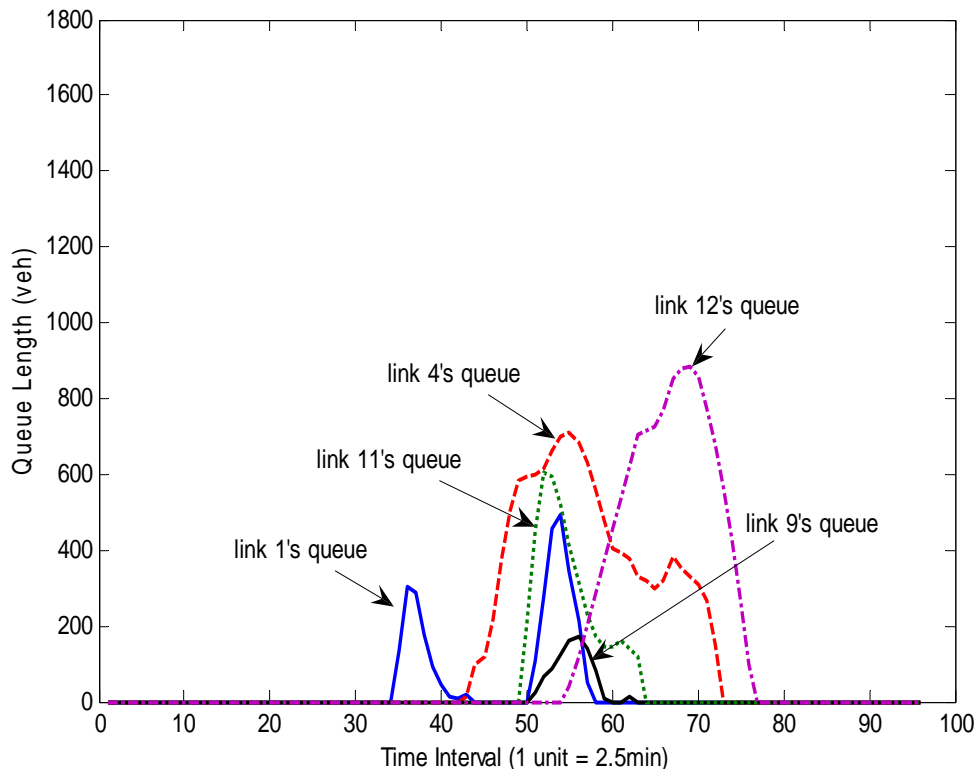


Figure 5.3 Queue Length under Tolling

Comparing Figure 5.1 with Figure 5.3, we observed that after tolling, the total queue length on the network (i.e. the queuing links 1, 4, 9, 11 and 12) reduced substantially, especially on link 1. This can only happen for two situations. One is that travelers change their departure time and the other is that travelers choose alternative links. In this example, the alternative link is link 7. From Figure 5.4 below, by comparing link inflow rate before and after tolling, one can see that there is not much traffic diversion from link 1 to link 7. However, after tolling, travelers do change the departure time to alleviate the heavy queuing on link 1.

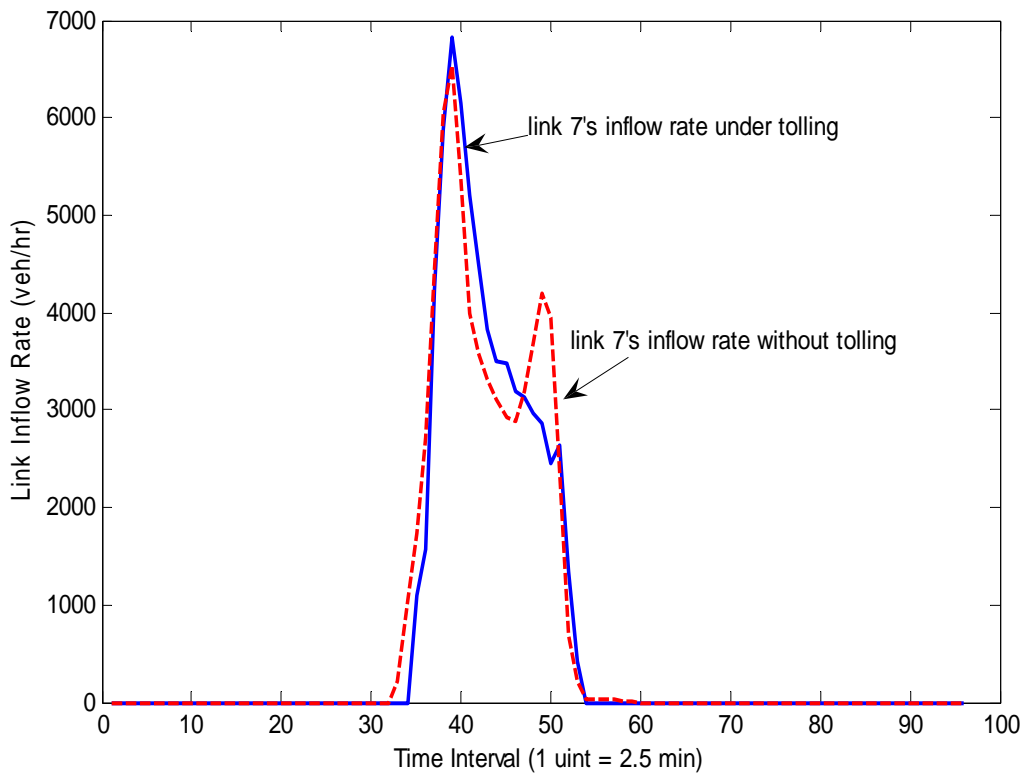
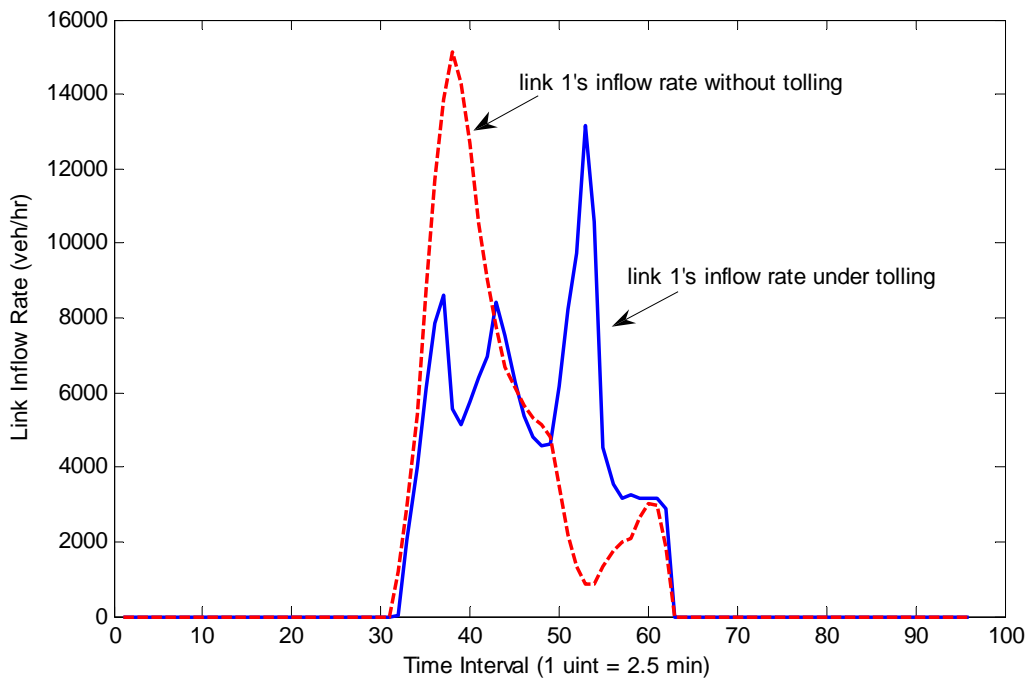


Figure 5.4 Link Inflow Rate before and after Tolling (Link 1 and 7)

5.3.2 Application of PS-Toll

Pattern search algorithm as introduced in Chapter 3 is applied to solve the dynamic road pricing problem. The network and parameters setting are the same as those in Section 5.3.1. Figure 5.5 and Figure 5.6 below show the convergence of total queue volume by using PS-Toll from different starting points. We find that the solution strongly depends on the starting point selection. If starting point is set as (0,0,0), the minimum of total queue volume is around 45026. If starting point is set as (1,1,1), the minimum of total queue volume is around 35621. Both results are greater than 30930 achieved by using DE-Toll in Section 5.3.1. The computation time is 10 hours by using PS-Toll from starting point (0,0,0) and 16 hours from starting point (1,1,1). However, the computation time is around 120 hours by using DE-Toll.

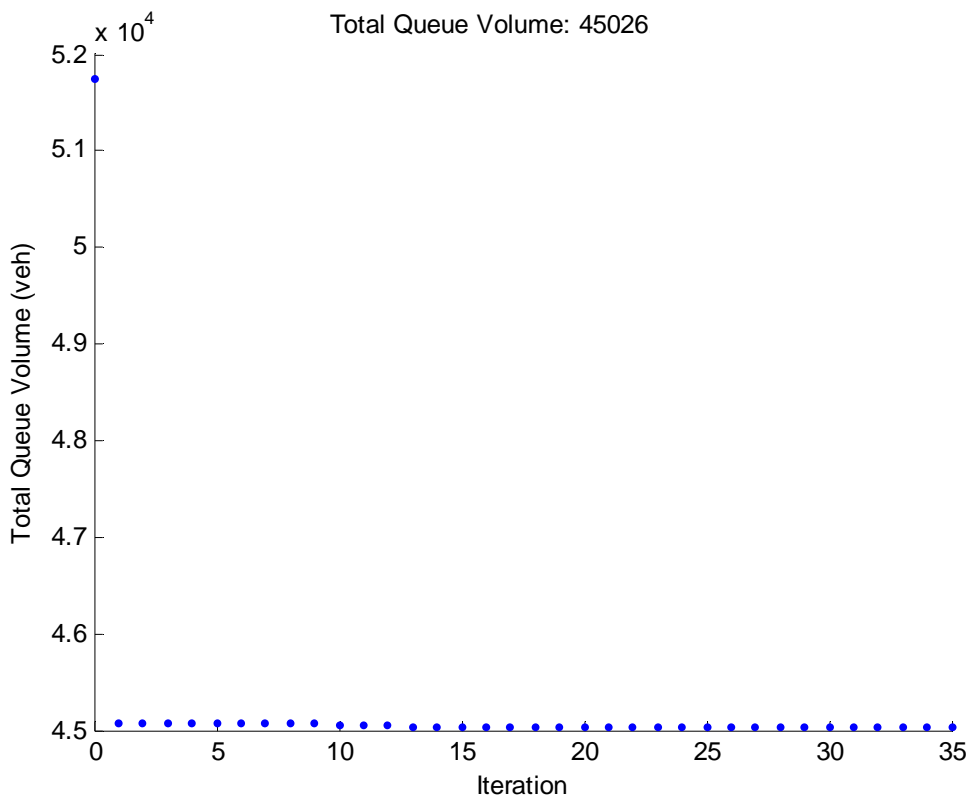


Figure 5.5 Convergence of Total Queue Volume by Using PS-Toll from Starting Point (0,0,0)

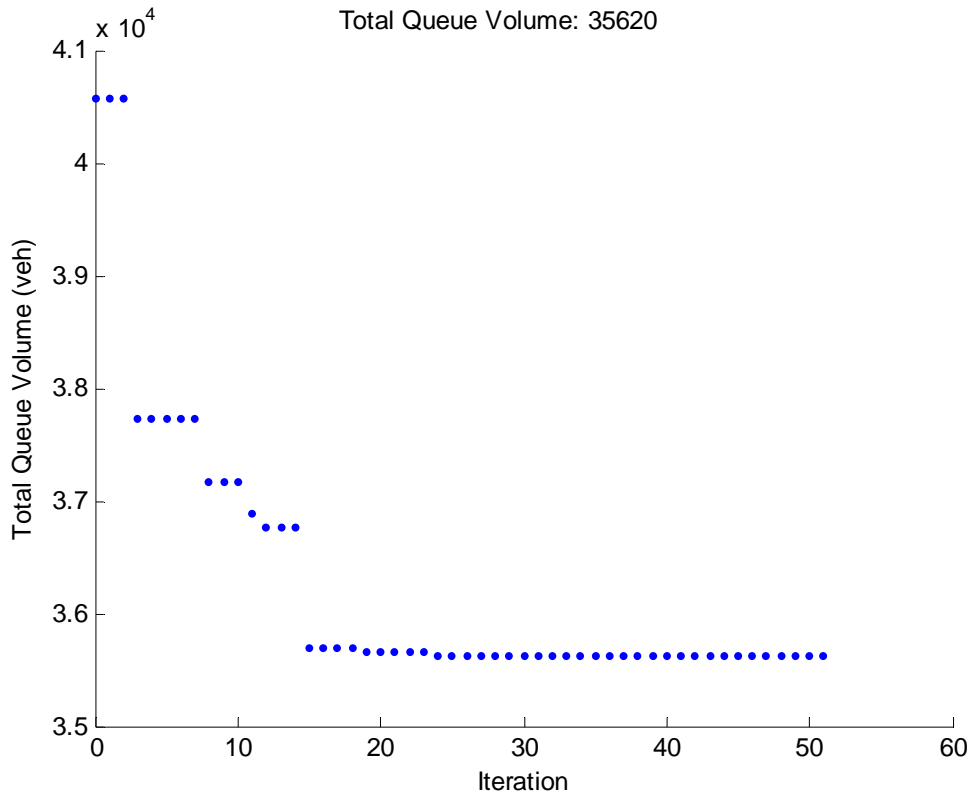


Figure 5.6 Convergence of Total Queue Volume by Using PS-Toll from Starting Point (1,1,1)

5.3.3 Sensitivity Analysis

In the numerical example, some of the pre-determined parameters are set as: the travel demand $F^{AC} = 10,000$ veh, $F^{BC} = 5,000$ veh, unit cost of travel time $\alpha = \$6.4/\text{hr}$, unit cost of scheduled delay time-early $\beta = \$3.9/\text{hr}$, unit cost of scheduled delay time-late $\gamma = \$15.21/\text{hr}$, and impact of road pricing on total travel cost $\mu = 0.1$. Recalling from Section 5.3.1 it has been found that by using DE-Toll method, total queue is minimized to 30930, tolling rates are set as $tol(1) = \$0.67, tol(2) = \$1.99, tol(3) = \$1.83$. Let each of the

pre-determined parameters vary 5%, the total queue and tolling rates are shown in Table 5.2. The percent of change of each variable are also listed. As the Table 5.2 shows, queue size is more sensitive to unit cost of travel time than to others parameters. Though the tolled link 1 does not line on any path between O-D pair BC, variation of demand of O-D pair BC can still result in changing of queuing and tolling rates.

Table 5.2. Sensitivity Analysis of Pre-determined Parameters

	Queue (veh)	Toll(1) \$	Toll(2) \$	Toll(3) \$
Original parameters	30930	0.67	1.99	1.83
α increased 5%	38143	0.98	1.92	2.00
Percent of change	23.32%	46.27%	-3.52%	9.29%
β increased 5%	32211	0.01	1.88	1.91
Percent of change	4.14%	-98.51%	-5.53%	4.37%
γ increased 5%	33221	0.38	1.96	1.91
Percent of change	7.41%	-43.28%	-1.51%	4.37%
F^{AC} increased 5%	36587	0.00	1.98	2.00
Percent of change	18.29%	-100.00%	-0.50%	9.29%
F^{BC} increased 5%	33462	0.00	1.81	1.93
Percent of change	8.19%	-100.00%	-9.05%	5.46%
μ increased 5%	33288	0.41	1.88	1.82
Percent of change	7.62%	-38.81%	-5.53%	-0.55%

5.4 SUMMARY

In this chapter, we present the dynamic road pricing model incorporating dynamic traffic assignment. The toll level is assumed to directly increase path

cost under tolling period. The tolling period is assumed to be fixed. The tolling amount and duration is displayed to the travelers before the start of traffic assignment. We also find that DE-Toll is superior to PS-Toll through a numerical example. Other pre-determined parameters include inputs for DTA model which were introduced in Chapter 4, such as: the travel demand, unit cost of travel time, unit cost of scheduled delay time-early, and unit cost of scheduled delay time-late. Sensitivity analysis of these parameters is carried. More complicated applications will be discussed in Chapter 7. In this chapter, we have only one design objective which is to minimize the total queue volume on the network during the simulation time period. In fact, policy makers are always faced with different design objectives at the same time. In the next Chapter, the multi-objective road pricing design problem will be discussed.

CHAPTER 6 MULTI-OBJECTIVE PROBLEM FOR DYNAMIC ROAD PRICING

6.1 INTRODUCTION

In the last two chapters, the dynamic road pricing problems have been established and solved for a single objective MPEC (Mathematical Program with Equilibrium Constraints) problem. Traffic manager can set a proper pricing strategy to alleviate the total queue on the network, or to minimize total traffic delay from a systematic viewpoint. In a real traffic network design, multi-objective problems are not uncommon. This requires the traffic manager to choose the best compromise between those objectives which may not be compatible. This Chapter establishes a multi-objective dynamic road pricing problem. After reviewing some of the evolutionary-based approaches, one of them named “Non-dominated Sorting Genetic Algorithm” (NSGA-II) will be used to tackle the problem.

6.2 PROBLEM SETTING

The standard formulation of a multi-objective problem is to find a vector $\mathbf{x}^* = [x_1^*, x_2^*, \dots, x_n^*]^T$ to optimize a vector function

$$\mathbf{f}(\mathbf{x}) = [f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_k(\mathbf{x})]^T \quad (6.1)$$

where $\mathbf{x} = [x_1, x_2, \dots, x_n]^T$ is the vector of decision variables.

The solutions would satisfy the following m inequality and p equality constraints:

$$g_i(\mathbf{x}) \geq 0 \quad i = 1, 2, \dots, m \quad (6.2)$$

$$h_i(\mathbf{x}) = 0 \quad i = 1, 2, \dots, p \quad (6.3)$$

Pareto Front

Pareto optimum says that \mathbf{x}^* is the Pareto optimal if there exists no feasible vector \mathbf{x} which would decrease some criteria without causing a simultaneous increase in at least one other criterion. Usually, Pareto optimum gives not a single solution, but a set of solutions called the non-dominated solutions. For instance, consider the case of a bi-objective problem to minimize a vector function $\mathbf{f}(\mathbf{x}) = [f_1(\mathbf{x}), f_2(\mathbf{x})]^T$. In Figure 7.1, Ω is the feasible region and the bold line is the Pareto optimum called the Pareto front.

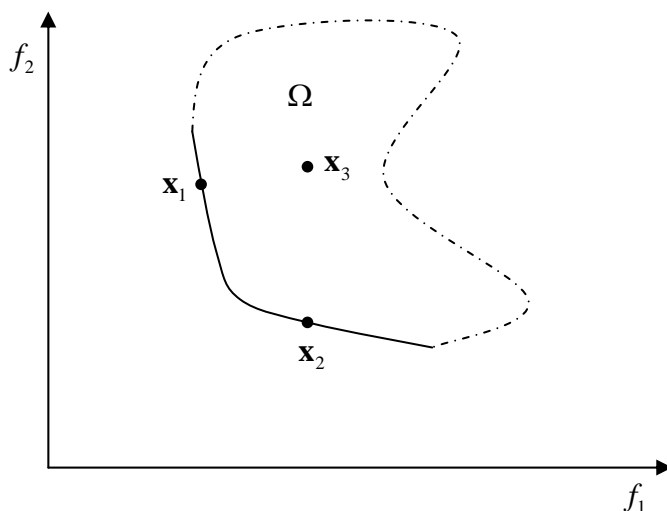


Figure 6.1 An Example of Pareto Optimum

In this example, $\mathbf{x}_1, \mathbf{x}_2$ are the points in the Pareto front. They are not dominated by each other. However, \mathbf{x}_3 is dominated by both of them as the objective function

f_1 and f_2 can achieve smaller value at point $\mathbf{x}_1, \mathbf{x}_2$ than at point \mathbf{x}_3 . Thus, \mathbf{x}_3 is not a Pareto optimal solution.

6.3 EVOLUTIONARY-BASED APPROACHES FOR MULTI-OBJECTIVE PROBLEM

6.3.1 Approaches That Use Aggregating Functions

As general genetic algorithm (GA) needs to evaluate the fitness of only one single objective function, a simple attempt to tackle multi-objective optimization problem is to reduce the multi-objective functions to a single objective function which is usually termed as aggregating function.

Weighted Sum Approach

This method consists of adding all the objective functions together using different weighting coefficients for each one of them. This means that a multi-objective problem is transformed into an optimization problem with single objective function as shown below:

$$\min \sum_{i=1}^k w_i f_i(\mathbf{x}) \quad (6.4)$$

where $w_i \geq 0$ are the weighting coefficients representing the relative importance of the objectives. It is usually assumed that

$$\sum_{i=1}^k w_i = 1 \quad (6.5)$$

As the weighting coefficients change, the results obtained by solving the single objective function (6.4) can vary significantly. It is necessary to solve the same problem for different values of w_i . After that, one has to choose the appropriate solution based on intuition. Friesz et al. (1993) adopted the weighted sum approach with the simulated annealing method to solve the multi-objective network design problem. The main strength of this method is its computational efficiency. However, it faces several weaknesses. The first is that it is difficult to determine the appropriate weighting coefficient if one does not have enough information about the problem. Secondly, it may not generate true Pareto optimal solutions if the feasible space and objective function are not convex (Chankong and Haimes, 1983).

Goal Programming Method

The development of a goal programming method for a linear model can be credited to Charnes and Cooper (1961) and Ijiri (1965). In this method, the decision maker has to specify a target or goal for each objective. These values are incorporated into the problem as conditional conditions. The objective function is then singled out as to minimize the absolute deviations from the targets to the objectives. The simplest case of such an objective function can be shown as:

$$\min \sum_{i=1}^k |f_i(\mathbf{x}) - T_i| \quad (6.6)$$

where T_i denotes the target or goal for the objective function $f_i(\mathbf{x})$. A more general formulation of goal programming method is to add the weighting sum of p th power of the deviation $|f_i(\mathbf{x}) - T_i|$ together to a single objective function (Ignizio, 1976). The goals T_i are preset by the decision maker. If the goals are chosen in the feasible domain, the method will yield a dominated solution (Duchstein, 1984).

The ε -Constraint Method

This method (Haimes, et al., 1971) is to minimize one objective function (the most preferred or primary) and to consider the other objective functions as hard constraints bound by some allowable levels ε_i . The levels ε_i are then varied to generate Pareto optimal solutions. The method can be stated as follows:

- (1) Find the minimum of *rth* objective function, i.e., find \mathbf{x}^* such that

$$f_r(\mathbf{x}^*) = \min_{\mathbf{x} \in \Omega} f_r(\mathbf{x}) \quad (6.7)$$

subject to additional hard constraints:

$$f_i(\mathbf{x}) \leq \varepsilon_i \quad i = 1, 2, \dots, k; i \neq r \quad (6.8)$$

where ε_i are the values of the objective functions which one wishes not to exceed.

- (2) Repeat step (1) for different values of ε_i until a set of acceptable solutions is compiled.

This method allows the designer to determine the complete Pareto solutions, but only if all possible values ε_i are used.

6.3.2 Other Not Pareto-based Approaches

To overcome the difficulty involved in the aggregating approaches where multiple objectives are transformed into a single objective function, much work has been

devoted to develop alternatives based on the population policies. Some of the most popular approaches are presented below.

VEGA

Schaffer (1985) used an extension of the simple Genetic Algorithm (GA) which he called the Vector Evaluated Genetic Algorithm (VEGA) to tackle the multi-objective optimization problem. This method differs with GA on the selection procedure. In each generation, a number of sub-populations will be generated by performing proportional selection according to each objective function in turn. For example, a multi-objective problem has k objective functions. Total population size is N_2 . k sub-populations will be generated with the size N_2/k each. These sub-populations will be shuffled to the generation of a new population with size N_2 achieved by using the general GA mutation and crossover operators. This process is illustrated in Figure 6.2.

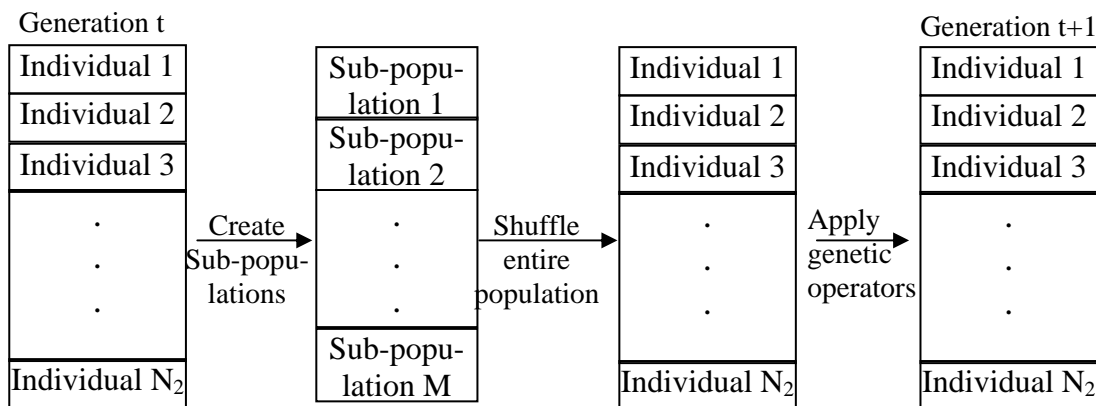


Figure 6.2 Schematic of VEGA Selection

The main strength of this method is its simplicity. However, it has its own weaknesses. Schaffer used proportional fitness assignment and the fitness components were proportional to each objective. Therefore, the resulting expected

fitness corresponded to a linear combination of the objectives. When the feasible region is non-convex, this method produces no Pareto optimal solutions.

Lexicographic Ordering Method

In this method (Fourman, 1985), designers are required to rank the objectives in the order of importance. Then the final solution is obtained by solving a series of optimization problem with a single objective function. Let the subscripts of the objectives indicate not only the objective function number, but also the priority of each objective. Thus, $f_1(\mathbf{x}), \dots, f_k(\mathbf{x})$ denote the objective functions in the order of most importance to the least importance. Then the first problem is formulated as

$$\min f_1(\mathbf{x}) \quad (6.9)$$

subject to

$$g_j(\mathbf{x}) \leq 0 \quad j = 1, 2, \dots, m \quad (6.10)$$

and its solution \mathbf{x}_1^* is obtained and denote $f_1^* = f_1(\mathbf{x}_1^*)$. Then the second problem is formulated as

$$\min f_2(\mathbf{x}) \quad (6.11)$$

subject to

$$g_j(\mathbf{x}) \leq 0 \quad j = 1, 2, \dots, m \quad (6.12)$$

$$f_1(\mathbf{x}) = f_1^* \quad (6.13)$$

and the solution of this problem is \mathbf{x}_2^* . Note that $f_1(\mathbf{x}_2^*) = f_1^*$. Denote $f_2^* = f_2(\mathbf{x}_2^*)$. This procedure is repeated until all the k objectives have been considered. The i th problem is formulated as

$$\min f_i(\mathbf{x}) \quad (6.14)$$

Subject to

$$g_j(\mathbf{x}) \leq 0 \quad j = 1, 2, \dots, m \quad (6.15)$$

$$f_l(\mathbf{x}) = f_l^* \quad l = 1, 2, \dots, i-1 \quad (6.16)$$

This approach is relatively simple as compared to VEGA. However, it requires the designer to make a preference of all objective functions. In the presence of many objective functions, this approach will favour to some extent certain objectives resulting in the inability to generate the whole range of Pareto optimal solutions.

6.3.3 Pareto-based Approaches

Goldberg (1989) was the first to suggest the use of the non-dominated ranking selection to move a population towards a Pareto front in a multi-objective optimization problem. The idea is to set a sub-population which is non-dominated from the remaining population with the highest rank. Then another set of a sub-population from the remaining population is determined and is assigned the next highest rank. This process continues until all the population is ranked. Goldberg also suggested the use of some kind of niching technique to keep the GA from converging to a single point on the front. Some recent algorithms include Multiple Objective Genetic Algorithm (MOGA) proposed by Fonseca and Fleming (1993), Non-Dominated Sorting Genetic Algorithm (NSGA) by Srinivas and Deb (1993), and Niche Pareto Genetic Algorithm (NPGA) by Horn and Nafpliotis (1993). As

cited by Deb (1998), the main weakness of MOGA is that it prohibits two different vectors with the same objective function values to coexist simultaneously in the population. This is apparently undesirable. However, NSGA allows multiple equivalent solutions to coexist despite that it is more inefficient than MOGA. NPGA does not apply Pareto selection to the entire population, but only to a segment of it at each run. Its main strength is its computational efficiency.

The algorithms mentioned above represent the early development of Pareto-based approach to multi-objective optimization problem. However, Laumanns et al. (2002) showed that none of these algorithms can converge to a true Pareto front because of the lack of elitism strategy. Thus, various methods in designing the elitism strategy have been proposed. Some of them include Strength Pareto Evolutionary Algorithm (SPEA) by Zitzler and Thiele (1999), Pareto Archived Evolutionary Strategy (PAES) by Knowles and Corne (2000) and Non-Dominated Sorting Genetic Algorithm-II (Deb et al., 2002). Deb et al. (2002) also found that NSGA-II outperforms SPEA and PAES in terms of finding a diverse set of solutions and in converging near to the true Pareto front. In view of this, NSGA-II is applied in this research study to tackle the multi-objective problem for the optimal toll setting. The next section introduces the details of NSGA-II.

6.3.4 NSGA-II Algorithm

The step-by-step procedure is shown below:

Step 1. Combine parent and offspring population: $R_t = P_t \cup Q_t$;

Step 2. Population R_t is sorted by using the fast-non-dominated sorting (to be described in Section 6.3.5) and to generate the front F_i ($i = 1, 2, \dots$);

Step 3. Set new population $P_{t+1} = \emptyset$ and $i=1$. Until $|P_{t+1}| + |F_i| \leq N_2$, perform

$$P_{t+1} = P_{t+1} \cup F_i \text{ and } i = i + 1, \text{ where } N \text{ is the population number;}$$

Step 4. Choose the first $(N - |P_{t+1}|)$ elements from F_i by using crowded-comparison operator (\prec_n) (to be described in Section 6.3.6) to fill P_{t+1} ;

Step 5. Generate offspring population Q_{t+1} from P_{t+1} by using general GA selection, mutation and crossover, in which the selection criteria are based on crowded-comparison operator (\prec_n).

Initially, a random parent population P_0 is created. The population is sorted based on the non-domination as described in subsection 6.3.5. Each solution is assigned a fitness equal to their non-domination rank. Thus, minimization of fitness is assumed. The usual binary tournament selection, recombination, and mutation operators are used to create a child population Q_0 of size N_2 . For $t \geq 1$, as shown in Figure 6.3, the algorithm which firstly generates a population of size $2 * N_2$ is formed as $R_t = P_t \cup Q_t$. Q_t from P_t is generated by using the general GA algorithm with the selection criteria based on the crowded-comparison operator (\prec_n). R_t is then sorted according to the non-domination sorting which is introduced in Section 6.3.5. The purpose of the non-domination sorting is to generate fronts F_i with different ranks. In step 3, the best non-dominated set F_i is included in the new P_{t+1} according to their ranks. The size of F_1 is usually smaller than the population size N_2 . Suppose F_{j_0} is the last front whose member can be entirely included in the new P_{t+1} and the current size of P_{t+1} is less than N_2 , step 4 shows how to find members from F_{j_0+1} to complete the generation of P_{t+1} . This process is introduced in Section 6.3.6. In step 5, new offspring population

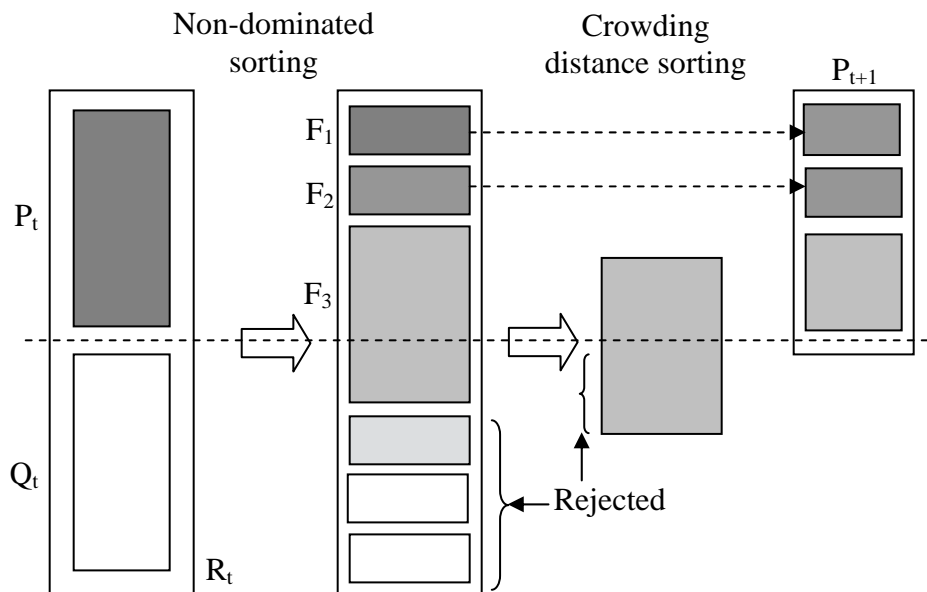


Figure 6.3 Procedure of NSGA-II

6.3.5 A Fast Non-dominated Sorting

Non-dominated sorting of a population is a key elitist in NSGA-II. Figure 6.4 shows an example of a non-dominated sorting with 2 objectives. Three fronts F_1 , F_2 , and F_3 are demonstrated to be given with different ranks. The rank of F_1 is set to 1, rank of F_2 is 2, and rank of F_3 is 3. It is obvious that the smaller the rank of a front, the better the member of the front satisfying the multi-objective optimization. Members of each front are non-dominated mutually as explained in Figure 6.1. Deb et al. (2002) gave a fast non-dominated sorting method which is shown in Figure 6.5.

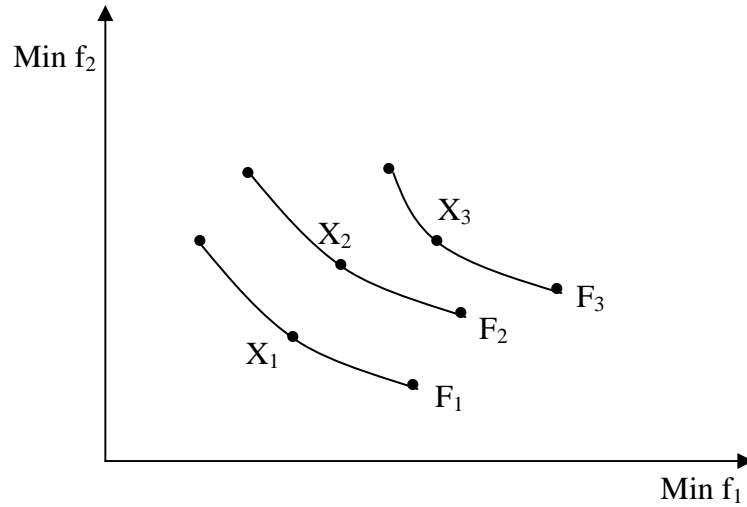


Figure 6.4 An Example of Non-dominated Sorting

A Fast Non-dominated Sorting Method

For each $p \in P$

Set $S_p = \phi$, $n_p = 0$

For each $q \in P$

If $(p \prec q)$ then

$$S_p = S_p \cup \{q\}$$

Else if $(q \prec p)$ then

$$n_p = n_p + 1$$

If $n_p = 0$ then

$$Rank(p) = 1$$

$$F_1 = F_1 \cup \{p\}$$

Set counter $i = 1$

While $F_i \neq \phi$

$$Q = \phi$$

For each $p \in F_i$

For each $q \in S_p$

$$n_q = n_q - 1$$

If $n_q = 0$ then

$$Rank(q) = i + 1$$

$$Q = Q \cup \{q\}$$

$$i = i + 1$$

$$F_i = Q$$

Figure 6.5 A Fast Non-dominated Sorting Method

Parts of the notations in Figure 6.5 are explained as follows. For any two members, p and q from the population P , let $p \prec q$ if q dominates p . Let n_p denotes the number of solutions which dominate the solution p . Let S_p denotes the set of solutions that p dominates.

6.3.6 Crowded-comparison Operator (\prec_n)

The crowded-comparison operator represents the selection criteria embedded in the algorithm. In principle, wide spread solutions have a higher chance to be selected. So, there is a need to estimate the density of solutions surrounding a particular solution. One can use the average distance of two points on either side of this point along with each of the objectives. This quantity represents an estimate of the perimeter of the cuboid formed by using the nearest neighbors as the vertices (called ‘crowding distance’). Figure 6.6 shows an example with the 2 objectives.

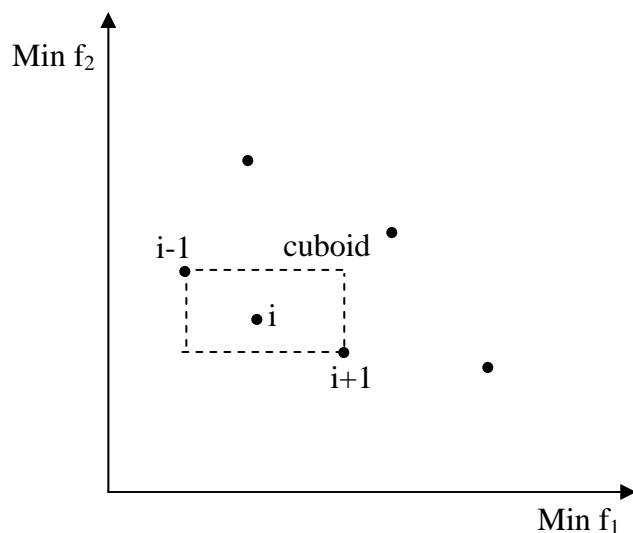


Figure 6.6 An Example of Crowding-distance Calculation

The following gives the step-by-step procedure to calculate the crowding distance.

Step 1 set the number of solution in F as $n = |F|$ and for each member $p \in F$, set the initial distance $d(p) = 0$;

Step 2 For each objective function m , sort the solution set in descending order of f_m ;

Step 3 For each objective function m , the boundary solutions (solutions with the smallest and largest functions) are assigned an infinite distance value. $d(I_1^m) = d(I_n^m) = \infty$, for other solutions $j = 2, \dots, n-1$,

$$d(p) = d(p) + (f_m^{(I_{j+1}^m)} - f_m^{(I_{j-1}^m)}) / (f_m^{\max} - f_m^{\min})$$

$d(p)$ denote the crowding distance; Index I_j^m denotes the solution index of the j th member in the sorted list based on the objective function m ; f_m^{\max} and f_m^{\min} are the maximum and minimum values of the objective function m .

After defining the crowding distance for each member in the population, one can define the crowded-comparison operator. Assume each member has two attributes:

- (a) non-domination rank
- (b) crowding distance

Crowded-comparison operator is defined as:

For any two members p and q ,

$$p \prec_n q \text{ if } \text{rank}(p) < \text{rank}(q) \text{ or } ((\text{rank}(p) = \text{rank}(q)) \text{ and } (d(p) > d(q)))$$

That means the solution with lower a rank is more preferable. If two solutions have the same rank, the solution located in a less crowded region is preferred.

6.4 MULTI-OBJECTIVE OPTIMAL TOLL DESIGN AND EXAMPLE

In this section, a bi-objective road pricing model using the NSGA-II algorithm is illustrated.

The network is the same with the example illustrated in the last chapter. Tolling link is still confined to link 1.

One objective is to have minimum queue. The other is maximum revenue generated by tolling. More specifically, $obj_1 = \sum_k \sum_a q_a(k)$, and

$$obj_2 = \sum_k \sum_a q_a(k) * tol(k)$$

Thus, the bi-objective road pricing can be expressed as:

$$\min \begin{pmatrix} obj_1 \\ -obj_2 \end{pmatrix} \quad (6.17)$$

\mathbf{f} solves VI problem

$$\sum_k \sum_{rs} \sum_p c_p^{rs}(k, \mathbf{f}^*) \cdot [f_p^{rs}(k) - f_p^{rs*}(k)] \geq 0 \quad (6.18)$$

Satisfying constraints:

$$\Omega = \left\{ \mathbf{f} \geq 0 : \sum_p \sum_k \delta \cdot f_p^{rs}(k) = F^{rs} \right\} \quad (6.19)$$

Here, $c_p^{rs}(k, \mathbf{f})$ in the VI problem (6.18) should be replaced by $Tc_p^{rs}(k, \mathbf{f})$ as shown in the last Chapter. Λ is a set of constraints for tolling variables.

Before one can apply the NSGA-II algorithm, one should tackle this problem by brute-force. Continuing with the last Chapter, tolling variables $tol(j)$ ($j = 1, 2, 3$) are applied on the same three different time windows accordingly. The 3 different time windows are intervals from [36, 41], [42, 47] and [48, 53]. One can note that the values of tolling variables are set to between 0 and 2. Suppose every tolling change is 0.1, there are $21 \times 21 \times 21 = 9261$ instances in total. For these 9261 pricing scenarios, the queue size and the revenue is illustrated in Figure 6.7. Note that the vertical axis is plotted as negative revenue for a simplistic demonstration.

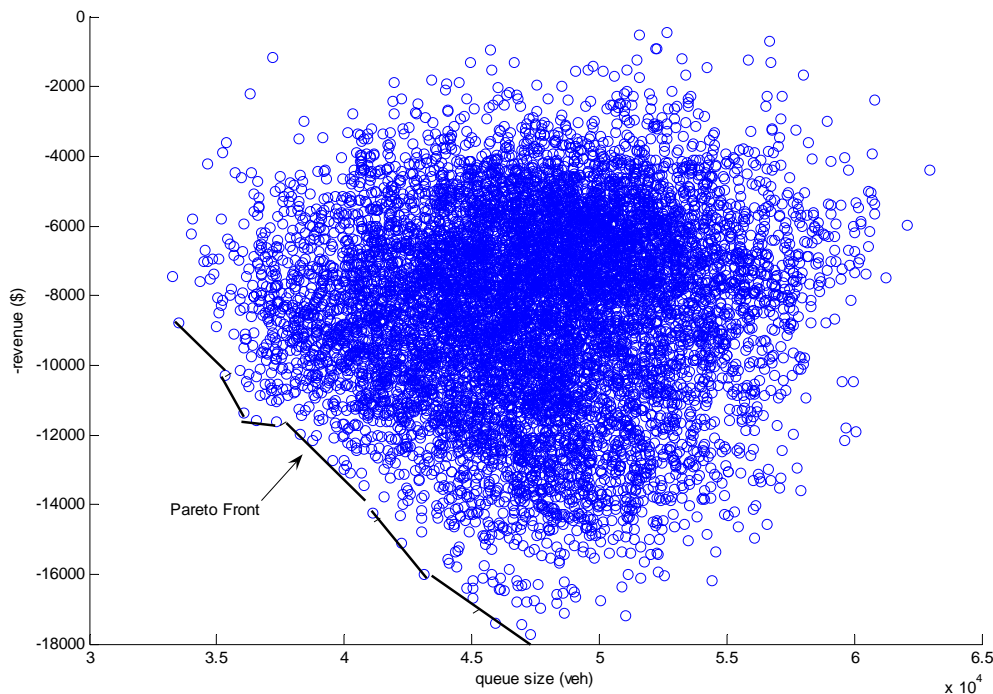


Figure 6.7 Discrete Results of Bi-objective Road Pricing

In the left corner of Figure 6.7, the Pareto Front is shown in solid line. In Figure 6.8, the Pareto Front is generated by the NSGA-II algorithm. Compared with Figure 6.7, one can see that the results by NSGA-II are satisfactory. The computation times for the brute-force method and NSGA-II algorithm are 18 hours and 8 hours separately. This example also shows the importance of multi-objective consideration in practice. In Figure 6.7, different queue size can achieve the same toll revenue.

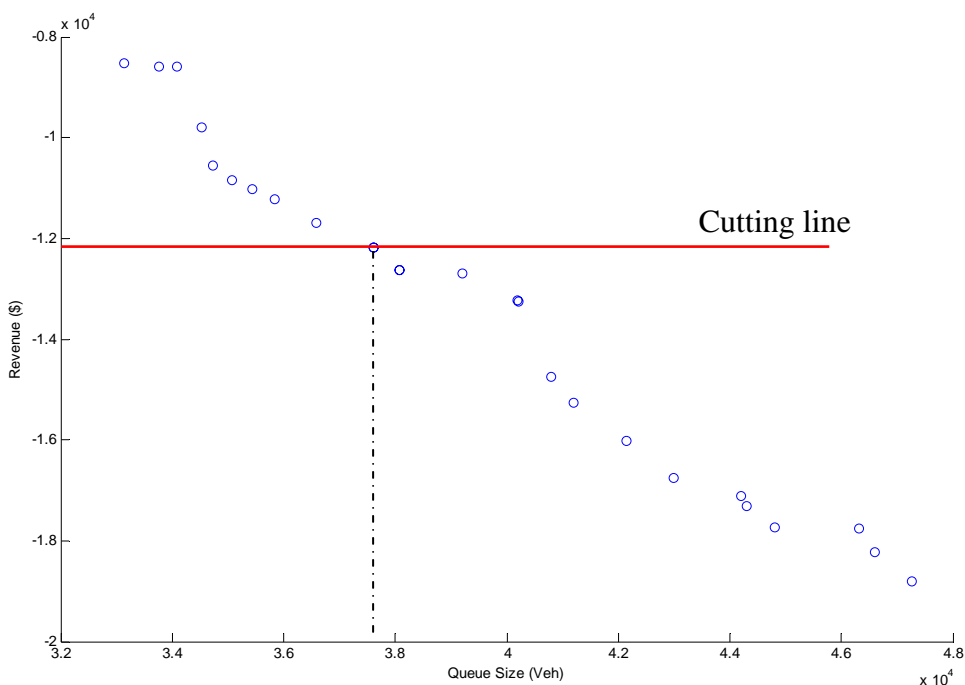


Figure 6.8 Pareto Front Generated by NSGA-II

6.5 SUMMARY

In this chapter, we have discussed a more practical dynamic road pricing design problem. Facing with multiple objectives is not uncommon in reality. Some of the

objectives may be in conflict with each other. How to measure the trade-off between them is of practical concern. After reviewing the evolutionary algorithms to deal with the multi-objective optimization problem, we apply NSGA-II in our grid network example. One of the objectives is to minimize the total queue volume on the network, while the other is to maximize the toll revenue. Basically, these two objectives do not accommodate each other. By applying NSGA-II, a quantitative relationship between these two objectives can support traffic manager for decision making. A traffic manager can choose his cutting line and then derived the appropriate tolling level. For example, in Figure 6.8, the cutting line shows the required minimal toll revenue is \$12,175, and the resulting minimal queue size is 37,614 vehicles. The corresponding tolling level is : $tol(1) = \$1.37$; $tol(2) = \$1.65$; $tol(3) = \$2.00$.

CHAPTER 7 CONCLUSIONS AND FUTURE RESEARCH

7.1 PRACTICAL ROAD PRICING

Road pricing has been in Singapore for more than 30 years from the early Area License Scheme (ALS) in 1975 to the recent Electronic Road Pricing (ERP) system in 1998. The current ERP rates depend on location, vehicle type and time of the day. Every three months the tolling level are reviewed and adjusted to maintain a targeted traffic speed on expressways and roads inside the Restricted Zone (RZ) or Central Business District (CBD) at 45-65 km/hr and 20-30 km/hr, respectively. This empirical pricing scheme is easy to implement. However, it has its own limitations. First, it does not take into account traffic evolution on un-tolled links. Initially, the traffic manager makes a decision to toll the congested roads to persuade travelers to use other links. When too many travelers have diverted to the un-tolled links, these previously un-congested links may become congested after implementing the pricing schemes. Thus, this pricing scheme only achieves local optimum and not globally. This case has been verified by feedback gathered from regular drivers such as taxi drivers who contended that during tolling period, the traffic inside the CBD is rather smooth flowing while traffic outside CBD is rather congested. Secondly, the tolling levels are adjusted quarterly. It is too long to reflect the day-to-day traffic evolution. All these reasons give rise to considering dynamic road pricing on the entire network.

7.2 THEORETICAL APPROACH TO SOLVING DYNAMIC ROAD PRICING INCORPORATING DTA

The static optimal road pricing problem was first formulated as a Mathematical Problem with Equilibrium Constraints (MPEC) in this research study. Static traffic assignment describing traffic distribution on the network is taken into account. The behaviour of travelers is assumed to follow Wardrop's user equilibrium. To solve this problem, three different algorithms were tested: Differential Evolutionary (DE), Pattern Search (PS), and Non-linear Problem with Equilibrium Constraints (NLPEC) solver. For this well-defined problem, NLPEC solver showed the least computation time for a given hypothetical nine-node network.

To design a dynamic road pricing scheme, dynamic traffic assignment model is used to simulate the evolution of traffic. As there is no general formulation of Dynamic Traffic Assignment (DTA) model and this study focuses on dynamic road pricing on a network; a DTA model proposed by Huang and Lam (2002) on a network with queue was adopted as a base model. A new procedure to compute some inter-variables is proposed to increase the accuracy of the model. The behaviour of travelers follows the dynamic extension of Wardrop's user equilibrium, which is based on predictive travel time. FIFO conditions in dynamic situations are satisfied. Travellers can choose different departure times under different road pricing schemes. The existence of the model solution is guaranteed. A new algorithm based on a projection method was proposed. Comparing this new algorithm with the original route/time swapping algorithm through a numerical example, this algorithm presents a smoother traffic flow. In this DTA model, traffic distribution, queue size and departure time are all subject to changes in the tolling level. So, it can be treated as a good description of network evolution for the study of dynamic

road pricing.

For the dynamic road pricing model incorporating DTA model, it can still be modeled as an MPEC problem, albeit a not well-behaved one. In this MPEC setting, the DTA model was set as an equilibrium constraint. The single objective of this MPEC problem was to achieve a minimal queuing size on the network. The impact of tolling level was assumed to directly increase the travel cost formulation. Among the three MPEC algorithms, it was demonstrated that NLPEC could not be applied and DE was found to be more efficient in its computation in solving the dynamic road pricing model incorporating DTA. To evaluate the objective of the MPEC model in the process of applying DE algorithm, the resulting parametric DTA problem was solved using this algorithm based on the projection method proposed in this thesis. In this research, the amount of toll depends on the length of stay on the tolled link.

Lastly, multi-objective dynamic road pricing model applying the Non-dominated Sorting Genetic Algorithm (NSGA-II) was successfully illustrated to show how the problem would behave when confronted by the decision maker in reality. In the numerical example, two objectives were considered. One was to minimize the queue size on the network, while the other was to maximize the toll revenue. Besides, a quantitative tradeoff between these two objectives can help traffic manager to set the optimum toll charge to better manage the congestion level. In essence, the above conclusions were based on the assumptions of the DTA model involved. These assumptions include simplified link travel time function, idealistic travel behaviour modeling, pre-determined input parameters and illustration on a hypothetical nine-node network.

7.3 SOME THOUGHTFUL RECOLLECTIONS

Firstly, in order to achieve a more accurate representation of traffic distribution in a short time period, DTA models should be adopted. It allows travelers to adjust their departure time and route choice. Road pricing schemes should also be dynamic in nature in order to achieve the best design objectives.

Secondly, when decision maker is confronted with a design problem with multiple objectives, road pricing schemes must be carefully designed. One can recall from the example given in Chapter 6 (Figure 6.7) in which different queuing phenomena can result in the same revenue. Thus, the road pricing schemes optimizing queues should be used by the decision maker.

7.4 LIMITATIONS AND RECOMMENDATIONS

The dynamic road pricing model proposed in this research study largely depends on two crucial factors. One is the DTA model and the other is the tolling impact on travelers' route choice behavior.

There are a few limitations of this study. The DTA model developed in this research study is based on a number of assumptions which include a single user class, a single mode and the dynamic Wardrop's user equilibrium concept. Future research should explore the possibility of relaxing the DTA assumptions and try other tolling impact models. For static road pricing, Meng and Yang (2002) has considered road pricing with multiple user classes. Other than the concept of deterministic user equilibrium, stochastic user equilibrium can better represent travelers' route choice behaviour. One example of this DTA model is Han (2003); applying this DTA as an equilibrium constraint of MPEC for

dynamic road pricing is another direction of modeling improvement.

In DTA model, point queuing is adopted instead of physical queuing. Models employing point queuing cannot describe queue propagation phenomena accurately. To incorporate physical queuing is a challenge in future work. In Equation (4.8), travel time is assumed linear with queue length and maximum link exit flow rate s_a is assumed constant. These assumptions may not be true in reality. These assumptions can be relaxed. If s_a is a continuous function of link inflow rate $u_a(k)$, the variational inequality formulation of DTA problem is still valid. The reasoning is same with that in Section 4.3. However, the method of calculating link exit flow $v_b^{rsp}(k)$ proposed in Section 4.4 may not be valid if s_a is not constant. Other methods to estimate $v_b^{rsp}(k)$ have to be considered. A possible maximum exit flow rate function can be like below:

$$s_a(u_a(k)) = \begin{cases} e_a & \text{if } u_a(k) < c_0 \\ e_a - (u_a(k) - c_0)^{1/3} & \text{if } c_0 \leq u_a(k) < c_1 \\ e_a - (c_1 - c_0)^{1/3} & \text{if } u_a(k) \geq c_1 \end{cases} \quad (7.1)$$

c_0 , c_1 and e_a are constant value. This functions shows that maximum exit flow rate is fixed when inflow rate is under some critical value c_0 . If inflow rate is bigger than some critical value c_0 , maximum exit flow rate may decline with the increase of the inflow rate. If inflow rate is bigger than some critical value c_1 , maximum exit flow rate becomes constant again.

The linearity between the travel time and queue size can also be relaxed. For example, the travel time can be expressed as follows:

$$t_a(k) = t_a^0 + \frac{q_a^2(k)}{\delta \cdot s_a(u_a(k))} \quad (7.2)$$

Applying Equations (7.1), (7.2), (4.11) and (4.12), the following expression of queue can be derived:

$$\frac{q_a^2(k)}{\delta s_a(u_a(k))} - \frac{q_a^2(k-1)}{\delta s_a(u_a(k-1))} = \frac{u_a(k)}{v_a(k+t_a(k))} \quad (7.3)$$

In the dynamic road pricing problem formulated in this study, travelers' route choice was based on the path cost which was a function of travel time and delayed time as suggested in Equation (4.18). The parameters introduced to express this function are: α , the unit cost of travel time; β , the unit cost of schedule delay time-early; γ , the unit cost of schedule delay time-late. Tolling effect is directed to increase the path cost by introducing the parameter μ as suggested in Equations (5.1) and (5.2). In reality, two things have to be considered thoroughly. One is that travelers cannot have the same values for these four parameters. The other is that it is difficult to express the path cost as a function of travel time, delayed time and tolling level determinedly. In other words, for a small change in tolling level, travel time or delayed time, travelers may think their path costs remain unchanged. That is, the path cost generated in this study is overly sensitive to the tolling level. A promising treatment of the tolling effect, travel time and delayed time is to introduce fuzzy logic in the path cost formulation. This can lead to a more representative model, thus one of the directions for future work.

In our model, total traffic demand F^{rs} is assumed fixed. However, this may not be true when toll is applied. In addition, travellers may incur departure time changing cost. Both have not been considered in our model.

The network used in this study is a nine-node network with 8 paths. In

Chapter 5, the total time interval is set as 96. So, for the variational inequality formulation of DTA model, the number of variables is $8 \times 96 = 768$. In the projection-based algorithm proposed in Chapter 4, nonlinear optimization solver CONOPT (Drud, 1992) was used to solve the quadratic programming problem (Equation 4.25). Currently, CONOPT can solve large scale problem with variables more than 20,000 (Drud, 1992). Thus, if total time interval is not changed, projection-based algorithm can solve DTA model over a network with 200 paths. This study uses the same nine-node network with Huang and Lam (2002) to compare the different algorithms. Application on larger network is recommended for future study.

To make the dynamic road pricing applicable, one has to calibrate and validate the assumptions and parameters used in the model. Maximum exit flow rate function (Equation 7.1) may be generated from traffic historic data. The unit cost of travel time α ; the unit cost of schedule delay time-early β ; the unit cost of schedule delay time-late γ , and the tolling effect μ may be approximated by conducting statistical survey.

As carried out in this study, tolling links and different tolling intervals were determined beforehand. In our example, tolling is applied during time interval [38, 53]. As shown in Fig. 5.4, the traffic in this period has dropped substantially. The results are not applicable in the real world. To best explore the dynamic road pricing schemes, tolling links and tolling intervals should be set as variables. This is an optimal tolling location and tolling level problem. One simple model extension is to produce a mixed 1-0 optimization problem of the MPEC. Each tolling link and tolling interval are variables with only two values: "1" means that the tolling link or the tolling interval is selected, while "0" otherwise. However, it is tremendously hard to obtain a solution. Trying some mixed integer programming techniques like branch and bound may be the

direction to solve this problem.

The DTA model used in this study can only lead to local optimal results. The proposed projection-based algorithm to solve DTA models may not be efficient for a large-scale network because of the large dimension of variables. More efforts should be put to design an efficient algorithm that gives reasonable solutions.

REFERENCES

1. Aashtiani, H. Z. and Magnanti, T. L., (1981). "Equilibria on a Congested Transportation Network." *SIAM Journal on Algebraic and Discrete Methods*, Vol. 2, No. 3, pp 213-226.
2. Alfa, A. S., (1986). "A Review of Models for the Temporal Distribution of Peak Traffic Demand." *Transportation Research Part B*, Vol. 20, No. 6, pp 491-499.
3. Arnott, R., de Palma, A. and Lindsey, R., (1990). "Departure Time and Route Choice for the Morning Commute." *Transportation Research Part B*, Vol. 24, No. 3, pp 209-228.
4. Audet, C. and J. E. Dennis, Jr., (2004), "Mesh Adaptive Direct Search Algorithms for Constrained Optimization." Rice CAAM Technical Report TR04-02.
[Online] Available:
<http://www.caam.rice.edu/caam/trs/2004/TR04-02-rev10-04.pdf>
[Access Date 12/07/2006]
5. Ben-Akiva, M., Bierlaire, M., Burton, D., Koutsopoulos, H. and Mishalani, R., (2001). "Network State Estimation and Prediction for Real-time Traffic Management." *Network and Spatial Economics*, Vol. 1, pp 293-318.
6. Berg, J. T. and Bhatt, K., (1998). "A report on U.S. Experience with the Congestion Pricing Pilot Program." *Proc. Int. Conf. on Transportation into the Next Millennium*, September 9-11, Singapore, pp 161-169.
7. Bard, J. F., (1988). "Convex Two-level Optimization." *Mathematical Programming*, Vol. 40, No. 1, pp 15-27.
8. Beckmann, M. J., McGuire, C. B. and Winsten, C. B., (1956). Studies in the Economics of Transportation, Yale University Press, New Haven, Conn.
9. Bliemer, M. C. J. and Bovy, P. H. L., (2003). "Quasi-Variational Inequality

- Formulation of the Multiclass Dynamic Traffic Assignment Problem.”
Transportation Research Part B, Vol. 37, No. 6, pp 501-519.
10. Booker, A. J., Dennis, J. E., Frank, Jr. P. D., Serafini, D. B., Torczon, V. and Trosset, M. W., (1999), “A Rigorous Framework for Optimization of Expensive Functions by Surrogates.” *Structural Optimization*, Vol. 17, No. 1, pp 1-13.
 11. Braid, R., (1996). “Peak–Load Pricing of a Transportation Route with an Unpriced Substitute.” *Journal of Urban Economics*, Vol. 40, No. 2, pp 179-197.
 12. Button, K. J., (1993). Transportation Economics, 2nd edition, Edward Elgar, England.
 13. Carey, M., (1986). “A constraint Qualification for a Dynamic Traffic Assignment Model.” *Transportation Science*, Vol. 20, No. 1, pp 55-58.
 14. Carey, M., (1987). “Optimal Time-varying Flows on Congested Networks.” *Operations Research*, Vol. 35, No. 1, pp 58-69.
 15. Carey, M., (1992). “Nonconvexity of the Dynamic Traffic Assignment Problem.” *Transportation Research Part B*, Vol. 26, No. 2, pp 127-133.
 16. Carey, M. and Srinivasan, A., (1993). “Externalities, Average and Marginal Costs, and Tolls on Congested Networks with Time-varying Flows.” *Operations Research*, Vol. 41, No. 1, pp 217-231.
 17. Chankong, V. and Haimes, Y. Y., (1983). Multiobjective Decision Making Theory and Methodology, North-Holland, New York.
 18. Charnes, A. and Cooper, W. W., (1961). Management models and Industrial Applications of Linear Programming, Vol. 1, John Wiley, New York.
 19. Chen, Y. and Florian, M., (1995). “The Nonlinear Bilevel Programming Problem: Formulations, Regularity and Optimality Conditions.” *Optimization*, Vol. 32, No. 2, pp 193-202.
 20. Chin, K. K., (2002). “Road Pricing Singapore's Experience.” *Proc. of*

IMPRINT-EUROPE Thematic Network: “Implementing Reform on Transport Pricing: Constraints and Solutions: Learning from Best Practice, Brussels, Belgium.

21. Conn, A. R., Gould, N. I. M. and Toint, Ph. L., (1996). “Numerical Experiments with the lancelot Package (Release A) for Large-scale Nonlinear Optimization.” *Mathematical Programming*, Vol. 73, No.1, pp 73-110.
22. Corne, D., Dorigo, M. and Glover, F., (1999) New Ideas in Optimization, McGraw-Hill.
23. Dafermos, S., (1980). “Traffic Equilibrium and Variational Inequalities.” *Transportation Science*, Vol. 14, No. 1, pp 42-54.
24. Dafermos, S., (1982). “The General Multimodal Network Equilibrium Problem with Elastic Demand.” *Network*, Vol. 12, pp 57-72.
25. Dafermos, S. C., and Sparrow, F. T., (1969). “The Traffic Assignment Problem For a General Network.” *Journal of Research of the National Bureau of Standards*, Vol. 73, pp 91-118.
26. Daganzo, C. F., (1994). “The Cell Transmission Model: A Simple Dynamic Representation of Highway Traffic.” *Transportation Research Part B*, Vol. 28, No. 4, pp 269-287.
27. Daganzo, C. F., (1995a). “The Cell Transmission Model Part II: Network Traffic.” *Transportation Research Part B*, Vol. 29, No. 2, pp 79-83.
28. Daganzo, C. F., (1995b). “Properties of Link Travel Time Functions under Dynamic Loads.” *Transportation Research Part B*, Vol. 29, No. 2, pp 93-98.
29. Deb, K., (1998), Multi-Objective Genetic Algorithms: Problem Difficulties and Construction of Test Problems, Technical Report CI-49/98. Dortmund: Department of Computer science/LS11, University of Dortmund, Germany.
30. Deb, K., Pratap, A., Agarwal, S. and Meyarivan, T., (2002). “A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II.” *IEEE Transactions on Evolutionary Computation*, Vol. 6, No. 2, pp 182-197.

31. Dirkse, S. P., Ferris, M. C. and Meeraus, A., (2002). Mathematical Programs with Equilibrium Constraints: Automatic Reformulation and Solution via Constraint Optimization, Technical Report NA-02/11, Oxford University Computing Laboratory.
32. Downs, A., (1993). "Point of View: Implementing Peak-hour Road Pricing at Full Scale: Finding Solution to Practical Problems." TR News, No. 167, pp 7-9.
33. Drud, A. S., (1992). "CONOPT: A Large-Scale GRC Code." ORSA Journal on Computing, Vol. 6, pp 207-216.
34. Duckstein, L., (1984). "Multiobjective Optimization in Structural Design: The Model Choice Problem." in Atrek, E., Gallagher, R. H., Ragsdell, K. M., and Zienkiewicz, O. C. editors, New Directions in Optimum Structural Design, pp 459-481, John Wiley and Sons.
35. Facchinei, F., Jiang, H. and Qi, L., (1996). A Smoothing Method for Mathematical Programs with Equilibrium Constraints, Technical Report 03.96, Universita di Roma, "LaSapienza".
36. Ferris, M. C., (1998), MATLAB and GAMS: Interfacing Optimization and Visualization software, Mathematical Programming Technical Report 98-19. Computer Sciences Department, University of Wisconsin, Madison, Wisconsin.
37. Ferris, M. C. and Kanzow, C., (1999). Complementarity and Related Problems: A Survey, Mathematical Programming Technical Report 98-17, University of Wisconsin.
38. Ferris, M. C. and Munson, T. S., (2000), "Complementarity Problems in GAMS and the PATH solver." Journal of Economic Dynamics and Control, Vol. 24, No. 2, pp 165-188.
39. Ferris, M. C. and Pang, J. S., (1997). "Engineering and Economic Applications of Complementarity Problems." SIAM Review, Vol. 39, No. 4, pp 669-713.
40. Ferris, M. C. and Tin, L. F., (1998). "Nonlinear Programming Approach for a Class of Inverse Problems in elastoplasticity." Structural Engineering and

Mechanics, Vol. 6, No. 8, pp 857-870.

41. Fletcher, R. and Leyffer, S., (2002a). "Nonlinear Programming without a Penalty Function." *Mathematical Programming*, Vol. 91, No. 2, pp 239-269.
42. Fletcher, R., and Leyffer, S., (2002b). Numerical Experience with Solving MPECs as NLPs, Numerical Analysis Report NA/210, Department of Mathematics, University of Dundee, Dundee, UK.
43. Fletcher, R., Leyffer, S., Ralph, D. and Scholtes, S., (2002). Local convergence of SQP methods for Mathematical Programs with Equilibrium Constraints, Numerical Analysis Report NA/209, Department of Mathematics, University of Dundee, Dundee, UK.
44. Fonseca, C. M. and Fleming, P. J., (1993). Genetic Algorithms for Multi-objective Optimization: Formulation, Discussion and Generalization, In Stephanie Forrest, editor, *Proceedings of the Fifth International Conference on Genetic Algorithms*, pp 416-423, San Mateo, California.
45. Fourman, M. P.,(1985), "Companion of Symbolic Layout Using Genetic Algorithms." *Proc. of the First International Conference on Genetic Algorithms*, pp 141-153.
46. Friesz, T. L., Anandalingam, G., Mehta, N. J., Nam, K., Shah, S. J. and Tobin, R. L., (1993). "The multiobjective Equilibrium Network Design Problem Revisited: A simulated Annealing Approach." *European Journal of Operational Research*, Vol. 65, No. 1, pp 44-57.
47. Friesz, T. L., Bernstein, D., Smith, T. E. and Wie, B. W., (1993). "A Variational Inequality Formulation of the Dynamic Network User Equilibrium Problem." *Operations Research*, Vol. 41, No. 1, pp 179-191.
48. Friesz, T. L., Leque, J., Tobin, R. L. and Wie, B. W., (1989). "Dynamic Network Traffic Assignment Considered as a Continuous Time Optimal Control Problem." *Operations Research*, Vol. 37, No. 6, pp 893-901.
49. Goldberg, D. E., (1989). Genetic Algorithms in Search, Optimization and Machine Learning, Addison-Wesley Publishing Company, Reading, Massachusetts.

50. Gupta, S., Kalmanje, S., and Kockelman, K. M., (2006). "Road Pricing Simulations: Traffic, Land Use and Welfare Impacts for Austin, Texas." *Transportation Planning and Technology*, Vol. 29, No. 1, pp 1-23.
51. Harrison, B., (1986). "Electronic Road Pricing in Hong Kong: Estimating and Evaluating the Effects." *Traffic Engineering And Control*, Vol. 27, No. 1, pp 13-18.
52. Han, S. J., (2003), "Dynamic Traffic Modeling and Dynamic Stochastic User Equilibrium Assignment for General Road Network." *Transportation Research Part B*, Vol. 37, No. 3, pp 225-249.
53. Harker, P. T. and Pang, J. S., (1990), "Finite-dimensional Variational Inequality and Nonlinear Complementarity Problems: A Survey of Theory, Algorithms and Applications." *Mathematical Programming, Series B*, Vol. 48, No. 2, pp 161-220.
54. Haimes, Y. Y., Lasdon, L. S. and Wismer, D. A., (1971). "On a Bicriteria Formulation of the Problems of Integrated System Identification and System Optimization." *IEEE Transactions on Systems, man, cybernetics*, Vol. 1, No. 3, pp 296-297.
55. Hansen, P., Jaumard, B. and Savard, G., (1992). "New Branch-and-bound Rules for Linear Bilevel Programming." *SIAM Journal on Scientific and Statistical Computing*, Vol. 13, No. 5, pp 1194-1217.
56. Hau, T. D.,(1992a). "Economic Fundamentals of Road Pricing: A Diagrammatic Analysis." *Policy Research Working Paper*, The World Bank.
57. Hau, T. D., (1992b), "Congestion Charging Mechanisms for Roads: An Evaluation of Current Practice." *Policy Research Working Paper*, The World Bank.
58. Hearn, D. W., Lawphongpanich, S., and Nguyen, S., (1984). "Convex Programming Formulations of the Asymmetric Traffic Assignment Problem." *Transportation Research Part B*, Vol. 18, No. 4-5, pp 357-365.
59. Hearn, D. W. and Yildirim, M. B., (1999). "A Toll Pricing Framework for Traffic Assignment Problems with Elastic Demand." *Research Report 99-8*, Department of Industrial and Systems Engineering, University of Florida.

60. Hearn, D. W. and Yildirim, M. B., (2002). "A Toll Pricing Framework for Traffic Assignment Problems with Elastic Demand." in *Transportation and Network Analysis: Current Trends*, Gendreau, M. and Marcotte, P. (eds), Kluwer Academic Publishers, Netherlands.
61. Holland, E. P. and Watson, P. L., (1978). "Traffic Restraint in Singapore: Measuring the Impacts of Area License Scheme." *Traffic Engineering and Control*, Vol. 19, No. 1, pp 14-17.
62. Hooke, R. and Jeeves, T. A., (1961). "Direct Search Solution of Numerical and Statistical Problems." *Journal of the Association of Computing Machinery*, Vol. 8, No. 2, pp 212-229.
63. Horn, J. and Nafpliotis, N., (1993). Multiobjective Optimization Using the Niche Pareto Genetic Algorithm, Technical Report IlliGAI Report 93005, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA.
64. Huang, H. J. and Lam W. H. K., (2002). "Modeling and Solving the Dynamic User Equilibrium Route and Departure Time Choice Problem in Network with Queues." *Transportation Research Part B*, Vol. 36, No. 3, pp 253-273.
65. Ieromonachou, P., Potter, S. and Warren, J. P., (2006). "Norway's Urban Toll Rings: Evolving towards Congestion Charging?" *Transport Policy*, Vol. 13, pp 367-378.
66. Ignizio, J. P., (1976). Goal Programming and Extensions, Lexington Books, Lexington, MA.
67. Ijiri, Y., (1965). Management Goals and Accounting for Control, North-Holland, Amsterdam.
68. In't Veld, R. J., (1991), "Road Pricing: A Logical Failure." in Kraan D. J. and in't Veld, R. J., (eds), *Environmental Protection: Public or Private Choice*, Kluwer Academic Publishers, Dordrecht , pp 111-121.
69. Janson B. N., (1991). "Dynamic Traffic Assignment for Urban Networks." *Transportation Research Part B*, Vol. 25, No. 3, pp 143-161.

70. Jiang, H. and Ralph, D., (1997). Smooth SQP Methods for Mathematical Programs with Nonlinear Complementarity Constraints, Manuscript, Department of Mathematics, University of Melbourne.
71. Joksimovic, D., Bliemer M. and Bovy, P., (2005). "Optimal Toll Design Problem in Dynamic Traffic Networks with Joint Route and Departure Time Choice." *Journal of the Transportation Research Record*, No. 1923, pp 61-72.
72. Jones, P., (2002). "Addressing Equity Concerns in Relation to Road User Charging." *Proc. of Conference on Acceptability of Transport Pricing Strategies*, Dresden, Germany.
73. Knowles, J. D. and Corne, D. W., (2000). "Approximating the Non-dominated Front Using Pareto Archived Evolutionary Strategy." *Evolutionary Computation*, Vol. 8, No. 2, pp 149-172.
74. Kuwahara, M. and Akamatsu, T., (1993). "Dynamic Equilibrium Assignment with Queues for a One-to-many OD Pattern." *Proceedings of the 12th International Symposium on Transportation and Traffic Theory*, Berkely, Elsevier, New York, pp 185-204.
75. Kuwahara, M. and Akamatsu, T., (1997). "Decomposition of the Reactive Dynamic Assignment with Queues for a Many-to-many Origin-destination Pattern." *Transportation Research Part B*, Vol. 31, No. 1, pp 1-10.
76. Kuwahara, M. and Akamatsu, T., (2001). "Dynamic User Optimal Assignment with Physical Queues for a Many-to-many OD Pattern." *Transportation Research Part B*, Vol. 35, No. 5, pp 461-479.
77. Land Transport Authority, (1996). The White Paper-A world Class Land Transport System, Singapore.
78. Land Transport Authority, (2005). A Decade of Travel Made Easy, Annual Report, Singapore.
79. Laumanns, M., Thiele, L., Deb, K. and Zitzler, E., (2002). "Combining Convergence and Diversity in Evolutionary Multiobjective Optimization." *Evolutionary Computation*, Vol. 10, No. 3, pp 263-282.

80. Leape, J., (2006). "The London Congestion Charge." *Journal of Economic Perspectives*. Vol. 20, No. 4, pp 157-176.
81. Leyffer, S., (2002). The Penalty Interior Point Method Fails to Converge For Mathematical Programs with Equilibrium Constraints, Numerical Analysis Report NA/208, Department of Mathematics, University of Dundee.
82. Li, J., Fujiwara, O. and Kawakami, S., (2000). "A Reactive Dynamic User Equilibrium Model in Network with Queues." *Transportation Research Part B*, Vol. 34, No. 8, pp 605-624.
83. Li, M. Z. F., (1999). "Estimating Congestion Toll by Using Traffic Count data–Singapore’s Area Licensing Scheme." *Transportation Research Part E*, Vol. 35, No. 1, pp 1-10.
84. Li, M. Z. F., (2002). "The Role of Speed-flow Relationship in Congestion Pricing Implementation with and Application to Singapore." *Transportation Research Part B*, Vol. 36, No. 8, pp 731-754.
85. Lim, A. C., (2002). Transportation Network Design Problems: An MPEC Approach, Ph. D Dissertation, Department of Mathematics, Johns Hopkins University, Baltimore, Maryland.
86. Lindberg, G., (1995). "Road Pricing: Policy and Options for the Future." in: Johansson, B. and Mattson, L.-G. (Eds.), *Road Pricing: Theory, Empirical Assessment and Policy*, Kluwer, Boston, pp 205-221.
87. Liu, L. N. and McDonald, J., (1998). "Efficient Congest Tolls in the Presence of Unpriced Congestion: A Peak and Off-Peak Simulation Model." *Journal of Urban Economics*, Vol. 44, No. 3, pp 352-366.
88. Liu, L. N. and McDonald, J., (1999). "Economic Efficiency of Second-best Congestion Pricing Schemes in Urban Highway Systems." *Transportation Research Part B*, Vol. 33, No. 3, pp 157-188.
89. Lo, H. K. and Szeto, W. Y., (2002). "A Cell-based Variational Inequality Formulation of the Dynamic User Optimal Assignment Problem." *Transportation Research Part B*, Vol. 36, No. 5, pp 421-443.

90. Lo, H. K. and Szeto, W. Y., (2005). "Road Pricing Modeling for Hyper-congestion." *Transportation Research Part A*, Vol. 39, No. 7-9, pp 705-722.
91. Luo, Z. Q., Pang, J. S. and Ralph, D., (1996). Mathematical Programs with Equilibrium Constraints. Cambridge University Press, New York.
92. Marchand, M., (1968). "A Note on Optimal Tolls in an Imperfect Environment." *Econometrica*, Vol. 36, No. 3-4, pp 575-581.
93. May, A. D. and Sumalee, A., (2003), "One Step Forwards, Two Steps Back? An Overview of Road Pricing Application and Research outside the US." *Proc. of International Symposium on Road Pricing*, Key Biscayne, Florida.
94. Meng, Q., and Yang, H., (2002). "Benefit Distribution and Equity in Road Network Design." *Transportation Research Part B*, Vol. 36, No. 1, pp 19-35.
95. Menon, A. P. G., (2000). "ERP in Singapore - A Perspective One Year on." *Traffic Engineering And Control*, Vol. 41, No. 2.
96. Merchant, D. K. and Nemhauser, G. L., (1978a). "A Model and an Algorithm for the Dynamic Traffic Assignment Problems." *Transportation Science*, Vol. 12, No. 3, pp 183-199.
97. Merchant, D. K. and Nemhauser, G. L., (1978b). "Optimality Conditions for a Dynamic Traffic Assignment Model." *Transportation Science*, Vol. 12, No. 3, pp 200-207.
98. Nagurney, A., (1993). Network economics: a variational inequality approach, Kluwer Academic Publishers, Boston.
99. Nagurney, A. and Zhang, D., (1997a). "Projected Dynamic Systems in the Formulation, Stability Analysis, and Computation of Fixed-demand Traffic Network Equilibria." *Transportation Science*, Vol. 31, No. 2, pp 147-158.
100. Nagurney, A. and Zhang, D., (1997b). *Projected Dynamical Systems and Variational Inequalities with Applications*. Kluwer Academic Publishers, Boston, MA.

101. Nash, J. F., (1951). "Nocooperative games." *Annals of Mathematics*, Vol. 45, pp 286-295.
102. Newell, G. F., (1993a). "A simplified Theory of Kinematic Waves in Highway Traffic, Part II: General Theory." *Transportation Research Part B*, Vol. 27, No. 4, pp 289-304.
103. Newell, G. F., (1993b). "A simplified Theory of Kinematic Waves in Highway Traffic, Part III: Multi-destination Flows." *Transportation Research Part B*, Vol. 27, No. 4, pp 305-314.
104. Odeck, J. and Brathen, S., (2002). "Toll Financing in Norway: The Success, the Failures and Perspectives for the Future." *Transport Policy*, Vol. 9, No. 3, pp 253-260.
105. Pang, J. S., (1994), "Complementarity problems." in *Handbook of Global Optimization*, R. Horst and P. Pardalos, eds., Kluwer Academic Publishers, Norwell, MA.
106. Pang, J. S., Han, S. P. and Rangaraj, N., (1991). "Minimization of Locally Lipschitzian functions." *SIAM Journal of Optimization*, Vol. 1, No. 1, pp 57-82.
107. Patriksson, M., (1994). *The Traffic Assignment Problem: Models and Methods*, Utrecht, The Netherlands.
108. Patriksson, M., (2004). "Algorithms for Computing Traffic Equilibria." *Networks and Spatial Economics*, Vol. 4, No. 1, pp 23-38.
109. Pigou, A. C., (1920). Wealth and Welfare, Macmillan, London.
110. Pontryagin, L. S., Boltyanskii, V. A., Gramkrelidze, R. V. and Mischenko, E. F., (1962). The Mathematical Theory of Optimal Processes, John Wiley & Sons, New York.
111. Prager, W., (1954). "Problems of Traffic and Transportation." *Proceedings of the Symposium on Operations Research in Business and Industry*, pp 105-113.

112. Ran, B., Boyce, D. E. and Leblanc, L. J., (1993). "A New Class of Instantaneous Dynamic User-optimal Traffic Assignment Models." *Operations Research*, Vol. 41, No. 1, pp 192-202.
113. Ran, B. and Boyce, D. E., (1996). Modeling Dynamic Transportation Networks : An Intelligent Transportation System Oriented Approach. Springer, Germany.
114. Schade, J. and Baum, M., (2007). "Reactance or Acceptance? Reactions towards the Introduction of Road Pricing." *Transportation Research Part A*, Vol 41, No. 1, pp 41-48.
115. Schade, J. and Schlag, B., (2003). "Acceptability of Urban Transport Pricing Strategies." *Transportation Research Part F*, Vol. 6, No. 1, pp 45-61.
116. Schaffer, J. D.,(1985). "Multiple Objective Optimization with Vector Evaluated Genetic Algorithms." *Proceedings of the First International Conference on Genetic Algorithms*, pp 93-100.
117. Scheel, H. and Scholtes, S., (2000). "Mathematical Program with Complementarity Constraints: Stationarity, Optimality and Sensitivity." *Mathematics of Operations Research*, Vol. 25, No. 1, pp 1-22.
118. Seik, F. T., (1998). "A Unique Demand Management Instrument in Urban Transport: the Vehicle Quota System in Singapore." *Cities*, Vol. 15, No. 1, pp 27-39
119. Small, K. A., (1982). "The Scheduling of Consumer Activities: Work Trips." *American Economic Review* 72, pp 467-479.
120. Small, K. A. and Gomez-Ibanez, J. A., (1998). "Road Pricing for Congestion Management: the Transition from Theory to Policy." *Road Pricing, Traffic Congestion, and the Environment: Issues of Efficiency and Social Feasibility*, K. Button and E. T. Verhoef, eds., Edward Elgar Publishing, Massachusetts, pp 213-246.
121. Smith, M. J., (1979), "The Existence, Uniqueness and Stability of Traffic Equilibria." *Transportation Research Part B*, Vol. 13, No. 4, pp 295-304.

122. Smith, M. J., (1993). "A New Dynamic Traffic Model and the Existence and Calculation of Dynamic User Equilibrium on Congested Capacity-constrained Road Networks." *Transportation Research Part B*, Vol. 27, No. 1, pp 49-63.
123. Srinivas, P., (1994). System Optimal Dynamic Traffic Assignment in Congested Networks with Advanced Information Systems.
[Online] Available: <http://ce.ecn.purdue.edu/~peeta/dissertation.pdf>
[Access Data 12/07/2006]
124. Srinivas, N. and Deb, K., (1993). Multiobjective Optimization Using Non-dominated Sorting in Genetic Algorithms, Technical report, Department of Mechanical Engineering, Indian Institute of Technology, Kanpur, India.
125. Stackelberg, H. Von., (1952). The Theory of Market Economy, Oxford University Press, England.
126. Storn, R. and Price, K., (1995). Differential Evolution - A Simple and Efficient Adaptive Scheme for Global Optimization over Continuous Spaces. Technical Report TR-95-012, International Computer Science Institute, University of California at Berkeley.
[Online] Available:
<ftp://ftp.icsi.berkeley.edu/pub/techreports/1995/tr-95-012.pdf>
[Access Data 12/07/2006]
127. Stewart, K., (2007). "Tolling Traffic Links under Stochastic Assignment: Modelling the Relationship between the Number and Price Level of Tolloed Links and Optimal Traffic Flows." *Transportation Research Part A*, Vol. 41, No. 7, pp 644-654.
128. Sullivan, E. C., (2000). "Continuation Study to Evaluate the Impacts of the SR 91 Value-Priced Express Lanes: Final Report." Sacramento: California Department of Transportation.
[Online] Available:
http://ceenve.calpoly.edu/sullivan/SR91/final_rpt/FinalRep2000.pdf
[Access Data 12/07/2006]
129. Tong, C. O. and Wong, S. C., (2000). "A Predictive Dynamic Traffic Assignment Model in Congested Capacity-constrained Road Networks." *Transportation Research Part B*, Vol. 34, No. 8, pp 625-644.

130. Verhoef, E. T., (2002). "Second-best Congestion Pricing in General Networks: Heuristic Algorithms for Finding Second-best Optimal Toll Levels and Toll Points." *Transportation Research Part B*, Vol. 36, No. 8, pp 707-729.
131. Vickrey, W. S., (1993). "Point of View: Principles and Applications of Congestion Pricing." *TR News*, Vol. 167, pp 4-5.
132. Vickrey, W. S., (1969). "Congestion Theory and Transport Investment." *American Economic Review*, Vol. 59, No. 2, pp 251-260.
133. Walters, A. A., (1961). "The Theory and Measurement of Private and Social Cost of Highway Congestion." *Econometrica*, Vol. 29, No. 4, pp 676-699.
134. Wardrop, J. G., (1952). "Some Theoretical Aspects of Road Traffic Research." *Proceedings of the Institute of Civil Engineers, Part II*, pp 325-378.
135. Wie, B. W., Friesz, T. L. and Tobin, R. L., (1990). "Dynamic User Optimal Traffic Assignment on Congested Multi-destination Networks." *Transportation Research Part B*, Vol. 24, No. 6, pp 431-442.
136. Wie, B. W., Friesz, T. L., Tobin, R. L. and Bernstein, D., (1995). "A Discrete Time, Nested Cost Operator Approach to the Dynamic Network User Equilibrium Problem," *Transportation Science*, Vol. 29, No. 1, pp 79-92.
137. Wie, B. W. and Tobin, R. L., (1998). "On the Relationship between Dynamic Nash and Instantaneous User Equilibria." *Networks*, Vol. 32, No. 2, pp 141-163.
138. Wu, J. H., Chen, Y., and Florian, M., (1998). "The Continuous Dynamic Network Loading Problem: A Mathematical Formulation and Solution Method." *Transportation Research Part B*, Vol. 32, No. 3, pp 173-187.
139. Xu, Y., Wu, J .H., Florian, M., Macrotte, P. and Zhu, D. L., (1999). "Advances in the Continuous Dynamic Network Loading Problem." *Transportation Science*, Vol. 33, No. 4, pp 341-353.

140. Yang, H. and Huang, H. J., (1997). "Analysis of the Time-varying Pricing of a Bottleneck with Elastic Demand Using Optimal Control Theory." *Transportation Research Part B*, Vol. 31, No. 6, pp 425-440.
141. Yang, H. and Huang, H. J., (1998). "Principle of Marginal-cost Pricing: How Does It Work in a General Road Network?" *Transportation Research Part A*, Vol. 32, No. 1, pp 45-54.
142. Yang, H. and Lam, W. H. K., (1996). "Optimal Road Tolls under Conditions of Queuing and Congestion." *Transportation Research Part A*, Vol. 30, No. 5, pp 319-332.
143. Yang, H. and Meng, Q., (1998). "Departure Time, Route Choice and Congestion Toll in a Queuing Network with Elastic Demand." *Transportation Research Part B*, Vol. 32, No. 4, pp 247-260.
144. Yang, H., Meng, Q. and Lee, D. H., (2004). "Trial-and-error Implementation of Marginal-cost Pricing on Networks in the Absence of Demand Functions." *Transportation Research Part B*, Vol. 38, No. 6, pp 477-493.
145. Zitzler, E. and Thiele, L., (1999). "Multiobjective Evolutionary Algorithms: A Comparative Case Study and the Strength Pareto Approach." *IEEE Transactions on Evolutionary Computation*, Vol. 3, No. 4, pp 257-271.