



Terms and Conditions of Use of Digitised Theses from Trinity College Library Dublin

Copyright statement

All material supplied by Trinity College Library is protected by copyright (under the Copyright and Related Rights Act, 2000 as amended) and other relevant Intellectual Property Rights. By accessing and using a Digitised Thesis from Trinity College Library you acknowledge that all Intellectual Property Rights in any Works supplied are the sole and exclusive property of the copyright and/or other IPR holder. Specific copyright holders may not be explicitly identified. Use of materials from other sources within a thesis should not be construed as a claim over them.

A non-exclusive, non-transferable licence is hereby granted to those using or reproducing, in whole or in part, the material for valid purposes, providing the copyright owners are acknowledged using the normal conventions. Where specific permission to use material is required, this is identified and such permission must be sought from the copyright holder or agency cited.

Liability statement

By using a Digitised Thesis, I accept that Trinity College Dublin bears no legal responsibility for the accuracy, legality or comprehensiveness of materials contained within the thesis, and that Trinity College Dublin accepts no liability for indirect, consequential, or incidental, damages or losses arising from use of the thesis for whatever reason. Information located in a thesis may be subject to specific use constraints, details of which may not be explicitly described. It is the responsibility of potential and actual users to be aware of such constraints and to abide by them. By making use of material from a digitised thesis, you accept these copyright and disclaimer provisions. Where it is brought to the attention of Trinity College Library that there may be a breach of copyright or other restraint, it is the policy to withdraw or take down access to a thesis while the issue is being resolved.

Access Agreement

By using a Digitised Thesis from Trinity College Library you are bound by the following Terms & Conditions. Please read them carefully.

I have read and I understand the following statement: All material supplied via a Digitised Thesis from Trinity College Library is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of a thesis is not permitted, except that material may be duplicated by you for your research use or for educational purposes in electronic or print form providing the copyright owners are acknowledged using the normal conventions. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone. This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

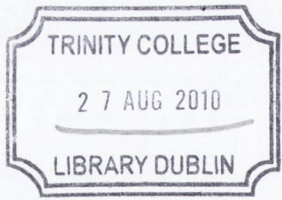
**Game Theoretic Formulations and Solution
Methods for Transport Network Reliability
and Reliable Network Design**

Liam O'Brien

**A dissertation submitted to the University of Dublin in the
partial fulfilment of the requirements for the Degree of
Doctor of Philosophy**

**Department of Civil, Structural & Environmental Engineering
Trinity College Dublin**

January 2010



DECLARATION

TH0818
8899.

I hereby declare that this dissertation, in whole or in part has not been submitted as an exercise for a degree at this or any other University. I further declare except where reference is given in the text, the work is entirely my own. I also give the library permission to lend or copy the thesis, upon request, for academic purposes.

Liam O'Brien.

Liam O'Brien

January 2010

DEDICATION

This work is dedicated to my late Grandmother, Mary Clarke, who passed away at the end of last year. Her sustained belief in the value of education and hard work, which she imparted to her family, is an enduring legacy.

SUMMARY

Transport network reliability has emerged as an issue of immense importance over the past decade leading many researchers to propose various reliability measures and approaches to assess transport network reliability and to design reliable networks. A number of modelling approaches have been proposed, including the game-theoretic approach. Traditionally, game theoretic approaches have relied on the outcome of a game played between network users seeking to minimise their travel costs and an origin destination (OD)-specific demon that seeks the opposite by damaging links on each OD pair in the network. The traffic assignment problem when formulated based on this approach is known as risk-averse traffic assignment. The expected network cost of this game at equilibrium can be treated as a measure of the overall network reliability while the link failure probabilities indicate links that are critical for network performance. The objective of this thesis is to propose new models of risk-averse traffic assignment, which assess the network reliability and consider many more possibilities and appropriate assumptions than existing approaches in order to capture the worst-case scenarios of network reliability and identify the vulnerable network links. Specific measures of network reliability are also defined and models are proposed for designing reliable transport networks. To solve some of the formulations, a solution method is proposed, which guarantees finite convergence arbitrarily close to the global optimum and does not rely on trial and error to guess the suitable initial solutions.

In this thesis, the proposed approaches are formulated into various mathematical programs. Chapter 2 proposes a non-linear complementarity problem (NCP) formulation of risk-averse traffic assignment with elastic demands. The NCP is transformed to a smooth and unconstrained mathematical program via a gap function and then solved by the newly proposed modified α -based branch and bound (α BB) algorithm. To improve the computation speed, a sampling method is suggested to calculate the α parameter for the lower bounding function. Through numerical studies, the convergence rate of the proposed solution method is illustrated. Chapter 3 proposes a more general network-specific demon formulation of risk-averse assignment via the NCP approach by relaxing the assumptions on the number/type of demons and the form of capacity reduction. This model imposes no restriction on the number of demons or number of links that can be damaged. Numerical studies were conducted to examine the effects of the network-specific demons and capacity reduction assumptions on expected network cost and reliability measures. Chapter 4 develops a link-based NCP formulation of risk-averse traffic assignment, proves its equivalence to the path-based approach, and in addition, defines specific reliability

measures. Numerical studies were conducted to demonstrate the existence of stochastic Braess' paradox and reliability paradox (analogues of Braess' paradox for expected network travel cost and reliability) in risk-averse assignment, and some further insights are provided into the implications for network design if these paradoxes are ignored. The NCP formulations in chapters 3 and 4 are transformed to smooth and unconstrained minimisation programs via a gap function and solved by the generalized reduced gradient (GRG) algorithm. A proof of the existence of solutions is given for the path and link-based approaches. Chapter 5 proposes a reliable continuous network design problem (RCNDP) and demonstrates important aspects of the design approach through numerical studies. Chapter 6 proposes two NDP formulations in a time-dependent framework: the conventional (i.e., without uncertainty) discrete network design problem over time (DNDP-T) and the corresponding reliable model (RDNDP-T). Numerical studies demonstrate the design approach of the RDNDP-T and the importance of introducing the time dimension to the NDP and discrete decision variables to the NDP over time. The models in chapters 5 and 6 are bi-level in nature but formulated as single-level problems and solved by the GRG algorithm and a combination of the GRG and branch and bound (B & B) algorithm, respectively.

The main results obtained are summarised as follows. In terms of the proposed solution method, the results show the factors which influence the rate of convergence. They also demonstrate the importance of the assumptions on the type and number of demons to assume, and how important it is to accurately capture the link capacity degradations. Relaxing these assumptions can allow many more operational states of a road network to be modelled and can provide a better (more pessimistic) approximation of the worst-case expected network cost for network reliability. Furthermore, the results showed that failure to account for paradoxical phenomena can imply that the network is more reliable than it actually is. In terms of designing reliable networks, the results show the ability of the RCNDP model to introduce higher capacity expansions on links vulnerable to failure and to consider many operational states to cater for the worst-case scenarios of network reliability. The results from the DNDP-T and RDNDP-T models highlight the practical importance of considering improvements to the network in the form of indivisible facilities, the importance of considering the time dimension in the DNDP in terms of providing better network performance, and the vulnerable network link improvements over time. Finally, using the risk-averse frameworks proposed in this thesis, one can assess the reliability of transport networks, develop a reliable network design, and analyse the effect of assumptions on traveller behaviour in each approach on network design and reliability.

ACKNOWLEDGEMENTS

I would like to take this opportunity to thank a number of people whose help and support greatly assisted me while completing this thesis.

Firstly, I am deeply grateful to my supervisor, Prof. Margaret O'Mahony, for her support throughout my time spent as a postgraduate research student and in particular for the many invaluable discussions and helpful comments which enabled me to complete this thesis.

I'd like to also acknowledge the help and guidance of my former supervisor, Dr. Wai Yuen Szeto, during the earlier stages of this work and for introducing me to the topic.

The financial support from the 'Trinity College Dublin Postgraduate Studentship' and from the 'Centre for Transport Research and Innovation for People' is gratefully acknowledged.

I would also like to acknowledge the help received from the faculty, staff and the postgraduate students of the Department of Civil, Structural and Environmental Engineering.

In addition, I'd like to thank Brian Marsden and Vivian O'Brien for giving their time so generously in helping to print the final draft of this thesis.

I am deeply indebted to so many friends who can always be counted on for their support and willingness to help others. In particular, I'd like to thank Tom, Brian, Bidisha and Vikram for their friendship and unequivocal support.

The most heartfelt thanks to Mary for the love, loyalty, encouragement and unwavering support she has shown me during all our years together and in particular during the times when it was most needed.

Finally, I am grateful to my family: my parents William and Anne, my brother Enda and sisters Deirdre, Clare and Vivian. Their constant support, encouragement and positive attitude have inspired me throughout my entire life. Their contribution in supporting me throughout the course of this thesis and my education in general, is impossible to measure.

TABLE OF CONTENTS

SUMMARY..... iv

ACKNOWLEDGEMENTS..... vi

TABLE OF CONTENTS..... vii

LIST OF FIGURES..... xii

LIST OF TABLES..... xiv

CHAPTER 1: INTRODUCTION AND LITERATURE REVIEW

1.1 Background..... 1

1.2 Literature Review..... 2

 1.2.1 Risk-Averse User Equilibrium Traffic Assignment..... 2

 1.2.2 Transport Network Reliability..... 4

 1.2.3 Braess’ Paradox..... 13

 1.2.4 The Transport Network Design Problem (NDP)..... 13

 1.2.5 Solution Methods..... 17

1.3 Objectives of the Thesis..... 19

1.4 Organisation of the Thesis..... 20

CHAPTER 2: RISK-AVERSE TRAFFIC ASSIGNMENT WITH ELASTIC DEMANDS: NCP FORMULATION AND SOLUTION METHOD FOR ASSESSING PERFORMANCE RELIABILITY

2.1 Introduction..... 23

2.2 Formulation..... 25

2.3 Modified α -Based Branch-And-Bound (α BB) Solution Method..... 32

 2.3.1 Similarities of the Modified and Conventional α BB Algorithms..... 32

 2.3.2 The Key Differences between the Modified and the Conventional α BB Algorithms..... 35

 2.3.3 Heuristic Sampling Method to Calculate α 37

 2.3.4 Steps of the Proposed Method Incorporating Sampling..... 37

2.4 Numerical Studies..... 40

 2.4.1 Base Scenario: The Convergence Rate of the Algorithm..... 42

 2.4.2 Scenario 2: Comparison of the Two Bound Improving Strategies..... 44

2.4.3	Scenario 3: The Effect of Sample Size on Solution Speed.....	46
2.4.4	Scenario 4: The Effect of Fixed User Defined α	47
2.5	Concluding Remarks.....	47

CHAPTER 3: GENERALISATION OF THE RISK-AVERSE TRAFFIC ASSIGNMENT

3.1	Introduction.....	49
3.2	Formulation.....	50
3.2.1	The User Problem.....	51
3.2.2	The Demon Problem.....	53
3.2.3	The Non-linear Complementarity Problem Formulation.....	54
3.2.4	The Proof of the Existence of Solutions.....	56
3.3	Numerical Studies.....	59
3.3.1	The Effect of Introducing Multiple Network-Specific Demons.....	60
3.3.2	The Effect of Introducing Capacity Reduction in some Proportion to the Number of Demons.....	64
3.3.3	The Demon Probabilities – the Demons’ Evil Behaviour.....	67
3.3.4	The Existence of Multiple Solutions.....	73
3.3.5	OD-Specific vs. Network-Specific Demons in a Multi-Commodity Network.....	74
3.4	Concluding Remarks.....	77

CHAPTER 4: MEASURING NETWORK RELIABILITY CONSIDERING PARADOXES: THE MULTIPLE NETWORK DEMON APPROACH

4.1	Introduction.....	79
4.2	Formulation.....	81
4.2.1	The Demon Problem.....	82
4.2.2	The User Problem.....	84
4.2.3	Link-Based Non-linear Complementarity Problem Formulation.....	89
4.3	The Reliability Measures.....	91
4.3.1	Encountered Reliability.....	91

4.3.2	Travel Time Reliability.....	93
4.4	Numerical Studies.....	95
4.4.1	The Effect of Multiple Network-Specific Demons and Multiple States of Link Damage to Reliability Measurement.....	95
4.4.2	Stochastic Braess' Paradox and Reliability Paradox in Risk-Averse User Equilibrium Traffic Assignment.....	99
4.5	Concluding Remarks.....	101

CHAPTER 5: DESIGNING RELIABLE TRANSPORT NETWORKS: THE MULTIPLE NETWORK DEMON APPROACH

5.1	Introduction.....	103
5.2	Formulation.....	106
5.2.1	Generalised Risk-Averse Traffic Assignment Constraints.....	106
5.2.2	Design Constraints.....	111
5.2.3	Improvement Cost Functions and Budgetary Constraint.....	111
5.2.4	Tolling Conditions, Revenue Functions and the Cost-Recovery Constraint.....	112
5.2.5	The Objective Function.....	113
5.2.6	The Multiple Network Demon RCNDP Formulations.....	113
5.2.7	The OD-Specific Demon RCNDP Formulation.....	114
5.3	Numerical Studies.....	115
5.3.1	The Effect of Multiple Network Demons in Reliable Transport Network Design.....	115
5.3.2	OD-Specific vs. Network-Specific Demons in the Design of a Multi-Commodity Network.....	122
5.3.3	Discussion.....	126
5.4	Concluding Remarks.....	126

CHAPTER 6: CONVENTIONAL AND RELIABLE FORMULATIONS OF THE DISCRETE NETWORK DESIGN PROBLEM OVER TIME

6.1	Introduction.....	128
6.2	Background to DNDP-T.....	129
6.3	Description of the NDP-T.....	132

6.4	Formulation of the DNDP-T Model.....	133
6.4.1	Framework and Assumptions.....	133
6.4.2	Time-Dependent Traffic Assignment Constraints.....	134
6.4.3	The Definitional Constraints.....	135
6.4.4	Travel Demand Functions.....	136
6.4.5	Design and Toll Constraints.....	137
6.4.6	Cost and Revenue Functions.....	138
6.4.7	Cost Recovery Constraints.....	139
6.4.8	The Objective Function.....	140
6.4.9	The DNDP-T Formulation.....	140
6.5	Solution Method.....	141
6.6	Numerical Studies Involving the DNDP-T.....	142
6.6.1	The Comparison of the DNDP-T and CNDP-T under the ACR Principle.....	142
6.6.2	The Comparison of the DNDP-T and DNDP under the OCR Principle.....	146
6.6.3	Comparison of OCR and ACR under the DNDP-T.....	151
6.6.4	Discussion.....	153
6.7	The RDNDP-T Model.....	153
6.8	Formulation of the RDNDP-T.....	154
6.8.1	Time-Dependent Risk-Averse Traffic Assignment Constraints.....	154
6.8.2	The Definitional Constraints.....	157
6.8.3	The Potential Travel Demand Function.....	158
6.8.4	Design Constraints.....	158
6.8.5	Improvement Cost Functions and Budgetary Constraints.....	158
6.8.6	The Objective Function.....	159
6.8.7	The RDNDP-T Formulation.....	159
6.9	Numerical Example Involving the RDNDP-T.....	160
6.9.1	The Design Approach of the RDNDP-T Model.....	161
6.10	Concluding Remarks.....	165

CHAPTER 7: CONCLUSIONS

7.1	Research Summary.....	167
7.2	Research Findings and Contributions.....	169

7.3	Critical Assessment.....	172
7.4	Recommendations for Future Research.....	174
7.4.1	Formulating the Problem as a Stackelberg Game.....	175
7.4.2	Application to Larger Networks.....	175
7.4.3	The Game of Increasing Adversity.....	175
7.4.4	Pessimistic versus Realistic Demon Link Selection Probabilities.....	176
7.4.5	Other Future Research Directions.....	176

REFERENCES

**APPENDIX A: OUTLINE OF THE GENERALIZED REDUCED
GRADIENT (GRG) ALGORITHM**

**APPENDIX B: LIST OF AUTHOR’S PUBLICATIONS AND
AWARDS**

LIST OF FIGURES

Figure 2.1:	The convex lower bounding function, $L(x)$, of the objective function, $G(x)$ in the one-dimension case.....	33
Figure 2.2:	The example network.....	41
Figure 2.3:	The convergence plot for the base scenario.....	41
Figure 2.4:	The convergence of damage probabilities and the rescaled total expected trip cost.....	42
Figure 2.5:	The convergence of the rescaled route flows.....	42
Figure 2.6:	The relationship among the number of iterations performed, the number of function evaluations and computation time.....	43
Figure 2.7:	A comparison of two bound improving strategies in terms of gap against computation time.....	44
Figure 2.8:	A comparison of two bound improving strategies in terms of number of partitions over iterations.....	44
Figure 2.9:	Comparison of the effect of the sample size on convergence speed.....	45
Figure 3.1:	Test network 1.....	58
Figure 3.2:	The total expected network cost for a range of congested conditions.....	59
Figure 3.3:	The total expected network cost for a range of demons from free-flow to congested conditions when a linear capacity reduction is assumed.....	63
Figure 3.4:	The total expected network cost for a range of demons from free-flow to congested conditions when a non-linear capacity reduction is assumed.....	64
Figure 3.5:	The total expected network cost vs. congestion levels for the three different assumptions on link capacity degradation for the three demon case.....	64
Figure 3.6:	Test network 2.....	65
Figure 3.7:	The link flow patterns for the case of one active demon.....	66
Figure 3.8:	The link damage probabilities for the one demon case.....	66
Figure 3.9:	The link flow patterns for the case of two active demons.....	68

Figure 3.10:	The maximum expected pay-off vs. level of congestion in the network.....	70
Figure 3.11:	Test network 3.....	72
Figure 4.1:	Braess' network.....	92
Figure 5.1:	Test network 1.....	113
Figure 5.2:	Test network 2.....	119
Figure 6.1:	The NDP-T.....	128
Figure 6.2:	Test network 1.....	139
Figure 6.3:	Test network 2.....	142
Figure 6.4a:	Link capacities under ACR.....	147
Figure 6.4b:	Link capacities under OCR.....	147
Figure 6.5:	Test network 3.....	158

LIST OF TABLES

Table 2.1:	Settings for the four scenarios.....	40
Table 2.2:	The number of function evaluations over iterations.....	43
Table 2.3:	The effect of the user defined α values on the computation performance.....	46
Table 3.1a:	Link capacities for the one demon case.....	60
Table 3.1b:	Link capacities for the two demon case.....	60
Table 3.2:	The damaged link capacities.....	62
Table 3.3:	Scenarios for two demons in network 2.....	67
Table 3.4:	The demons' link selection probabilities for the two demon case.....	69
Table 3.5:	Another set of the optimal demons' link selection probabilities.....	71
Table 3.6:	Link selection probabilities of the demons for case 1.....	73
Table 3.7:	Link selection probabilities of the demons for case 2.....	74
Table 4.1:	The effect of multiple network-specific demons and multiple states of link damage on network reliability measurement.....	94
Table 4.2:	The effect of Braess' paradox on network reliability measurement.....	96
Table 5.1:	The damaged link capacities.....	114
Table 5.2:	Route Flows for the different cases of capacity reduction.....	114
Table 5.3:	Demon link selection probabilities for the different cases of capacity reduction.....	115
Table 5.4:	Description of the scenarios of link damage.....	115
Table 5.5:	Scenario probabilities for the different cases of capacity reduction.....	116
Table 5.6:	ETSTT values for the different cases of capacity reduction.....	116
Table 5.7:	Link capacity improvements for the different cases of capacity reduction.....	117
Table 5.8:	Link tolls for the different cases of capacity reduction.....	117
Table 5.9:	Route flows for the network and OD-specific demons.....	120
Table 5.10:	Link flows for the network and OD-specific demons.....	120
Table 5.11:	Link selection probabilities of the network and OD-specific demons.....	121

Table 5.12:	Link capacity improvements for the RCNDP with different types of demon.....	121
Table 6.1:	Link improvements and tolls for DNDP-T and CNDP-T.....	140
Table 6.2:	Hourly demand and TDCS for DNDP-T and CNDP-T.....	141
Table 6.3:	Link improvements for the DNDP-T and DNDP under OCR.....	143
Table 6.4:	Link tolls for the DNDP-T and DNDP under OCR.....	143
Table 6.5:	Link flows for the DNDP-T and DNDP under OCR.....	143
Table 6.6:	Hourly travel demands and TDCS for the DNDP-T and DNDP under OCR.....	144
Table 6.7:	OD Travel Costs for DNDP-T and DNDP.....	146
Table 6.8:	Tolling Strategies under OCR and ACR.....	148
Table 6.9:	Link selection probabilities for the network and OD-specific demons over time.....	159
Table 6.10:	Link flows over time.....	159
Table 6.11:	Link capacity improvements for the RDNDP-T model.....	160

CHAPTER 1

Introduction and Literature Review

1.1 BACKGROUND

In the past decade the reliability of transport networks has emerged as an issue of great importance for researchers worldwide. This has occurred for a number of reasons. Firstly, the experience of events such as the Kobe earthquake of 1995 have led many researchers to identify the transport network as a critical element in acting as a lifeline to ferry emergency services to and from the scene of a major incident in the network. Secondly, increased economic activity worldwide has led to an increase in the values of time and increased users' expectations on how the road network should perform. Road networks are now required to be highly reliable to ensure that users can experience smooth travel under both normal and abnormal traffic conditions so that users can avoid unexpected delays due to disruptions such as adverse weather, traffic accidents, breakdowns, signal failures, road-works, and landslides.

This requirement has led many researchers to develop various reliability measures or indicators to assess the transport network reliability and to design reliable networks. These measures to evaluate the network reliability impacts have been studied under various different approaches. For transport systems, there are numerous models for helping to identify the consequence of system degradation (e.g., increased travel times, distances and costs), but there is little experience in the identification of the probability of disruption. One approach is to consider the worst-case scenario, where links in a network are selected for failure (in the sense of a reduction in link capacity) from a malevolent agency, which is seeking to cause the greatest possible impact. This approach, known as the game theory approach to network reliability, is the approach adopted in this thesis. Building on some initial work in the field, this approach is further developed in this thesis to assess the worse-case scenario of network reliability and to design transport networks with adequate capacity to deal with events which threaten its reliability. A new solution method is also developed to solve some of the proposed formulations.

The new methodologies, formulations, solution method, numerical results and insights given, significantly add to the knowledge currently available in the field. Before outlining the detailed objectives of each chapter of this thesis, a brief review of previous work in the field is provided.

1.2 LITERATURE REVIEW

The broad objectives of this thesis are to develop game theoretic formulations and a new solution method for transport network reliability and in addition, to propose as a result reliable network design models. The purpose of this review is to provide an outline of the most relevant work in the areas relating to these models. These areas include: risk-averse traffic assignment; transport network reliability approaches (and in particular the game theory approach); transport network design/reliable network design; and the solution methods used for some of these models. A number of issues are highlighted to outline the reasons for developing the models and solution method proposed in this thesis.

1.2.1 Risk-Averse User Equilibrium Traffic Assignment

1.2.1.1 Risk-Taking Behaviour under Uncertainty

A major factor in the determination and improvement of transport network reliability is a good understanding and representation of the complex interactions between travel behaviour and network performance, in particular the response of travellers to uncertainty or adverse situations, (Cassir et al., 2001). Disruptions to the road network, for example day-to-day disturbances, such as road repairs or indeed with events of even greater magnitude and impact such as natural disasters, degrade its performance and cause the network to be less reliable by reducing its link capacities and thus increasing travel times. Since degradations of the network are perceived as an added risk or increased cost to the user, the reliability of a particular route is expected to play an important role in influencing the travellers' route choice behaviour. Indeed, Jackson and Juncker (1981) identified reliability-related attributes to be some of the most important service attributes in a variety of situations. In a recent survey in the Los Angeles area (Abdel-Aty et al., 1997), travel time reliability was found to be one of the most important factors for route choices. About

54% of the respondents in the survey indicated that travel time reliability is either the most or second most important factor for choosing their primary commuting routes.

Faced with uncertainty about the travel times in their route choice, transport network users are required to make a trade-off between the travel cost (including travel time, early or late arrival penalty, and so on) and its uncertainty. This behaviour is known as risk taking behaviour (Yin et al., 2004). Many studies (e.g. Mirchandani and Soroush, 1987; Uchida and Iida, 1993; Boyce et al., 1998; Chen and Recker, 2000; Yin and Iida, 2001; Bell and Cassir, 2002; Yin et al., 2002; Lo and Tung, 2003; Chan and Lam, 2005; Sumalee et al., 2005) have been performed to model this risk-taking behaviour of route choices. One approach is the game theoretic approach (e.g. Bell, 2000, and Bell and Cassir, 2002).

1.2.1.2 Risk-Averse User Equilibrium Traffic Assignment

This game theoretic approach was employed by Bell and Cassir (2002) to model user equilibrium traffic assignment, known as risk-averse user equilibrium traffic assignment. This approach differs from the conventional approach to the incorporation of risk-averseness in traffic assignment in which safety margins are derived from the standard deviations of link costs. The size of the optimal safety margin is proportional to the standard deviation of link cost, where the constant of proportionality represents risk-averseness. A higher value of this constant implies a higher risk-averseness. Uchida and Iida (1993) provide a review of previous work in this area.

Bell and Cassir (2002) take quite a different approach. In their approach each network user is assumed to ‘play through’ all the things that can go wrong and consider his response in each case. The user is pessimistic and assumes that there is malevolence at work as opposed to assuming that things go wrong randomly. Bell (2000) envisaged a two-player, zero-sum, non-cooperative game between a network user and a fictitious ‘evil entity’ or ‘demon’ to model this. This game assumes that the network user seeks to minimise his trip cost whereas the demon seeks the opposite, that is, to maximise trip cost. Since the game is non-cooperative, the network user does not know which link state (damaged or undamaged) will be invoked by the demon and conversely the demon does not know which path the user will use. At the mixed-strategy Nash equilibrium, link choice probabilities are found which minimise expected trip costs with respect to link state probabilities, which in

turn maximise expected trip costs. According to Bell (2000), this equilibrium describes the outcome of a mental game played by a rational network user who takes a pessimistic view about link costs.

As Bell (2000) considered only one network user, the reactions of other users were not allowed for and the link costs were treated as constant. Bell and Cassir (2002) note that a pessimistic estimate of travel cost must take into account the reactions of other network users. Thus, they extended their approach leading to a $n + m$ player, non-cooperative game where n users seek their best routes and m origin-destination (OD)-specific demons penalise the network users as much as possible by damaging links in the network. The OD demon damages (reduces the link capacity) of each origin-destination pair separately. The formulation contains a series of programming problems solved for each origin-destination. Bell and Cassir (2002) show that the mixed-strategy Nash equilibrium of this problem describes a risk-averse user equilibrium traffic assignment.

The algorithm used to solve the formulations in Bell and Cassir (2002) is based on the Method of Successive Averages (MSA), (see Sheffi, 1985). This heuristic method adopted for solving the risk-averse user equilibrium assignment in Bell and Cassir (2002) without guaranteeing convergence, solves the user and demon problems iteratively until a mutually consistent solution is found.

The Bell and Cassir (2002) formulation is the starting point of this thesis. There are a number of limitations and problems with the modelling approach and solution method used to solve the Bell and Cassir (2002) formulation. These issues are examined in greater detail in sections 1.2.2.5 and 1.2.5.

1.2.2 Transport Network Reliability

The ability of society to function efficiently depends to a large degree on networks of many kinds including water supply, energy supply, communication, and in particular transportation. During particularly disruptive events the importance of the transportation network is best appreciated, as evidenced by the problems in New Orleans following 'Hurricane Katrina'. Natural or man-made events are of course only at one end of the scale of disruptive events. At the other end, network users are likely to face more routine

disruptive events such as heavy congestion, road repairs, inclement weather and traffic accidents. The effect of these adverse events on the transportation network has given rise to a whole field of study, broadly known as transport network reliability. Already much attention has been devoted to this subject including numerous symposia, books and special issues of journals (see for example: Bell and Iida, 1997, 2003; Bell and Cassir, 2001; Sumalee and Kurauchi, 2006).

From these studies it is clear that in the past decade the reliability of transport networks has become an increasingly important issue as improvements to the quality of life around the world lead to increases in the value of time. Consequently road networks are now required to be highly reliable to accommodate unexpected delays due to disruptions such as adverse weather, traffic accidents, breakdowns, signal failures, road-works, and landslides. According to Bell and Iida (1997), the initial impetus appears to have derived from the study of major natural events such as earthquakes which affect the connectivity of the road network. However, even a connected network may fail to provide an adequate level of service since disruptive events may cause unacceptable variations in the travel times faced by network users. This dimension of reliability is generally referred to in the literature as network performance reliability. A number of modelling techniques have been proposed to quantify the impact of variable network performance. Following the reviews of network reliability carried out by Nicholson et al. (2003) and Clark and Watling (2005), five broad classes of reliability methods/analysis techniques are identified and a brief review of four of these methods is provided here while a more in-depth review is provided for the modelling approach (game theory approach) adopted and further developed in this thesis.

1.2.2.1 Connectivity/Terminal Reliability Methods

In connectivity reliability methods (Bell and Iida, 1997; Asakura et al., 2001) each link of a network is assumed to have an independent, probabilistic, binary mode of operation. According to Clark and Watling (2005), this binary mode may be open/closed, or may more generally reflect some subjective definition of the successful function of a link, such as the flow to capacity ratio being less than some given value. The objective is to compute the probability that a particular path or Origin-Destination (OD) movement will be 'connected', or more generally, will 'function' as desired.

1.2.2.2 Travel Time Reliability Methods

Travel time reliability methods have received a lot of attention in the literature (see for example: Asakura and Kashiwadani, 1991; Asakura, 1996; Du and Nicholson, 1997; Bell et al., 1999; Yang et al., 2000; Clark and Watling, 2005). In general, these efforts make a continuous probabilistic treatment of link/path travel times. For example, Asakura and Kashiwadani (1991) propose a simulation-based method whereby an OD matrix is sampled and an equilibrium assignment performed for each sampled demand with the purpose of examining the impact of variability in OD demand levels. Bell et al. (1999) follow a similar approach but reduce computation time by employing equilibrium sensitivity analysis. Du and Nicholson (1997) adopt a similar approach to Bell et al. (1999) to a multi-modal equilibrium model to examine the sensitivity of their performance measure (system surplus) to unreliable events such as capacity degradation.

1.2.2.3 Capacity Reliability Methods

The third class encompasses methods to study capacity reliability, i.e., estimating the number of users who are unable to travel on a link because of the capacity restraint (for example: Chen et al., 2000, 2002). Chen et al. (2000) determine the maximum global OD multiplier such that the resulting link flows are assigned within their respective link capacities. They also allow for risk-taking behaviour in the route choice assignment model. Chen et al. (2002) employ sensitivity analysis to estimate the impact of variations in link capacities on equilibrium flows. They also extend this approach by mixing it with Monte Carlo simulation, in order to estimate sensitivities particularly when correlated link capacities are assumed.

1.2.2.4 Behavioural Reliability Methods

Clark and Watling (2005) classify this method whereby an effect on mean network performance is presumed to arise from the modified, mean behaviour of users in their attitude to the unpredictable variation and/or the 'risks' perceived. Various studies have been carried out, under an equilibrium framework, to assess the impact on the 'typical' route choice pattern (see for example: Lo and Tung, 2000; Gordon et al., 2001; Liu et al.,

2002). Other studies consider the impact on departure time choice (see for example: Noland et al., 1998; Noland and Polak, 2002).

1.2.2.5 Potential Reliability/Pessimistic Methods

These methods aim at examining the potential reliability of a network rather than aiming to model performance based on some defined probabilistic model. These are known also as ‘pessimistic methods’ since the aim is to identify potential weak points/problems and their effect. Berdica (2001, 2002) proposed various simple tests of network vulnerability, to examine the impact on various output measures (in equilibrium) of changes to the input variables of a network model. Berdica (2000) defines ‘vulnerability’ as the susceptibility to incidents that result in route closures and it increases as the probability and/or consequence of failing to meet user expectations increases. Vulnerability is essentially the opposite of reliability in this case as it increases as reliability decreases. This definition of vulnerability encompasses both probability and consequence.

D’Este and Taylor (2003) also considered ‘vulnerability’ as being appropriate for application to regional networks. They define a network as being vulnerable if degradation of a small number of network links (even a single link) causes significant reduction in network performance. Their study proposes a framework to establish where the network is most vulnerable, the degree of vulnerability and if a link or overall accessibility is vulnerable to degradation. They proposed a heuristic technique to identify potential vulnerability based on logit-based multi-path assignment.

Bell (2000) and Bell and Cassir (2002) adopt quite a different approach by supposing that the link performance probabilities arose from a ‘game’ between the users and an ‘evil entity’, suggesting they could be used as a cautious basis for network design when users are pessimistic about the performance. This approach, known as the game theory approach, is the approach adopted in this thesis and is now further reviewed in the paragraphs that follow.

The theory of games, initially developed by von Neumann, is a study of the interaction between rational decision makers. Game theory has been applied in many contexts and the field of transportation research is no different. Hollander and Prashker (2006) provide an in-depth review of non-cooperative game theory applications in the field of transport. They classified the games into four groups: games against a demon, between travellers, between authorities, and between travellers and authorities. Games against a demon are zero-sum games, in which a fictitious demon or 'evil entity' (sometimes also referred to in the literature as a 'spoiler' or 'network tester') seeks to maximise user costs. According to Bell et al. (2008), the solution describes optimal network usage from the perspective of a pessimistic user, highlighting points of vulnerability in the network.

As mentioned above, the game theoretic approach was first proposed in Bell (2000) as an approach to assess the performance reliability of transport networks. This approach uses the network cost resulting from a two-player zero-sum non-cooperative game as a measure of the performance reliability. Bell and Cassir (2002) extended this game concept to a multiplayer non-cooperative game to model the risk-averse behaviour in route choice assignment with fixed demand.

Advantages of the Game Theory Approach

The game theoretic approach enjoys at least three advantages over the more conventional approaches according to Bell (2000) and Bell and Cassir (2002). First, the game theoretic approach can determine links where the network users are most vulnerable to link failure under the assumption that the users are highly pessimistic or naturally very negative about the state of the road network. Second, the total expected network cost of the game gives a useful measure of network reliability. Different designs may be compared on the basis of this measure. Third, statistical distributions for link performance (such as delay, travel time or capacity) are not required, unlike the more conventional approaches. This information is very often either absent or not accurate enough to be used.

The Main Concept behind the Game Theory Approach

The main idea behind the game theoretic approach is based on the notion of a fictitious game played between on the one hand users who seek minimum cost routes, and on the other hand, an evil entity or OD-specific demon that seeks to maximise the total expected trip cost to the users by damaging links on each OD pair in the network. The traffic assignment problem when formulated based on this approach, known as risk-averse traffic assignment, consists of two sub-problems: the user problem and the demon problem. The user problem describes the non-cooperative behaviour of network users, whereas the demon problem describes the evil behaviour in the sense of trying to cause maximum damage to the users. The resulting expected trip cost at equilibrium can be treated as a measure of the overall network reliability while the link failure probabilities indicate links that are critical for network performance.

Expected Trip Cost as a Measure of Overall Network Reliability

Bell's games (Bell, 2000; Bell and Cassir, 2002) explain the demon's motivation plausibly by noting that the solution (i.e., the equilibrium expected trip cost) gives a pessimistic perspective which can be used as a reliability indicator. At equilibrium, the network users can no longer lower their expected trip costs by changing routes while the demon can no longer increase the expected trip costs faced by the user by damaging different links in the network. According to Bell (2000), one would expect that the network performance in the sense of expected trip cost (as opposed to realised trip cost) to deteriorate with increasing pessimism about the state of the network. This leads to the following intuitive definition of network reliability (from Bell, 2000):

Definition of Network Reliability: *A network is reliable if the expected trip costs are acceptable even if users are extremely pessimistic about the state of the road network.*

Bell (2000) proposes that the mixed strategy Nash equilibrium for his proposed game (reviewed above) offers a useful measure of the overall network reliability since it yields the expected trip cost when the user is extremely pessimistic about the state of the road network. Hollander and Prashker (2006) offer the following explanation:

'Since the driver cannot anticipate the decision of the evil entity, the equilibrium cost is the cost he/she will be willing to compromise on under the assumption of maximum obstructions. In an unreliable system the driver will compromise on a higher cost, hence the Nash equilibrium of the game is suggested as a useful measure of the network reliability'.

Therefore, the expected trip cost can be treated as a measure of the overall network reliability, while the link failure probabilities indicate links that are critical for network performance. Normative measures of network reliability can then be defined based on outputs from the game theory approach.

Measures of Network Reliability

A range of reliability measures have also been proposed to evaluate reliability impacts. One important advantage of the game theory approach (mentioned above) is that it can provide *normative* measures of network reliability. Such normative measures are useful in pointing out the most vulnerable/critical nodes (either origins or destinations), paths, and OD pairs in terms of travel time and encountering disruptions. Different designs may be compared on the basis of these measures. Many researchers have developed various reliability measures or indicators to assess the network reliability, including *encountered reliability* and *travel time reliability*. Bell and Schmöcker (2002) define encountered reliability as the probability of not encountering link degradations on the path with least perceived costs whereas according to Asakura and Ksahiwadini (1991), travel time reliability is commonly defined as the probability that a trip can be made successfully within a given acceptable time interval. Travel time reliability receives much more attention in the literature (e.g., Chen et al., 2002; Chen et al., 2007; Liu et al., 2004; Nie and Fan., 2006; Shao et al., 2008; Sumalee et al., 2006) than encountered reliability possibly because the notion of the former was proposed much earlier.

Both of these measures can be studied using the game theory approach and it is important that these measures are defined formally for the game theory approach in general and in particular for the approach developed in this thesis. Furthermore, it is essential that they are defined on the path, OD and network level so that the measures can be assessed from the point of view of the users and planners.

Apart from the two measures mentioned above, it is worthwhile to briefly mention that many researchers have also developed various other reliability measures or indicators under different modelling approaches to assess the network reliability. These include: connectivity reliability (see for example, Wakabayashi and Iida, 1992); capacity reliability (see Chen et al., 1999); travel time reliability (see Asakura and Kashiwadani, 1991); Yin and Ieda, 2001); travel cost reliability (see Schmöcker et al., 2002); flow decrement reliability (see Du and Nicholson, 1997); travel demand satisfaction reliability (see Lam and Zhang, 2000); vulnerability (see Berdica, 2000; D'Este and Taylor, 2003); and risk (see Dalziell and Nicholson, 2001).

Limitations of the Current Game Theoretic Approach

The game theory approach to assessing network reliability proposed in Bell (2000) and Bell and Cassir (2002) suffers from a number of restrictive assumptions and disadvantages from a modelling perspective which may limit its potential usefulness. Briefly, these include:

- They assume that the number of users is fixed. In reality, users may also decide not to travel at all. Therefore, it is important to incorporate elastic demand into the problem where appropriate.
- The formulation developed based on their approach only allows exactly one demon per OD pair in which each OD-specific demon is not free to damage links outside of its own specific OD pair. This one demon per OD pair assumption is not general enough. For example, it does not allow two links to be damaged on one OD pair but no link to be damaged on other OD pairs. It also restricts the number of demons or links to be damaged to be equal to the number of OD pairs and does not allow the number of links to be damaged in a network to be greater or smaller than the number of OD pairs. Therefore, a more general and flexible approach is needed.
- Following on from the previous point, the capacity reduction when link damage occurs is typically 50%. In other words only two operational states are accounted for – congested and un-congested. This clearly does not model the wide range of capacity levels.

- The effect of using or relaxing the capacity reduction assumption is not well studied or understood, with particular reference to the implications for network (or performance) reliability.
- Very little insight is provided in the literature into the demon behaviour in selecting links to damage or the existence of multiple solutions.
- The formulations are path based ones that require path enumeration. This is not attractive from a computational point of view. Developing link-based formulations is important particularly when more complicated models are formulated and when larger networks are considered.
- Another assumption of the game is that the users selfishly and non-cooperatively seek to maximise their own utilities by seeking the expected minimum cost routes without regard to the overall effect this may have on other network users. This point is dealt with in greater detail in section 1.2.3.
- The formulations developed thus far are to assess the network reliability but do not provide any information on optimal and reliable network improvements. Furthermore, the formulations are static and do not contain the time-dimension to facilitate long term planning. These issues are dealt with in greater detail in section 1.2.4.
- The solution approach used may not guarantee convergence. This point is addressed in greater detail in section 1.2.5.

Related Studies Applying the Game Theory Approach

Since the probabilities of attack or the probabilities of a link failure are in general unknown, game theory has proved to be a useful tool to analyse network reliability. It has also found some other interesting applications including: risk-averse routing of hazardous materials (see for example, Bell, 2004; 2007), the measurement of vulnerability in stochastic electronic/communication networks (see for example, Bell, 2003; Satayapiwat et al., 2008) and defending VIPs against attack (see for example, Bell et al., 2008; Kanturska et al., 2009). Other relevant studies include: Cassir and Bell (2001) and Cassir et al. (2003).

More recently, a number of publications and extensions arising from this thesis have also contributed to the field and the main ones include: O'Brien and O'Mahony (2010),

O'Brien and Szeto (2008), O'Brien et al. (2007a), O'Brien et al. (2006), Szeto et al. (2006, 2007, 2009).

1.2.3 Braess' Paradox

The risk-averse assignment game assumes that the users selfishly and non-cooperatively seek to maximise their own utilities by seeking the expected minimum cost routes without regard to the overall effect this may have on other network users. This assumption in the traditional user equilibrium problem can give rise to the seemingly counter-intuitive result known as "Braess' Paradox", (Braess, 1968). This paradox describes how adding a new link to a transportation network might not improve the operation of the system in the sense of reducing the total system travel time in the network. This phenomenon has received considerable attention in the transportation and more general scientific literature and some of the past efforts include: Arnott and Small (1994); Bass (1992); Dafermos and Nagurney (1984); Fisk (1979); and Murchland (1970), among countless others. Most of the past efforts can be found in Pas and Principio (1997) or Bazzan and Klugl (2005). However, to date no study has considered the existence of analogues of Braess' paradox in risk-averse user equilibrium assignment or the resulting implications for travel time and encountered reliability measurements.

1.2.4 The Transport Network Design Problem (NDP)

The transport network design problem (NDP) is a simplified, abstract version of a challenge often faced by the agencies responsible for managing a network of roads and highways – how to choose improvements or additions to the network, in a manner that makes efficient use of resources, in order to achieve some stated objective, such as reducing traffic congestion (Davis and Sanderson, 2001). The NDP encompasses a class of optimisation problems which models this process and a substantial body of research explores the various different formulations this problem can take, along with associated solution procedures. Comprehensive reviews of the NDP have been carried out in the past and these efforts include: Magnanti and Wong (1984), Boyce (1984) and Friesz (1985). Yang and Bell (1998) and Friesz and Shah (2001) provide more recent reviews on models and solution approaches.

Traditionally, the NDP is formulated as a bi-level programming problem. In this problem, the upper level is the network planners' problem, whereas the lower level is the network users' problem, which can be described by either the trip distribution/assignment problem or the traffic assignment problem. A review of the literature reveals the many different NDP formulations arising from different combinations of the lower and upper level problems. A number of different traffic assignment problems have been considered in the lower level problem, including deterministic user equilibrium (e.g., Meng et al., 2001), stochastic user equilibrium (Davis, 1994), and probabilistic user equilibrium (e.g., Lo, 2002).

Depending on the nature of the decision variables chosen for the problem of interest, the NDP can be briefly classified into two catalogues: the Continuous Network Design Problem (CNDP) and the Discrete Network Design Problem (DNDP). The former only includes continuous decision variables whereas the latter can contain *both continuous and discrete decision variables*. Moreover, the former takes the network topology as given and is concerned with optimising the network parameters whereas the latter is concerned with the network topology in addition to optimising network parameters. Indeed, the CNDP can be considered an approximation to the DNDP when the network topology is fixed, and has received considerably more attention in the literature (e.g. Chen et al., 2003; Chen and Yang, 2004; Chiou, 2005; Chootinan et al., 2005; Connors et al., 2005; Huang and Bell, 1998; Lim et al., 2004; Meng et al., 2001, 2004; and Yang and Meng, 2000) possibly due to the computational benefits arising from using a continuous approximation to the discrete problem especially for large transport networks. Abdulaal and LeBlanc (1979) concluded from their computational results that a continuous approximate solution could be obtained at about 25% of the cost of obtaining an exact discrete solution. Despite the computational difficulty associated with the DNDP, there are still some examples in the literature, including the determination of new roadway links to be added to the network (e.g. Chen and Alfa, 1991; LeBlanc, 1975), the bus network design problem (e.g. Wan et al., 2002), and the selection of pedestrianisation location (Wu et al., 2005).

Aside from being formulated with different decision variables, the upper level problem of the NDP has also been formulated with different objective functions. Past NDP efforts have considered: total travel time/cost minimisation (e.g., Maher et al., 2001); reserve

capacity maximisation (e.g., Wong and Yang, 1997); consumer surplus maximisation (e.g., Lam et al., 2002); and multi-objective optimisation (Tzeng and Tsaur, 1997).

In general, none of these efforts, however, consider uncertainty be it in terms of travel time, network conditions, demand etc. In the past decade, several researchers have tackled the NDP under uncertainty. Yin and Ieda (2002) analyzed the NDP with uncertain travel time. Waller and Ziliaskopoulos (2001) studied the problem of a dynamic NDP with demand uncertainty. Chen et al. (2003) introduced a mean-variance model to determine optimal tolls under the build-operate-transfer scheme with uncertain demand. Sumalee et al. (2006) developed a Reliable NDP (RNDP) considering demand and route choice variability. Their model allows the probability of users' route choice to change in response to a change in the network following the concept of probit stochastic user equilibrium. Asakura et al. (2003) formulated a stochastic network design problem that provides a set of link investment patterns for the most reliable network with highest network performance under uncertain conditions. Lo and Tung (2003) developed an approach to model network performance when its link capacities are subject to stochastic degradations arising from day-to-day traffic incidents which cause travel time variability. They characterised the users' route choice behaviour in the face of uncertain travel times with the notion of probabilistic user equilibrium. Other studies on reliable/stochastic NDP include: Chootinan et al. (2005) and Chen and Yang (2004).

However, very few studies consider the users' route choice behaviour in a *damaged network*. This is due in part to the difficulty in collecting reliable behavioural data in a degraded network. One such approach which can model route choice behaviour in a damaged network without relying on previous data relating to capacity reducing incidents is the game theoretic approach. *Therefore, the road network design problem should be studied under this approach to achieve a reliable network design to cater for the worse case scenarios from a network reliability point of view.*

In addition, all of these efforts optimise the network for a certain future time without clearly defining the time dimension within the formulation over the planning horizon (this is known as the static approach to NDP). Recently, researchers have explicitly considered the time dimension in the NDP. The time scale varies depending on the approach and problem of interest. For example, Heydecker (2002) used a very small time-scale to

capture the within-day dynamics such as fluctuation of demand and departure choice of travellers. Friesz and Shah (2001) define a medium time scale to model the route switching behaviour of network users from day to day. Lo and Szeto (2004) introduced the largest time scale by dividing the planning horizon into longer periods (typically periods of one year but other assumptions are also possible). They referred to this extension as time-dependent network design (NDP-T). The time-scale considered in the NDP-T is typically in years, as compared with the second-to-second scale of traffic dynamics, or the day-to-day scale of route choice dynamics. The flexible framework of the NDP-T permits the modelling of time-dependent travel demands, changing land-use patterns, and the gradually upgraded network over the planning horizon. With this extension, one can design for the optimal project initiation time, phasing and the financial arrangements over the planning period. Examples of previous work includes: equity issues over time and across generations (Szeto and Lo, 2006), and cost recovery issues over time (Lo and Szeto, 2009), among others.

However, all of these efforts model the capacity improvement decision variables as continuous (since this can be computationally advantageous). However, the disadvantage is that it may be difficult in practice to translate the continuous solution into capacity improvements which can be implemented. In many cases this problem is exacerbated when the continuous solution determines relatively small improvements on a large number of links or many small improvements on the same link over many years. These small improvements are not realistic to implement due to the fact that road capacity is mainly determined by the number of lanes and lanes are indivisible. Therefore, these problems should be avoided in future by proposing a time-dependent formulation of the DNDP.

Furthermore, none of the above efforts consider the users' route choice behaviour in a damaged network although there are examples in the literature on NDP-T which consider uncertainties in the financial environment (Szeto and Lo, 2005) and stochastic demand (Ukkusuri and Patil, 2009). Therefore, in order to design reliable improvements to the network which are practical to implement yet capture changing travel demands over time and caters for worst case scenarios from a network reliability perspective, a time-dependent reliable formulation of the DNDP should be proposed. Moreover, the NDP-T model enables the pursuit of important considerations that are otherwise difficult, if at all possible, with the traditional timeless (static) approach. For example, the critical or

vulnerable links may change over time and this can be captured under a time-dependent approach.

The literature on the NDP is immense, so while a brief review was provided here, a number of outstanding issues were nonetheless highlighted. Briefly, potential important gaps in knowledge that have been identified include the need to propose formulations of the reliable NDP which design the network for the worse case scenarios and which considers capacity improvements which are practical to implement yet reliable in a time-dependent framework. Later in this thesis these issues will be re-visited and further discussion is also provided.

Finally, a number of recent publications arising from this thesis have also contributed to the field of NDP (see for example: O'Brien and Szeto, 2007; O'Brien et al., 2007b).

1.2.5 Solution Methods

The MSA method adopted for solving the risk-averse user equilibrium assignment in Bell and Cassir (2002) without guaranteeing convergence, solves the user and demon problems iteratively until a mutually consistent solution is found. It is important, therefore, that in comparing with their approach that a method is proposed that that one does not rely on trial and error to guess the suitable initial solutions for getting global solutions. Furthermore, it is also important that any method proposed can guarantee finite convergence to a point arbitrarily close to the global minimum (ε -convergence). One such method which can do so is the conventional alpha-based branch and bound (α BB) algorithm in Adjiman et al. (1998a), which is based on the branch and bound scheme. A quick review is now given of the branch and bound scheme and the conventional α BB method.

1.2.5.1 Branch-and-Bound Scheme

The Branch and Bound scheme starts with finding the lower and upper bounds of the root problem (the original problem with the complete feasible region). If the bounds match, then an optimal solution has been found and the procedure terminates. Otherwise, the feasible region is divided into two or more regions, which is called *branching*. These sub-problems partition the feasible region. The *bounding* step is applied to the selected sub-

problem, which calculates the lower bound and the upper bound of the sub-problem. If an optimal solution is found to a sub-problem, it is a feasible solution to the original problem, but is not necessarily globally optimal. If the lower bound for a node (or a partition) exceeds the best-known upper bound, no globally optimal solution can exist in the subspace of the feasible region represented by the node. Therefore, the node becomes *fathomed* and can be removed from consideration. The branching and bounding steps are applied to each unsolved sub-problem until all nodes have been solved or pruned, or until the bounds match within some specified threshold.

1.2.5.2 The α based Branch-and-Bound Algorithm

The α BB algorithm in Adjiman et al. (1998a) is based on a branch-and-bound scheme. This algorithm is designed to solve general non-convex but twice-differentiable minimisation problems such as the one posed in chapter 2 of this thesis, and guarantees convergence to a point arbitrarily close to the global minimum (ε -convergence). The convergence hinges on the construction of a converging sequence of upper and lower bounds towards the global minimum. The upper bound can be obtained by calculating the original objective function value at some point, whereas the lower bound can be obtained through solving the convex relaxation of the original problem, in which the latter relies on constructing a valid convex lower bounding function using the α parameter. This construction is discussed in more detail in chapter 2 of this thesis.

It is important to develop a solution algorithm along the lines of the α BB algorithm which will be tailored for the mathematical programs developed in the thesis and which will enjoy some advantages over the conventional one to solve this problem (for example, smaller values of the α parameter for the lower bounding function comparing with those obtained by calculating the interval Hessian matrix, tighter fathomed rules, and be more efficient).

Moreover, it is important that the solution method guarantees finite convergence to a point arbitrarily close to the global minimum (in contrast to the heuristic method adopted for solving the risk-averse user equilibrium assignment in Bell and Cassir (2002) without guaranteeing convergence, which solves the user and demon problems iteratively until a mutually consistent solution is found). Comparing with their approach, it is also

advantageous if the solution method is one that does not rely on trial and error to guess the suitable initial solutions for getting global solutions.

1.3 OBJECTIVES OF THE THESIS

The broad objectives of this thesis may be outlined as follows:

- To develop new models of risk-averse traffic assignment and as a result new approaches to assess the network reliability considering many more possibilities than existing approaches, including: incorporating elastic behaviour of the travel demand into the problem; new assumptions on the type and destructive action of the fictitious player or demon and the resulting form of capacity reduction that may occur; and modelling of many more operational states of the road network. These approaches capture worst-case scenarios of network reliability and identify the critical network links (the links, which when damaged, can cause the greatest amount of disruption to the users).
- To develop a solution method with guaranteed ε convergence to solve some of the proposed approaches.
- To make the proposed approaches more amenable for application to larger networks and real-life problems by formulating, where appropriate, link-based models.
- To propose a number of reliability measures based on these approaches and to demonstrate the existence of paradoxical phenomena, known as stochastic Braess' paradox and reliability paradox, in the existing and proposed approaches to risk-averse traffic assignment.
- To develop models of reliable road network design which prioritise improvements to the network links vulnerable to failure and which consider many operational states to cater for the worst-case scenarios of network reliability. In addition, time-dependent conventional and reliable network design models are proposed which can facilitate long-term planning and capacity improvements which are practical to implement.

In the next section the specific objectives and contributions of each chapter are described in greater detail.

1.4 ORGANISATION OF THE THESIS

The thesis is presented in six chapters following this chapter.

To assess the reliability of the network accurately and design a reliable transport network properly, one should consider the risk taking and elastic behaviour of travel demand. **Chapter 2** develops a Non-linear Complementarity Problem (NCP) formulation of risk-averse traffic assignment considering this behaviour. The NCP formulation is transformed to a smooth and unconstrained mathematical program via a Fischer function and then solved by the proposed modified α BB (branch and bound) algorithm, while guaranteeing finite convergence to a point arbitrarily close to the global minimum (ε -convergence). To speed up the computation, a sampling method is suggested. A small numerical study is set up to illustrate the properties and the performance of the algorithm. The results show that the bound improving strategy, the sampling size on estimating the value of α for the convex lower bounding function, and choice of α have a strong influence of the rate of convergence.

In **chapter 2** and in traditional game theoretic approaches to measuring transport network reliability, the risk-averse traffic assignment is described by a game played between network users who seek minimum cost routes and demons that seek to impose maximum costs on the network users by damaging links in the network. This problem assumes the presence of one and only one active Origin Destination (OD) specific demon in each OD pair, and furthermore assumes the capacity reduction to be 50% if the link is selected for damage by one or more OD-specific demons. In **chapter 3**, these two assumptions are relaxed and a multiple network demon formulation is proposed in which each demon is free to select any link to damage. Numerical studies are carried out to: examine the effects of relaxing the two assumptions on expected network cost, give an insight into the network demon behaviour in selecting links to damage, demonstrate the existence of multiple solutions to the proposed game, and compare the link selection behaviour of the OD-specific and network-specific demons and their impacts on expected network cost. Overall, the results indicate the importance of the assumptions used to expected network cost and reliability measures, and provide some further insights into the nature of the route choice game.

Chapter 3 proposed a more general game theoretic formulation via the NCP approach by relaxing the assumptions on the number/type of demons and the form of capacity reduction. This proposed formulation extended the existing game theoretic formulation to allow demons (called network-specific demons or simply network demons) to be free to select any links to damage and to impose no restriction on the number of demons or number of links to be damaged. The game is typically expressed as a path-based formulation, which is computationally intensive since the formulation requires path enumeration. Another assumption of the game is that the users selfishly and non-cooperatively seek to maximise their own utilities by seeking the expected minimum cost routes without regard to the overall effect this may have on other network users. This assumption in the traditional user equilibrium problem can give rise to the seemingly counter-intuitive result known as “Braess’ Paradox”(Braess, 1968). In **chapter 4**, a *link-based* multiple network demon formulation via the NCP approach is proposed where each demon is free to select any link to damage. Under this framework, this chapter examines the effects of the proposed model on total expected network cost and reliability measures, as well as demonstrating through specific examples paradoxical phenomena that if one adds a road to a network then all users may be worse off in terms of total expected network cost and/or travel time reliability. Overall the results indicate the importance of the assumptions used to total expected network cost and reliability measures, and provide some insights into the problem of ignoring these paradoxical phenomena in reliable and robust network design.

To design a reliable transport network properly it is necessary to take into account users reactions to incidents which cause disruption in different parts of the network. In order to design a reliable network to deal with such events it is necessary to capture the users’ route choice behaviour under degraded conditions. Furthermore, for practical reasons, it is imperative that any improvement scheme should give priority to links in the network which are most vulnerable to failure. Therefore, in **chapter 5**, a reliable continuous network design problem (RCNDP) is formulated, in which the link flow is assigned to different routes considering the risk-taking behaviour of the network users when confronted by network uncertainty. To capture this behaviour, this chapter poses the RCNDP under the multiple network demon formulation of risk-averse traffic assignment. The ability of this model to introduce higher capacity expansions on links vulnerable to

failure and consider many operational states to cater for the worst-case scenarios of network reliability is demonstrated through two numerical studies.

Over time there will be inevitable changes in the travel demand patterns in a transport network. To capture these changing demands and design a gradually upgraded network it is necessary to introduce the time dimension into the conventional transport network design problem (NDP). Previous efforts have extended the continuous network design problem to cater for such considerations. However, the continuous problem does not always indicate improvements that are practical. Therefore, in **chapter 6**, a model of the time-dependent discrete network design problem is developed. The importance of the model is illustrated by two numerical studies, which demonstrate the importance of introducing the time dimension to the NDP and the importance of introducing discrete decision variables to the NDP over time. An example on NDP under two cost-recovery principles is also provided to illustrate the application of the model.

Furthermore, over time transport networks are expected to remain highly reliable to ensure that users can experience smooth travel under both normal and abnormal traffic conditions. To design a reliable transport network properly it is necessary to take into account users reactions to incidents which cause disruption in different parts of the network *over time*. In order to design a reliable network to deal with such time-varying events it is necessary to capture the users' route choice behaviour under degraded conditions with the time dimension explicitly considered. Furthermore, for practical reasons, it is imperative that any improvement scheme be possible to implement. Therefore, in **chapter 6**, a time-dependent reliable discrete network design problem is formulated where the users' behaviour in route choice is modelled through the game theoretic approach. Some important aspects of this model are demonstrated through a numerical example including the design approach of the proposed model under a budget and the selection and design of the vulnerable links over time under different assumptions on the type of demons assumed and the corresponding form of link capacity reduction they may cause.

Chapter 7 is the concluding chapter of the thesis. A summary of the works presented in the thesis is given along with the conclusions drawn. In addition, a discussion of the results is provided and a critical assessment of the work is given. Finally, this chapter also offers perspectives on future research.

CHAPTER 2

Risk-Averse Traffic Assignment with Elastic Demands: NCP Formulation and Solution Method for Assessing Performance Reliability

2.1 INTRODUCTION

The importance of a reliable transportation network has been appreciated for some time now. According to Chen et al. (2002), a stable (and reliable) transportation system would provide a competitive edge in the global economy. Disruptions to the road network, for example day-to-day disturbances, such as road repairs or indeed with events of even greater magnitude and impact such as natural disasters, degrade its performance and cause the network to be less reliable by reducing its link capacities and thus increasing travel times. Since degradations of the network are perceived as an added risk or increased cost to the user, the reliability of a particular route is expected to play an important role in influencing the travellers' route choice behaviour. In chapter 1, a number of studies were highlighted which identified the importance of reliability-related attributes and which showed that travel time reliability was found to be one of the most important factors for route choices.

Faced with uncertainty about the travel times in their route choice, transport network users are required to make a trade-off between the travel cost (including travel time, early or late arrival penalty, and so on) and its uncertainty. This behaviour is known as risk taking behaviour (Yin et al., 2004). Chapter 1 listed many studies that have been performed to model this risk-taking behaviour of route choices. One approach is the game theoretic approach (e.g. Bell, 2000, and Bell and Cassir, 2002).

This approach enjoys a number of advantages over the more conventional approaches as already described in chapter 1 of this thesis. The game theoretic approach was employed by Bell and Cassir (2002) to model user equilibrium traffic assignment, known as risk-averse user equilibrium traffic assignment, but they assume that the number of users is fixed. In reality, users may also decide not to travel at all. This chapter extends Bell and

Cassir's (2002) formulation by incorporating this elastic behaviour in the problem. This proposed problem is indeed equivalent to the mixed strategy Nash equilibrium game when the number of *potential* users is large, similar to the fixed demand case. This can be easily proven by modifying the original transport network through the addition of a virtual route connecting from each origin to each destination. With this modification, the result from the fixed demand case applies.

Like the fixed demand case, the proposed risk-averse user equilibrium traffic assignment with elastic demands consists of two sub-problems: the user problem and the demon problem. The user problem describes the non-cooperative and elastic behaviour of the potential travel demand, whereas the demon problem describes the evil behaviour in the sense of trying to cause maximum damage to the users. The two sub-problems are combined and formulated as a Non-linear Complementarity Problem (NCP) in this chapter.

To solve for solutions, the NCP formulation is transformed to a smooth and unconstrained mathematical program via the smooth Fischer function, and then solved by the proposed modified α -based Branch and Bound (α BB) algorithm. The modified α BB algorithm differs from the conventional one (e.g. Maranas and Floudas, 1994a, b, and Adjiman et al., 1998 a, b) in determining lower bounds, terminating the algorithm, and defining fathomed partitions. It is tailored for this mathematical program and enjoys some advantages over the conventional one to solve this problem, such as smaller values of the α parameter for the lower bounding function comparing with those obtained by calculating the interval Hessian matrix, and tighter fathomed rules. Furthermore, it is more efficient based on limited computation experience. Nonetheless, both methods guarantee finite convergence to a point arbitrarily close to the global minimum (ε -convergence), and contrasts to the heuristic method adopted for solving the risk-averse user equilibrium assignment in Bell and Cassir (2002) without guaranteeing convergence, which solves the user and demon problems iteratively until a mutually consistent solution is found. When compared with their approach, the advantage of the proposed method is that one does not rely on trial and error to guess the suitable initial solutions for getting global solutions. This can be regarded as one of the contributions of this chapter. To speed up the computation, a sampling method is suggested to calculate the α parameter for the lower bounding function.

To illustrate the convergence rate of the algorithm, a small test network is used, while the formulation and the algorithm can be used for a general network with multiple OD pairs. The results show that the partition selecting strategy, the sampling size on estimating the value of α , and choice of α , have a strong influence on the rate of convergence.

Finally, the formulation, solution method and results provided in this chapter are based upon a paper published by this author in the journal 'Networks and Spatial Economics' (see Appendix B for further details). The rest of this chapter is structured as follows: section 2.2 describes the proposed formulation of risk-averse user equilibrium traffic assignment with elastic demands; section 2.3 provides details of the proposed solution method; section 2.4 illustrates the convergence rate of the proposed solution method through numerical experiments on a simple test network. Finally, section 2.5 reports some concluding remarks.

2.2 FORMULATION

A general transportation network with multiple links, routes, and Origin-Destination (OD) pairs is considered. The network is modified by constructing one "virtual" route to connect the origin and destination of each OD pair. There are N^{rs} homogenous players between each OD pair, each of whom decides whether to travel, and if so, by which route. Each travel player is assumed to seek the minimum cost route whereas each non-travel player is assumed to select the sole artificially constructed "virtual" route. There is also a demon intent on causing maximum delays (and hence increased costs) to the travel players by damaging one and only one link on the whole network except the virtual links. This assumption can be easily relaxed in the future but here one demon is considered just for the ease of illustration. Moreover, it is assumed that no player knows which link a demon will choose to damage, and no demon knows in advance which link a user may decide to use. Furthermore, each link is assumed to have only two costs depending on whether it is damaged or not. With this setting, the proposed risk-averse user equilibrium traffic assignment problem with elastic demands is formulated by applying the results in Nash (1951) and Bell and Cassir (2002).

The proposed problem comprises two sub-problems: The user problem and the demon problem. The user problem can be viewed as a non-cooperative game, in which each homogenous player tries to select the route (including the virtual route) with minimum expected trip cost. This user problem can be approximated to deterministic traffic assignment when the number of homogenous players is large (Bell and Cassir, 2002). The first-order condition of deterministic traffic assignment can be expressed as the following non-linear complementarity conditions:

$$h_j \left[\sum_k g_{jk}(\mathbf{h}) q_k - \sum_{rs} \delta_j^{rs} \left(\min_m \delta_m^{rs} \left(\sum_k g_{mk}(\mathbf{h}) q_k \right) \right) \right] = 0, \quad (2.1)$$

$$\sum_k g_{jk}(\mathbf{h}) q_k - \sum_{rs} \delta_j^{rs} \left(\min_m \delta_m^{rs} \left(\sum_k g_{mk}(\mathbf{h}) q_k \right) \right) \geq 0, \quad (2.2)$$

$$h_j \geq 0, \quad (2.3)$$

$$e^{rs} \left[t^{rs}(e^{rs}) - \min_m \left(\delta_m^{rs} \left(\sum_k g_{mk}(\mathbf{h}) q_k \right) \right) \right] = 0, \quad (2.4)$$

$$t^{rs}(e^{rs}) - \min_m \left(\delta_m^{rs} \left(\sum_k g_{mk}(\mathbf{h}) q_k \right) \right) \geq 0, \quad (2.5)$$

$$e^{rs} \geq 0, \quad (2.6)$$

where h_j is the flow on route j ; $\mathbf{h}$ is the route flow vector; $g_{jk}(\mathbf{h})$ represents the cost of route j in scenario k based on the flow vector $\mathbf{h}$; q_k is the probability of damage to a link in scenario k ; δ_j^{rs} is the route-OD incidence indicator. $\delta_j^{rs} = 1$ if j connects OD pair rs , and $\delta_j^{rs} = 0$ otherwise; e^{rs} is the flow on the virtual route between OD pair rs satisfying the flow conservation:

$$e^{rs} = N^{rs} - \sum_j \delta_j^{rs} h_j; \quad (2.7)$$

N^{rs} denotes the total number of homogenous network players between OD pair rs ; $t^{rs}(e^{rs})$ is an increasing virtual route travel cost function.

Conditions (2.1)-(2.3) are the non-linear complementarity conditions for the routes on the network, excluding the virtual routes. According to (2.3), flows must be non-negative, implying that route j can carry flow ($h_j > 0$) or carry no flow ($h_j = 0$). If route j connecting OD pair rs carries flow ($h_j > 0$), according to (2.1), the corresponding expected route travel cost $\sum_k g_{jk}(\mathbf{h})q_k$ must be equal to the minimum expected travel cost $\sum_{rs} \delta_j^{rs} \left(\min_m \delta_m^{rs} \left(\sum_k g_{mk}(\mathbf{h})q_k \right) \right)$ between OD pair rs , as the term in the square bracket in (2.1) must equal zero. If route j carries no flow ($h_j = 0$), the corresponding expected route travel cost must be greater than or equal to the minimum expected travel cost based on (2.2).

Similarly, conditions (2.4)-(2.6) are the non-linear complementarity conditions for the virtual routes. According to the non-negativity condition (2.6), flows on the virtual route must be non-negative. If the virtual route between OD pair rs carries flow ($e^{rs} > 0$), according to (2.4), the corresponding route travel cost $t^{rs}(e^{rs})$ must be equal to the minimum expected OD travel cost $\min_m \delta_m^{rs} \left(\sum_k g_{mk}(\mathbf{h})q_k \right)$. If the virtual route carries no flow ($e^{rs} = 0$), the corresponding route travel cost must be greater than or equal to the minimum expected OD travel cost by the definition of (2.5).

The cost of route j in scenario k , $g_{jk}(\mathbf{h})$, appeared in (2.1), (2.2), (2.4), and (2.5) is the total cost on every link that is on this route:

$$g_{jk}(\mathbf{h}) = \sum_a \delta_{ja} t_{ak}(v_a), \quad (2.8)$$

where $t_{ak}(v_a)$ denotes the flow-dependent cost on link a in scenario k ; δ_{ja} is 1 if route j uses link a , and 0 otherwise. The link flow v_a , in (2.8) is given by:

$$v_a = \sum_j \delta_{ja} h_j. \quad (2.9)$$

The above equation states that the link flow is simply the sum of route flows using that link.

The demon problem depicts the objective of the demon who seeks a mixed strategy to maximize the total expected trip cost to the users. The problem can be written as the non-linear complementarity conditions:

$$q_k \left(\max_w \sum_j g_{jw}(\mathbf{h}) h_j - \sum_j g_{jk}(\mathbf{h}) h_j \right) = 0, \quad (2.10)$$

$$\max_w \sum_j g_{jw}(\mathbf{h}) h_j - \sum_j g_{jk}(\mathbf{h}) h_j \geq 0, \quad (2.11)$$

$$q_k \geq 0, \quad (2.12)$$

where the probability q_k must satisfy:

$$\sum_k q_k - 1 = 0. \quad (2.13)$$

and be non-negative as shown in (2.12) by definition, and $g_{jk}(\mathbf{h})$ follows the definition in (2.8).

Conditions (2.10)-(2.11) actually represent the necessary and sufficient conditions of the Nash equilibrium of a non-cooperative, mixed-strategy game for the demon (see Nash 1951). According to (2.10), if scenario k is selected ($q_k > 0$), the maximum total expected network cost, $\max_w \sum_j g_{jw}(\mathbf{h}) h_j$, must be equal to the total expected network cost in

scenario k , $\sum_j g_{jk}(\mathbf{h}) h_j$, through the condition $\left(\max_w \sum_j g_{jw}(\mathbf{h}) h_j - \sum_j g_{jk}(\mathbf{h}) h_j \right) = 0$. If scenario k is not selected ($q_k = 0$), the maximum total expected network cost must be greater than or equal to the total expected network cost in scenario k according to (2.11).

The risk-averse traffic assignment equilibrium with elastic demands is to find q_k, h_j such that (2.1)-(2.13) are satisfied. This problem can be re-expressed as the Non-Linear Complementarity Problem (NCP). Since the proposed formulation considers elastic

demand, $e^{rs} = 0$ (the fixed demand case) and $e^{rs} = N^{rs}$ (the case of no one deciding to travel) are not considered here. The flow on the virtual route e^{rs} is thus greater than zero, which leads to $t^{rs}(e^{rs}) - \min_m \delta_m^{rs} \left(\sum_k g_{mk}(\mathbf{h}) q_k \right) = 0$ or

$$\min_m \delta_m^{rs} \left(\sum_k g_{mk}(\mathbf{h}) q_k \right) = t^{rs}(e^{rs}). \quad (2.14)$$

Substituting (2.14) into (2.1) and (2.2) gives:

$$h_j \left[\sum_k g_{jk}(\mathbf{h}) q_k - \sum_{rs} \delta_j^{rs} t^{rs}(e^{rs}) \right] = 0, \text{ and} \quad (2.15)$$

$$\sum_k g_{jk}(\mathbf{h}) q_k - \sum_{rs} \delta_j^{rs} t^{rs}(e^{rs}) \geq 0. \quad (2.16)$$

To deal with (2.10), (2.11) and (2.13), first let

$$\pi = \max_w \sum_j g_{jw}(\mathbf{h}) h_j. \quad (2.17)$$

Then, (2.10)-(2.11) can be expressed as:

$$q_k \left(\pi - \sum_j g_{jk}(\mathbf{h}) h_j \right) = 0, \text{ and} \quad (2.18)$$

$$\pi - \sum_j g_{jk}(\mathbf{h}) h_j \geq 0. \quad (2.19)$$

Next, both sides of equation (2.13) are multiplied by π giving:

$$\pi \left(\sum_k q_k - 1 \right) = 0. \quad (2.20)$$

As link travel cost must be greater than zero, the maximum total travel cost must be greater than zero, and the term inside the bracket in (2.20) must be equal to zero at optimality. Equality (2.13) must therefore hold at optimality. Adding (2.20) to the problem (2.1)-(2.13) will not alter the optimality condition. For mathematical completeness, two more

conditions are also introduced to the original problem:

$$\pi \geq 0, \quad (2.21)$$

$$\sum_k q_k - 1 \geq 0. \quad (2.22)$$

They include $\pi > 0$ and $\sum_k q_k - 1 = 0$ as special cases and consequently do not change the optimality condition of the original problem (2.1)-(2.13).

The risk-averse traffic assignment equilibrium with elastic demands is thus to find q_k, h_j, π such that (2.3)-(2.9), (2.12), (2.15)-(2.16), and (2.18)-(2.22) are satisfied, and can be expressed as a NCP: to find $\mathbf{x}^* \geq \mathbf{0}$ such that:

$$\mathbf{F}(\mathbf{x}^*) \geq \mathbf{0}, \text{ and} \quad (2.23)$$

$$\mathbf{x}^{*T} \cdot \mathbf{F}(\mathbf{x}^*) = 0, \quad (2.24)$$

where the asterisk associated with the variable refers to the optimal solution; $\mathbf{x} = \begin{bmatrix} h_j, \forall j \\ q_k, \forall k \\ \pi \end{bmatrix}$;

$$\mathbf{F}(\mathbf{x}) = \begin{bmatrix} \sum_k g_{jk}(\mathbf{h}) q_k - \sum_{rs} \delta_j^{rs} t^{rs}(e^{rs}) \\ \pi - \sum_j g_{jk}(\mathbf{h}) h_j, \forall k \\ \sum_k q_k - 1 \end{bmatrix}; \text{ the cost on route } j, g_{jk}(\mathbf{h}), \text{ is defined by (2.8)-}$$

(2.9), and e^{rs} is defined by the flow conservation (2.7). To solve the above NCP efficiently by the existing local optimisation methods, $\mathbf{x}$ and $\mathbf{F}(\mathbf{x})$ in (2.23)-(2.24) are redefined though rescaling as follows:

$$\mathbf{x} = \begin{bmatrix} h'_j, \forall j \\ q_k, \forall k \\ \pi' \end{bmatrix}, \text{ and;} \quad (2.25)$$

$$\mathbf{F}(\mathbf{x}) = \begin{bmatrix} \sum_k g_{jk}(\mathbf{h}) q_k - \sum_{rs} \delta_j^{rs} t^{rs}(e^{rs}) \\ \pi' - \frac{\sum_j g_{jk}(\mathbf{h}) h_j}{\pi_{\max}}, \forall k \\ \sum_k q_k - 1 \end{bmatrix}, \quad (2.26)$$

where $h'_j = \frac{h_j}{\sum_{rs} \delta_j^{rs} N^{rs}}$; $\pi' = \frac{\pi}{\pi_{\max}}$; $\pi_{\max}$ is the upper bound of the maximum expected total

network trip cost; $g_{jk}(\mathbf{h})$ is defined by (2.8)-(2.9), and e^{rs} is defined by (2.7).

The NCP (2.23)-(2.26) can be transformed to a minimisation problem via a gap function. Let Ω be the solution set to the NCP formulation and $\psi = \{\mathbf{x} \geq \mathbf{0}, \mathbf{F}(\mathbf{x}) \geq \mathbf{0}\}$. A function $G(\mathbf{x})$ is the Gap function for the NCP if the following three conditions are satisfied:

- (i) $G(\mathbf{x}) \geq 0$;
- (ii) $G(\mathbf{x}) = 0 \Leftrightarrow \mathbf{x} \in \Omega$, and;
- (iii) $\min_{\mathbf{x} \in \psi} G(\mathbf{x}) = 0$ is a global minimum.

These properties mean that if $G(\mathbf{x})$ can be found to be zero, then one has a solution to the NCP. On the other hand, if the NCP has no solution, then $G(\mathbf{x})$ must have global solutions with positive objective values (Theorem 3.1 in Kanzow and Fukushima, 1996). In essence, the gap function provides a convenient measure of convergence for developing solution algorithms.

In this chapter, the one in Lo and Chen (2000) is adopted:

$$G(x) = \sum_i \varphi(x_i, F_i(x)), \quad (2.27)$$

where

$$\varphi(a, b) = \frac{1}{2} \varphi^2(a, b), \text{ and} \quad (2.28)$$

$\phi(a, b)$ is the *Fischer Function* defined as follows:

$$\phi(a, b) = \sqrt{(a^2 + b^2)} - (a + b). \quad (2.29)$$

One property of the gap function (2.27) is that (2.27) is *continuously differentiable* and unconstrained (Lo and Chen, 2000) which permits algorithms that use the gradient and the second derivatives of $G(\mathbf{x})$ for solutions. This property together with property (iii) allows one to develop the proposed method based on the conventional α BB algorithm to solve for global solutions from any initial feasible solutions.

By adopting the gap function (2.27), the NCP (2.23)-(2.26) can be reformulated into a minimisation program:

$$\min_{\mathbf{x}} G(\mathbf{x}), \quad (2.30)$$

where $G(\mathbf{x})$ is defined by (2.27)-(2.29).

2.3 THE MODIFIED α -BASED BRANCH-AND-BOUND (α BB) SOLUTION METHOD

To solve the minimisation program (2.30), this chapter develops a solution method based on a modification of the conventional α BB in Adjiman et al. (1998a). Their similarities and differences are summarised in the following two subsections.

2.3.1 Similarities of the Modified and Conventional α BB Algorithms

Both α BB methods which are based on a branch-and-bound scheme, are designed to solve general non-convex but twice-differentiable minimisation problems such as the one posed in this chapter, and guarantee convergence to a point arbitrarily close to the global minimum (ε -convergence). The convergence hinges on the construction of a converging sequence of upper and lower bounds towards the global minimum. The upper bound can be obtained by calculating the original objective function value at some point, whereas the lower bound can be obtained through solving the convex relaxation of the original

problem, in which the latter relies on constructing a valid convex lower bounding function $L(\mathbf{x})$.

This convex lower bounding function, as illustrated in figure 2.1, can be obtained by augmenting the objective function $G(\mathbf{x})$ using the same principles proposed by Maranas and Floudas (1992, 1993):

$$L(\mathbf{x}) = G(\mathbf{x}) + \alpha \sum_{i=1}^n (x_i^{LB} - x_i)(x_i^{UB} - x_i), \quad (2.31)$$

where x_i^{UB} and x_i^{LB} are the upper and lower bounds on variable x_i , and α is defined by:

$$\alpha \geq \max \left\{ 0, -\frac{1}{2} \min_{i, x_i^{LB} \leq x_i \leq x_i^{UB}} \lambda_i(\mathbf{x}) \right\}, \quad (2.32)$$

where the $\lambda_i(\mathbf{x})$'s are the eigenvalues of $H_G(\mathbf{x})$, the Hessian matrix of the function $G(\mathbf{x})$. According to (2.32), α is a non-negative parameter which must be greater than or equal to the negative of one half of the minimum eigenvalue of the Hessian matrix over a region. Its magnitude greatly influences the convergence rate of both algorithms, and can be estimated via a number of methods, including solving an optimisation problem with the use of the concept of the measure of a matrix (e.g. Maranas and Floudas, 1994b), and calculating the minimum eigenvalue interval Hessian matrix derived from natural interval extension (e.g. Adjiman et al., 1998a).

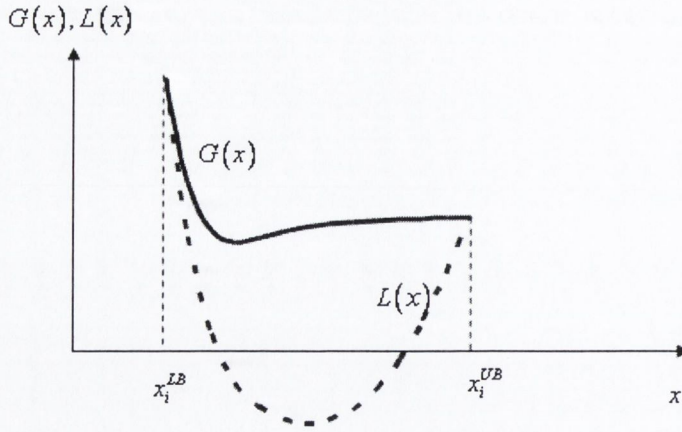


Figure 2.1: The convex lower bounding function, $L(x)$, of the objective function, $G(x)$ in the one-dimension case

Comparing with $G(\mathbf{x})$, $L(\mathbf{x})$ in (2.31) is made convex by adding the following term to $G(\mathbf{x})$:

$$\alpha \sum_{i=1}^n (x_i^{LB} - x_i)(x_i^{UB} - x_i). \quad (2.33)$$

This has the effect of overpowering the non-convexity characteristics of $G(\mathbf{x})$ with the addition of the term 2α to all of its eigenvalues,

$$\lambda_i^L = \lambda_i^{G(\mathbf{x})} + 2\alpha, \quad (2.34)$$

where, λ_i^L and $\lambda_i^{G(\mathbf{x})}$ represent the i^{th} eigenvalues of $L(\mathbf{x})$ and $G(\mathbf{x})$ respectively. This convex underestimator $L(\mathbf{x})$ has five main properties that help both conventional and modified α BB algorithms in locating the global solution. These properties, whose proofs can be found in Maranas and Floudas (1994b), are summarised below:

1. $L(\mathbf{x})$ is always a valid under-estimator of $G(\mathbf{x})$ inside the box constraints (i.e. $\forall \mathbf{x} \in [\mathbf{x}^{LB}, \mathbf{x}^{UB}], L(\mathbf{x}) \leq G(\mathbf{x})$).
2. $L(\mathbf{x})$ matches $G(\mathbf{x})$ at all corner points.

3. $L(\mathbf{x})$ is convex in $[\mathbf{x}^{LB}, \mathbf{x}^{UB}]$.
4. The maximum separation between $L(\mathbf{x})$ and $G(\mathbf{x})$ is bounded and proportional to α and to the square of the diagonal of the current box constraints.
5. The under-estimators constructed over supersets of the current sets are always less tight than the under-estimator constructed over the current box constraints for every point within current box constraints. Therefore the values of different definitions of $L(\mathbf{x})$ at any point define a non-decreasing sequence, provided that $L(\mathbf{x})$ is constructed over a tighter box of constraints each time.

Based on properties (1) and (3), $L(\mathbf{x})$ is a convex lower bounding function of $G(\mathbf{x})$ in which $L(\mathbf{x})$ equals $G(\mathbf{x})$ at all corner points (property 2). Property (4) defines what the maximum gap between the upper and lower bounds of $G(\mathbf{x})$ is, which can be used as a stopping criterion. Property (5) is used to refine the lower and upper bounds on the global minimum in each iteration. It implies that the value of $L(\mathbf{x})$ at every point, and therefore at its global minimum, is increased by restricting the box constraints within which it has been defined.

2.3.2 The Key Differences between the Modified and the Conventional α BB Algorithms

Despite the above similarities, the modified α BB algorithm differs from the conventional one in the terminating criterion for the algorithm, the rule to fathom partitions, and the bound improving strategy.

2.3.2.1 Terminating Criterion for the Algorithm

The gap value is used to measure how close to the optimal solution the current solution is and to determine when to terminate the algorithm. More specifically, if $G(\mathbf{x})^{UB} \leq \varepsilon$, where ε is the tolerance, then the solution procedure is stopped. This terminating criterion differs from the conventional one, which stops the algorithm based on the difference between the

upper and lower bounds (property 4 of the lower bounding function). Indeed, this stopping criterion is more restrictive than $\left(G(\mathbf{x})^{UB} - G(\mathbf{x})^{LB}\right) \leq \varepsilon$ as $G(\mathbf{x})^{LB}$ is negative.

2.3.2.2 The Rule to Fathom Partitions

If the lower bound is greater than zero (i.e. $L(\mathbf{x}) > 0$), the corresponding partition becomes fathomed and can be removed from consideration. The reason is as follows: Since it is known that $L(\mathbf{x}) > 0$ and zero must be the global minimum according to property (iii) of the gap function) (i.e., $0 = \min_{\mathbf{x} \in \mathcal{V}} G(\mathbf{x})$), one has $L(\mathbf{x}) > 0 = \min_{\mathbf{x} \in \mathcal{V}} G(\mathbf{x})$ or $L(\mathbf{x}) > \min_{\mathbf{x} \in \mathcal{V}} G(\mathbf{x})$. The last condition together with $G(\mathbf{x}) \geq L(\mathbf{x})$ (property 1 of the lower bounding function) gives $G(\mathbf{x}) \geq L(\mathbf{x}) > 0 = \min_{\mathbf{x} \in \mathcal{V}} G(\mathbf{x})$ or $G(\mathbf{x}) > \min_{\mathbf{x} \in \mathcal{V}} G(\mathbf{x})$, which means that the partition does not contain the global solution and therefore can be removed. This fathoming rule is more restrictive than the rule to fathom the partition when the lower bound is greater than the upper bound as in the conventional α BB algorithm, and can speed up the computation process theoretically.

2.3.2.3 Bound Improving Strategy

At each iteration, only the sub-rectangle responsible for the *supremum* (not *infimum* as in the conventional α BB algorithm) of all the minima of $L(\mathbf{x})$ over all *unfathomed* sub-rectangles is halved. This proposed strategy can still generate a non-decreasing sequence of the lower bound, provide a tighter bound at each iteration, and speed up the solution process. Note that generating a tighter non-decreasing sequence of the lower bound is possible here since the global minimum is known. Without knowing the global minimum for the general case, the proposed bounding strategy is usually not able to generate a non-decreasing sequence of the lower bound, thereby causing inefficiency.

If it is desired to further improve the computation speed of the modified algorithm, then α can be determined using the proposed heuristic sampling method discussed below.

2.3.3 Heuristic Sampling Method to Calculate α

Within the sub-rectangle, N points are selected. At each point, the Hessian and the corresponding eigenvalues are calculated. The minimum eigenvalue is calculated from those eigenvalues, and α is obtained by:

$$\alpha = \max \left\{ 0, -\frac{1}{2} \min \lambda_i(\mathbf{x}_j), \forall i, j \right\}. \quad (2.35)$$

From limited computation experience here, this sampling method is more efficient than the approach used in the literature, but the resulting lower bounding function may not be convex because the sampling method may underestimate α in (2.32).

This sampling method was incorporated in the modified α BB algorithm to determine the solutions in most studied scenarios in section 2.4. The basic steps of this approach are depicted in the next subsection.

2.3.4 Steps of the Proposed Method Incorporating Sampling

Step 1: Initialisation. A convergence tolerance ε is chosen and the iteration counter, $Iter$, is set to one. Appropriate global bounds x_i^{LB}, x_i^{UB} on x_i are chosen and local bounds $x_i^{L, Iter}, x_i^{U, Iter}$ for the first iteration are set to be equal to the global ones. An initial current point $\mathbf{x}^c = [x_i^{c, Iter}, \forall i]$ is selected by $x_i^{c, Iter} = \frac{x_i^{LB} + x_i^{UB}}{2}$. The lower bound on the global minimum $G(\mathbf{x})^{LB}$ is initialised (set to be a very large negative number), whereas the upper bound $G(\mathbf{x})^{UB}$ is set to be $G(\mathbf{x}^c)^{UB}$. Set the sample size N .

Step 2: Convergence checking. If $G(\mathbf{x})^{UB} \leq \varepsilon$ or $(G(\mathbf{x})^{UB} - G(\mathbf{x})^{LB}) \leq \varepsilon$, then stop, and the global minimum solution, and the solution points are:

$$G(\mathbf{x})^* = G(\mathbf{x})^{UB, Iter^*}, \text{ and;}$$

$$x_i^* = x_i^{c, Iter^*}, i = 1, \dots, n,$$

where:

$$Iter'' = \arg \left\{ G(\mathbf{x})^{UB,I} = G(\mathbf{x})^{UB} \right\}, I = 1, \dots, Iter'.$$

Step 3: *Partitioning the current rectangle and selecting a branching variable:* The current rectangle $[x_i^{L,Iter}, x_i^{U,Iter}]$, $i = 1, \dots, n$, is partitioned into the following two sub-rectangles ($r = 1, 2$):

$$\begin{bmatrix} x_1^{LB,Iter} & x_1^{UB,Iter} \\ \vdots & \vdots \\ x_{l^{Iter}}^{LB,Iter} & \frac{(x_{l^{Iter}}^{LB,Iter} + x_{l^{Iter}}^{UB,Iter})}{2} \\ \vdots & \vdots \\ x_n^{LB,Iter} & x_n^{UB,Iter} \end{bmatrix} \text{ and } \begin{bmatrix} x_1^{LB,Iter} & x_1^{UB,Iter} \\ \vdots & \vdots \\ \frac{(x_{l^{Iter}}^{LB,Iter} + x_{l^{Iter}}^{UB,Iter})}{2} & x_{l^{Iter}}^{UB,Iter} \\ \vdots & \vdots \\ x_n^{LB,Iter} & x_n^{UB,Iter} \end{bmatrix},$$

where l^{Iter} corresponds to the variable with the scaled longest side in the initial rectangle,

$$l^{Iter} = \arg \max_n \frac{(x_i^{UB,Iter} - x_i^{LB,Iter})}{(x_i^{UB} - x_i^{LB})}.$$

Step 4: α determination. For each sub-rectangle, randomly select N points inside it to

calculate its Hessian, $\mathbf{H}(\mathbf{x}_j) = \left[\frac{\partial^2 G(\mathbf{x}_j)}{\partial x_k \partial x_l} \right]$, at the selected point $\mathbf{x}_j$, and the corresponding

eigenvalues, $\lambda_i^{G(\mathbf{x}_j)}$, $j = 1, \dots, N, i = 1, \dots, n$, by solving

$$|\mathbf{H}(\mathbf{x}_j) - \lambda \mathbf{I}| = 0,$$

where $\mathbf{I}$ and λ are the identity matrix and the vector of eigenvalues. α_r is then determined by (2.35).

Step 5: *Determining the solutions of convex problems in two sub-rectangles.* For each sub-rectangle, solve the following convex non-linear optimisation problem by using a non-linear solver:

$$L(\mathbf{x})_{sol}^{r,Iter} = \min_{\mathbf{x}} L(\mathbf{x})^r = G(\mathbf{x}) + \alpha^r \sum_{i=1}^n (x_i^{LB,Iter} - x_i)(x_i^{UB,Iter} - x_i).$$

If the solution $L(\mathbf{x})_{sol}^{r,Iter}$ is less than zero, then it is stored along with the value of the variable x_i at the solution point $x_{i,sol}^{r,Iter}$.

Step 6: Updating the iteration counter $Iter$ and lower bound $G(\mathbf{x})^{LB}$. The iteration is increased by one,
 $Iter \leftarrow Iter + 1$.

The lower bound $G(\mathbf{x})^{LB}$ is updated to the *maximum* solution over the stored ones from previous iterations. Furthermore, the selected solution is erased from the stored set:

$$G(\mathbf{x})^{LB} = L(\mathbf{x})_{sol}^{r', Iter'},$$

where:

$$L(\mathbf{x})_{sol}^{r', Iter'} = \max_{r, I} L(\mathbf{x})_{sol}^{r, I}, r = 1, 2, I = 1, \dots, Iter - 1.$$

Step 7: Determining the current point $x_i^{c, Iter}$ and bounds $x_i^{LB, Iter}$ and $x_i^{UB, Iter}$ on x_i . The current point is selected to be the solution point of the previously found minimum solution in Step five:

$$x_i^{c, Iter} = x_{i, sol}^{r', Iter'}, i = 1, \dots, n.$$

The current rectangle becomes the sub-rectangle containing the previously found solution.

Step 8: Revising the upper bound $G(\mathbf{x})^{UB}$. $G(\mathbf{x}^c)^{UB}$ is calculated at $\mathbf{x}^c = [x_i^{c, Iter}, \forall i]$ and the upper bound $G(\mathbf{x})^{UB}$ is updated as follows: $G(\mathbf{x})^{UB} = \min[G(\mathbf{x})^{UB}, G(\mathbf{x}^c)^{UB}]$. Return to step 2.

The proof of ε convergence to the global minimum is provided below:

Proposition 1: The modified αBB algorithm is convergent to the global minimum given that (2.32) is satisfied.

Proof.

Theorem 4 in Maranas and Floudas (1994b) is applied to complete the proof. To apply this theorem, the following needs to be checked:

1. (2.32) is satisfied, and
2. Our lower bound at each iteration is a *valid* lower bound.

By assumption, (2.32) is always satisfied.

Moreover, our defined lower bound at each iteration is the largest *non-negative* $L(\mathbf{x})$ calculated by solving the lower bound problems, which is less than the global minimum of zero. So it is a valid lower bound. As a result, Theorem 4 in Maranas and Floudas (1994b) applies. This ends the proof.

It is worthwhile to mention that there is often a trade-off between the convergence and efficiency. The proposed sampling method does not guarantee that (2.32) is always satisfied as the minimum eigenvalue is determined based on a limited number of samples. Introducing the sampling method is just to enhance the computation speed. However, if the efficiency is sacrificed, the convergence of the proposed algorithm can be achieved by determining α though the methods in Adjiman et al. (1998a) since (2.32) is always satisfied. It is also believed that if the proposed method adopts the same α determination method as in Adjiman et al. (1998a), the proposed method is more efficient as a better lower bound is used while both methods can guarantee convergence. In addition, if a large number of samples are used (which slows down the computation process), (2.32) is usually satisfied. One more point is that the conventional α BB, genetic algorithms, and other global optimisation methods are guaranteed to solve the minimisation program (2.30), as well as Bell and Cassir (2002)'s one. However, the computation speed is always questionable.

2.4 NUMERICAL STUDIES

To illustrate the computation performance of the modified α BB algorithm under different settings, four scenarios are set up as shown in Table 2.1. The base scenario is set up to illustrate how the initial solution converges to the optimal one and for comparison purposes with the rest of the scenarios, whereas Scenarios 2, 3 and 4 respectively investigate the effects of bound improving strategies, the number of samples, and the fixed α values on the efficiency of the proposed algorithm.

Table 2.1: Settings for the four scenarios

Scenarios	Bound improving strategy	α determination method	Number of Samples
Base scenario (scenario 1)	Proposed strategy	Sampling	25
Scenario 2	Conventional strategy	Sampling	25
Scenario 3	Proposed strategy	Sampling	Varies
Scenario 4	Proposed strategy	User defined value	0

For all the above scenarios, a small network is adopted here for ease of result exposition. The network consists of two links, two nodes, two routes, and one OD pair, as shown in figure 2.2. A virtual route, represented by the dotted arrow, is also shown in this figure, which conceptually carries the non-travel demand between OD pair (1,2). The travel cost $t^{rs}(e^{rs})$ on this virtual route is assumed to satisfy:

$$t^{rs}(e^{rs}) = \frac{e^{rs}}{s^{rs}}, \quad (2.36)$$

where s^{rs} is an OD-specific parameter. The travel costs on the other two links are assumed to satisfy the BPR type relationship:

$$t_{ak}(v_a) = t_a^0 \left(1 + 0.15 \left(\frac{v_a}{c_{ak}} \right)^4 \right), \quad (2.37)$$

where c_{ak} and t_a^0 are the capacity in scenario k and free flow travel cost of link a respectively. The parameter values on these functions are as follows:

- Parameter in the virtual route cost function: $s^{12} = 43.95$ vph per euro.
- Link capacities: $c_{ak} = 500$ vph under normal circumstances, and 250 vph otherwise.
- Free flow travel costs: $t_1^0 = t_2^0 = \text{€ } 10$.

These functions and values are chosen for illustrative purposes. The results were obtained in a Pentium III PC assuming that the number of homogenous players N^{rs} is 2000. The tolerance level is set to be 5×10^{-13} unless specified.

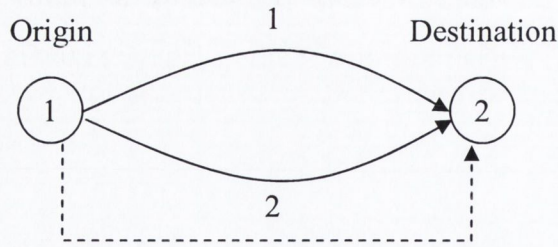


Figure 2.2: The example network

2.4.1 Base Scenario: The Convergence Rate of the Algorithm

Figure 2.3 shows the convergence plot of the modified α BB algorithm for the base scenario with the x-axis to be the number of iterations, and the y-axis the gap value in log scale. This shows a sharp decrease of the gap value (in log scale) in the early iterations and the gap value decreases slowly. This is because most of the partitions appeared in the later iterations and therefore cannot improve the upper bound or gap value. Some of them become fathomed, whereas most of the others cannot produce a better upper bound. A lot of the time is spent on identifying the partitions that can improve the gap value.

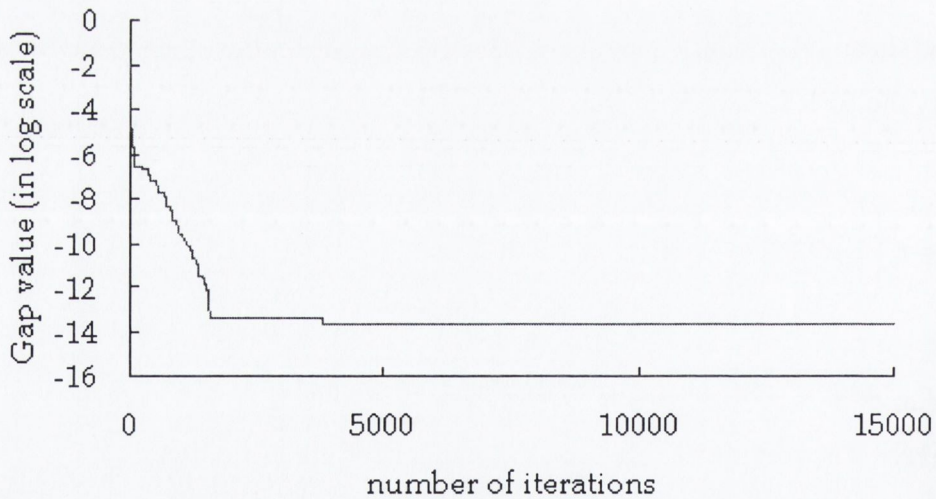


Figure 2.3: The convergence plot for the base scenario

Figures 2.4 and 2.5 reveal how the initial solution converges to the global optimal solution. The symmetry of the network is reflected by the link damage probabilities and the route flows. Both routes carry the same (rescaled) flow and the link damage probabilities are the same at equilibrium. Moreover, at equilibrium, the sum of the probabilities is equal to one,

in which the reason is obvious. However, the sum of the rescaled flows is less than one, meaning that some potential users decide not to travel or travel by other modes.

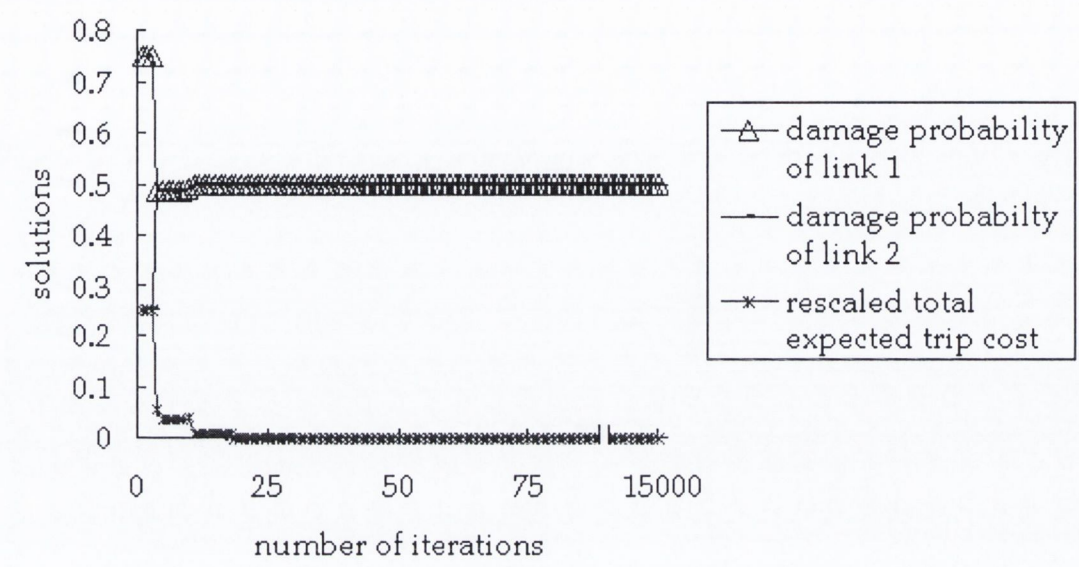


Figure 2.4: The convergence of damage probabilities and the rescaled total expected trip cost

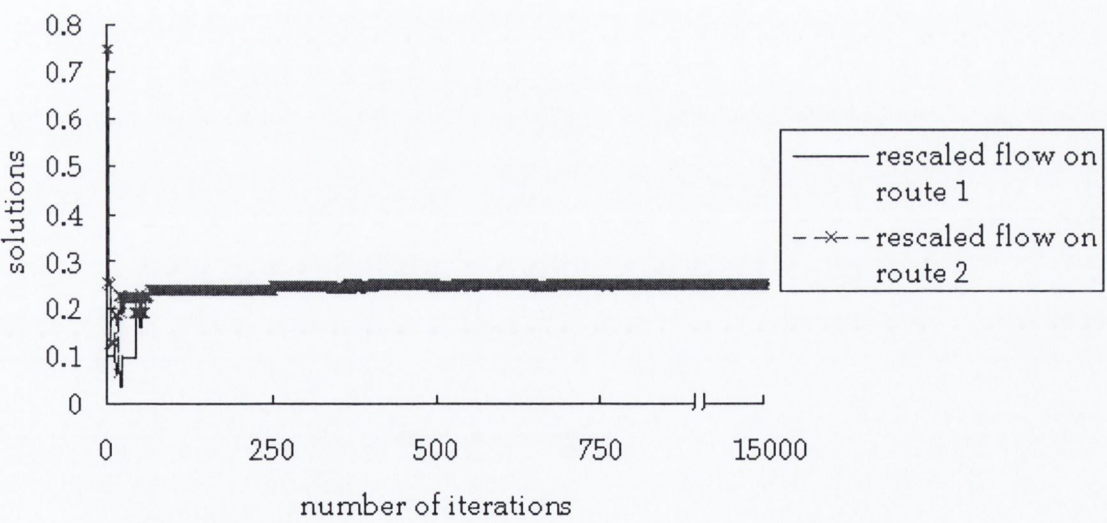


Figure 2.5: The convergence of the rescaled route flows

Figure 2.6 shows the number of function evaluations and the computation time required by a Pentium III PC, against the number of iterations performed. According to this figure, the greater the number of iterations performed, the greater the number of function evaluations

and computation time. Their relationships are approximately linear. This is expected, since in the algorithm, each (major) iteration performs a number of function evaluations, and each function evaluation takes a certain amount of time. However, the number of function evaluations (and hence the computation time) is not constant in each (major) iteration (Table 2.2), as the number of function evaluations for solving the convex lower bounding function is changing over iterations. Nevertheless, the number of iterations performed, and the number of function evaluations can give a rough idea of the computation time required.

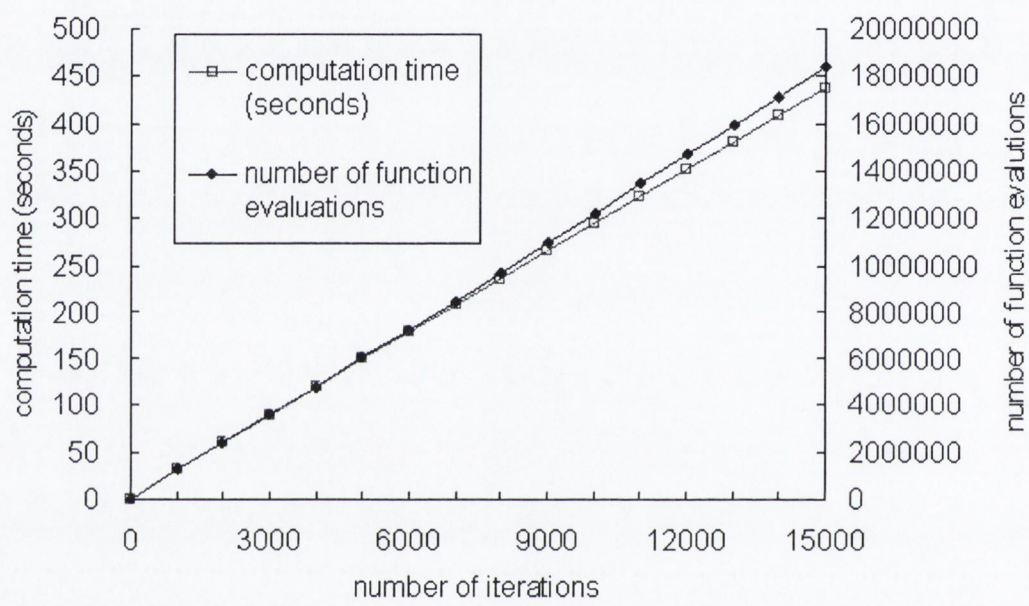


Figure 2.6: The relationship among the number of iterations performed, the number of function evaluations and computation time

Table 2.2: The number of function evaluations over iterations

Iteration count	1	2	3	4
No. of function evaluations	1198	1208	1230	1319

2.4.2 Scenario 2: Comparison of the Two Bound Improving Strategies

To illustrate the effect of the partition strategy on the efficiency of the algorithm, scenario 2 is set up and the results are given in figure 2.7. According to this figure, the bound improving strategy greatly affects the performance of the algorithm. In particular, for the same computation time, the proposed strategy performs better than the conventional one as in the original BB algorithm in Maranas and Floudas (1994 a, b), as revealed by the fact

that the former creates a smaller gap and a better lower bound. Moreover, the proposed strategy always creates much fewer partitions (figure 2.8) as the selected partition always has the value of $L(\mathbf{x})$ closer to zero, making it easier to have the lower bound greater than zero and become fathomed.

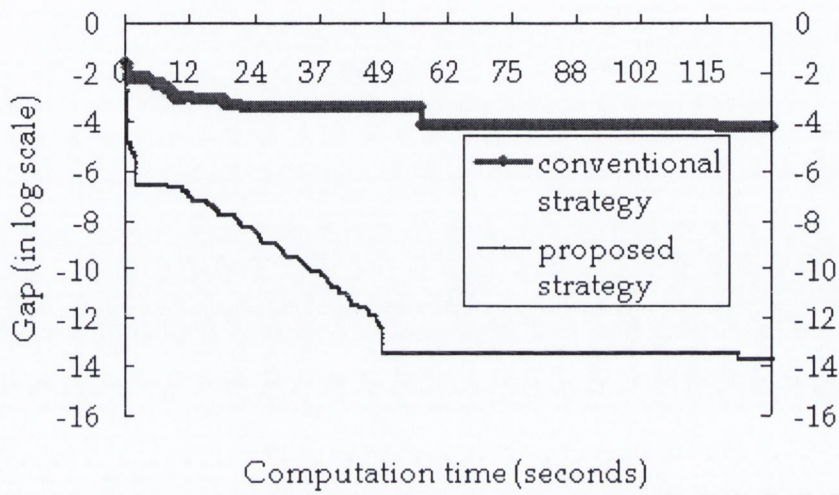


Figure 2.7: A comparison of two bound improving strategies in terms of gap against computation time

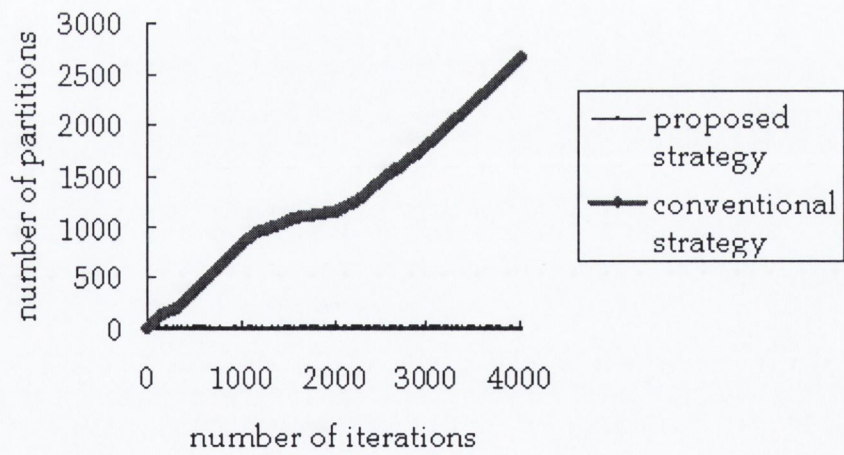


Figure 2.8: A comparison of two bound improving strategies in terms of number of partitions over iterations

2.4.3 Scenario 3: The Effect of Sample Size on Solution Speed

Figure 2.9 compares the convergence behaviour of the algorithm among various sample sizes. The results shows that the greater the number of samples, the greater the computation time required in general, but the trend is not strictly monotonic, and there exist scenarios where opposite trends can be observed. This is because the computation time depends heavily on two components: the quality of the under-estimator (the lower bounding function), and the number of function evaluations in solving the convex lower bounding function. In general, a tighter under-estimator and fewer number of function evaluations in solving the under-estimator result in shorter computation times. However, a tighter under-estimator requires a better estimate of the α value, and a better α estimate requires more samples, which in turn requires more function evaluations and hence more computation time. There is a trade-off between the number of function computations in estimating α and that solving the convex under-estimator. But the result here seems to indicate that few samples always perform better. One point worth mentioning is that overestimating α usually results in an increased number of iterations, whereas underestimating α may result in not being able to obtain global solutions of the original problem. The latter is because the resulting lower bounding function is not convex, and the local optimisation algorithm may fail to find a global minimum of the lower bounding function.

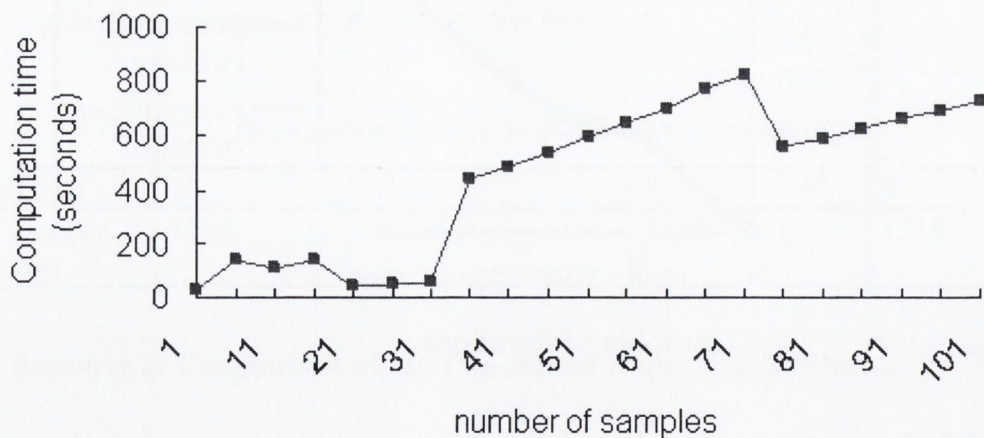


Figure 2.9: Comparison of the effect of the sample size on convergence speed

2.4.4 Scenario 4: The Effect of Fixed User Defined α

The computation time (in seconds) of a given α value to the computation performance is shown in table 2.3. In general, too large or too small α values hinder the attainment of global solutions within 40000 iterations. An α value that is too large results in creating too many partitions which cannot be solved quickly within a finite number of iterations. On the other hand, an α value that is too small results in a failure to obtain global solutions.

This is due to the fact that α is not large enough to make the underestimator convex or nearly convex, and hence the local optimisation algorithm cannot solve for global solutions, or that the small α causes numerical instability and hence the local optimisation algorithm cannot solve for even local optimal solutions in some cases. These are the problems of using user defined α , and are left for future studies. However, for the cases that the optimal solutions can be found here, the computation time required is much less than that of the base scenario (437 seconds), as function evaluations are not required to estimate α . This suggests that when the function evaluation is expensive, one can choose a small user defined α instead of calculating α through sampling for some of the iterations to reduce the computation time.

Table 2.3: The effect of the user defined α values on the computation performance

α value	Less than 0.05	0.05	0.1	0.15	0.2	0.25	0.3	Larger than 0.3
Computation time (seconds)	Cannot be found within 40000 iterations	32	77	131	171	224	88	Cannot be found within 40000 iterations

2.5 CONCLUDING REMARKS

This chapter develops an NCP formulation to model the risk-averse traffic assignment with elastic demands, and proposes a finite ε convergence method to solve for global solutions. A numerical study is also performed to illustrate the effects of bound improving strategies, sample size, and user defined α on the solution speed. The major contributions of this chapter are three-fold: first, it captures the elastic demand in the risk-averse assignment.

Second, a solution method with guaranteed ε convergence is developed. Third, a sampling method is developed for determining α .

CHAPTER 3

Generalisation of the Risk-Averse Traffic Assignment

3.1 INTRODUCTION

In chapter 2 of this thesis and in traditional approaches, the main idea behind the game theoretic approach is based on the notion of a fictitious game played between on the one hand users who seek minimum cost routes, and on the other hand, an OD-specific demon that seeks to maximise the total expected network cost to the users by damaging links on each OD pair in the network. The traffic assignment problem when formulated based on this approach, known as risk-averse traffic assignment, consists of two sub-problems: the user problem and the demon problem. The user problem describes the non-cooperative behaviour of network users, whereas the demon problem describes the evil behaviour in the sense of trying to cause maximum damage to the users. However, the formulation developed based on this approach only allows exactly one demon per OD pair in which each OD-specific demon is not free to damage links outside of its own specific OD pair. This one demon per OD pair assumption is not general enough. For example, it does not allow two links to be damaged on one OD pair but no link to be damaged on other OD pairs. It also restricts the number of demons or links to be damaged to be equal to the number of OD pairs and does not allow the number of links to be damaged in a network to be greater or smaller than the number of OD pairs. Furthermore, the capacity reduction when link damage occurs is typically 50%. In other words only two operational states are accounted for – congested and un-congested. This clearly does not model the wide range of capacity levels.

In this chapter a more general game theoretic formulation is proposed by relaxing the above two assumptions. The existing game theoretic formulation is extended to allow demons (called network-specific demons or simply network demons) to be free to select any links to damage and to impose no restriction on the number of demons or number of links to be damaged. The impact of this extension on the total expected network cost is assessed. Moreover the effect of the assumption of 50% capacity reduction is challenged,

especially when there are multiple demons involved, and when more than one demon selects a particular link to damage. In addition, the effect of using or relaxing this assumption is examined with particular reference to the implications for network (or performance) reliability. A brief study is also conducted to: provide a better insight into the demon behaviour in selecting links to damage; illustrate the existence of multiple solutions; and compare the results obtained from OD-specific and network-specific demons.

Finally, the formulations and results provided in this chapter are based upon a paper by this author for the book series ‘Transportation and Traffic Theory’ which was published following the ‘17th International Symposium on Transportation and Traffic Theory’ (see Appendix B for further details). The remainder of this chapter is organised as follows: section 3.2 proposes a multiple network demon formulation; section 3.3 details the numerical studies while section 3.4 provides some concluding remarks.

3.2 FORMULATION

A general transportation network with multiple links, routes, and Origin-Destination (OD) pairs is considered. There are N^{rs} homogenous players between each OD pair rs , each whom it is assumed seeks minimum cost routes among all other alternatives (hereafter they are referred to in this chapter as the ‘users’). There are also M demons with the freedom to roam the network without restriction and with the capability to select any link to damage, including links already selected for damage by other demon(s). It is assumed that the network-specific demons are intent on causing maximum delays (and hence increased costs) to the users by damaging link(s) in the network. By so doing the demons maximise their expected pay-off which is the total network cost faced by the users. Moreover, it is assumed that no user knows which link(s) the demon(s) will choose to damage, and none of the demons know in advance which link a user may decide to use. In general, it is assumed that each link is to have two costs depending on whether it is damaged or not. However, for the purposes of some of the studies in this chapter, this assumption is relaxed so that the link costs increase with an increasing number of demons that select a particular link to damage. This is due to the assumption that the link capacity reduces linearly or non-linearly with respect to the number of demons choosing that particular link. Furthermore, both N^{rs} and M are assumed to be fixed. With this setting, the proposed risk-averse user

equilibrium traffic assignment problem with multiple network demons can be formulated by applying the results in Nash (1951) and Bell and Cassir (2002).

Similar to Bell and Cassir (2002) and Szeto et al. (2006)/chapter 2 of this thesis, the proposed formulation comprises two main sub-problems: the user problem and the demon problem.

3.2.1 The User Problem

The user problem can be viewed as a non-cooperative game in which each homogenous player tries to select the route with minimum expected trip cost. The user problem can be approximated to deterministic traffic assignment when the number of homogenous users is large (Bell and Cassir, 2002). The first order condition of deterministic traffic assignment can be expressed as the following non-linear complementarity conditions:

$$h_j^{rs} \left[\sum_k g_{jk}^{rs}(\mathbf{h}) q_k - \min_d \left(\sum_k g_{dk}^{rs}(\mathbf{h}) q_k \right) \right] = 0, \quad (3.1)$$

$$\sum_k g_{jk}^{rs}(\mathbf{h}) q_k - \min_d \left(\sum_k g_{dk}^{rs}(\mathbf{h}) q_k \right) \geq 0, \quad (3.2)$$

$$h_j^{rs} \geq 0, \quad (3.3)$$

where h_j^{rs} is the flow on route j between OD pair rs ; $\mathbf{h}$ is the route flow vector; $g_{jk}^{rs}(\mathbf{h})$ represents the cost of route j between OD pair rs in scenario k based on the flow vector $\mathbf{h}$; d is the minimum expected travel cost route between OD pair rs ; q_k is the probability of scenario k occurring.

According to (3.1), if route j connecting OD pair rs carries flow ($h_j^{rs} > 0$), the corresponding expected route travel cost $\sum_k g_{jk}^{rs}(\mathbf{h}) q_k$ must be equal to the minimum expected travel cost $\min_d \left(\sum_k g_{dk}^{rs}(\mathbf{h}) q_k \right)$ between OD pair rs , as the term in the square brackets in (3.1) must equal zero. If route j carries no flow ($h_j^{rs} = 0$), the corresponding

expected route travel cost must be greater than or equal to the minimum expected travel cost based on (3.2). Condition (3.3) is the non-negativity condition of route flows.

The route flow h_j^{rs} in (3.1)-(3.3) must satisfy flow conservation, expressed as:

$$N^{rs} = \sum_j \delta_j^{rs} h_j^{rs}, \quad (3.4)$$

where N^{rs} is the travel demand of OD pair rs ; δ_j^{rs} is the route-OD incidence indicator.

$\delta_j^{rs} = 1$ if j connects OD pair rs , and $\delta_j^{rs} = 0$ otherwise.

The cost of route j between OD pair rs in scenario k , $g_{jk}^{rs}(\mathbf{h})$, in (3.1) and (3.2) is the total cost on every link that is on this route:

$$g_{jk}^{rs}(\mathbf{h}) = \sum_a \delta_{ja} t_{ak}(v_a), \quad (3.5)$$

where δ_{ja} is 1 if link a is on route j , and 0 otherwise; $t_{ak}(v_a)$ denotes the flow-dependent cost on link a in scenario k and is defined by the Bureau of Public Roads (BPR) function in (3.6) below:

$$t_{ak}(v_a) = t_a^0 \left(1 + 0.15 \left(\frac{v_a}{c_{ak}} \right)^4 \right), \quad (3.6)$$

where t_a^0 is the free-flow travel time of link a ; c_{ak} is the capacity of link a in scenario k ; v_a is the link flow given in (3.7) below:

$$v_a = \sum_{rs} \sum_j \delta_{ja} h_j^{rs}. \quad (3.7)$$

The above equation states that the link flow is simply the sum of the route flows using that link.

The scenario probability in (3.1)-(3.2), q_k , depends on the link selection probability of all demons. Let $k = (l_1, \dots, l_m, \dots, l_M)$, l_m be the link selected for damage by demon m , $p_{l_m}^m$ be the probability of demon m selecting l_m to damage, and M be the number of network-specific demons. The scenario probability q_k is defined by:

$$q_k = \prod_{m=1}^M p_{l_m}^m. \quad (3.8)$$

3.2.2 The Demon Problem

The demon problem describes the objectives of the demons. All demons seek the mixed strategy to maximise their individual expected pay-offs or the total expected network cost to the users. The problem for multiple network demons can be written as the following non-linear complementarity conditions:

$$p_{l_m}^m \left\{ \begin{array}{l} \max_{w_m} \sum_{l_1=1}^n \dots \sum_{l_q=1, q \neq m}^n \dots \sum_{l_M=1}^n \left(\prod_{i=1, i \neq m}^M p_{l_i}^i \right) \left[\sum_j g_{j, l_1, \dots, w_m, \dots, l_M}^{rs}(\mathbf{h}) h_j^{rs} \right] \\ - \sum_{l_1=1}^n \dots \sum_{l_q=1, q \neq m}^n \dots \sum_{l_M=1}^n \left(\prod_{i=1, i \neq m}^M p_{l_i}^i \right) \left[\sum_j g_{j, l_1, \dots, l_m, \dots, l_M}^{rs}(\mathbf{h}) h_j^{rs} \right] \end{array} \right\} = 0, m = 1, \dots, M, l_m = 1, \dots, n, \quad (3.9)$$

$$\left\{ \begin{array}{l} \max_{w_m} \sum_{l_1=1}^n \dots \sum_{l_q=1, q \neq m}^n \dots \sum_{l_M=1}^n \left(\prod_{i=1, i \neq m}^M p_{l_i}^i \right) \left[\sum_j g_{j, l_1, \dots, w_m, \dots, l_M}^{rs}(\mathbf{h}) h_j^{rs} \right] \\ - \sum_{l_1=1}^n \dots \sum_{l_q=1, q \neq m}^n \dots \sum_{l_M=1}^n \left(\prod_{i=1, i \neq m}^M p_{l_i}^i \right) \left[\sum_j g_{j, l_1, \dots, l_m, \dots, l_M}^{rs}(\mathbf{h}) h_j^{rs} \right] \end{array} \right\} \geq 0, m = 1, \dots, M, l_m = 1, \dots, n, \quad (3.10)$$

$$p_{l_m}^m \geq 0, \quad (3.11)$$

where n is the number of links in the network; $g_{j, l_1, \dots, l_m, \dots, l_M}^{rs}(\mathbf{h})$ is the cost of route j between OD pair rs in scenario $k = (l_1, \dots, l_m, \dots, l_M)$; l_m and M follow earlier definitions; w_m is the link selected for damage by demon m that gives the demon maximum expected pay-off; i and q are indices for demons. $p_{l_m}^m$ is the probability of demon m selecting l_m to damage, must satisfy:

$$\sum_{l_m} p_{l_m}^m - 1 = 0, \quad (3.12)$$

and be non-negative by definition as shown in (3.11).

Conditions (3.9)-(3.10) represent the necessary and sufficient conditions of the Nash equilibrium of a non-cooperative, mixed strategy game for the demons (see Nash 1951). The second square bracket term in (3.9) is the total network cost in scenario $(l_1, \dots, l_m, \dots, l_M)$. The second term is the total expected network cost when demon m selects link l_m to damage (or the expected pay-off to demon m when it selects link l_m to damage) and considers the link selection strategies of all other demons. The first term is the maximum expected pay-off to demon m or the maximum total expected network cost. When demon m selects a particular link to damage with some probability ($p_{l_m}^m > 0$) then the term inside the braces must be zero and the maximum expected pay-off to that demon (the first term inside the braces of (3.9)) must be equal to the total expected network cost (the second term inside the braces of (3.9)) when the demon selects that link to damage. If a particular demon does not choose a particular link l_m to damage ($p_{l_m}^m = 0$), then the maximum expected pay-off to the demon must be greater than or equal to the total expected network cost when the demon selects link l_m to damage according to (3.10).

3.2.3 The Non-linear Complementarity Problem Formulation

The risk-averse traffic assignment with multiple network-specific demons is to find h_j^{rs} and $p_{l_m}^m$ such that (3.1)-(3.12) are satisfied. This problem can be re-expressed as a Non-linear Complementarity Problem (NCP). To simplify the NCP, two notations are introduced. Let the minimum expected OD travel costs faced by users (or the second term in (3.2)) be ϕ^{rs} :

$$\phi^{rs} = \min_d \left(\sum_k g_{dk}^{rs}(\mathbf{h}) q_k \right), \quad (3.13)$$

and let the maximum expected pay-off to demon m (the first term inside the braces of (3.9) and (3.10)) be π^m :

$$\pi^m = \max_{w_m} \sum_{l_1=1}^n \dots \sum_{l_q=1, q \neq m}^n \dots \sum_{l_M=1}^n \left(\prod_{i=1, i \neq m}^M p_{l_i}^i \right) \left[\sum_j g_{j, l_1, \dots, w_m, \dots, l_M}^{rs}(\mathbf{h}) h_j \right]. \quad (3.14)$$

The risk-averse traffic assignment with multiple network-specific demons can then be expressed as a NCP: to find $\mathbf{x}^*$ such that:

$$\mathbf{x}^* \geq \mathbf{0}, \mathbf{F}(\mathbf{x}^*) \geq \mathbf{0}, \text{ and} \quad (3.15)$$

$$\mathbf{x}^{*T} \cdot \mathbf{F}(\mathbf{x}^*) = 0, \quad (3.16)$$

where the asterisk associated with the variable refers to the optimal solution;

$$\mathbf{x} = \begin{bmatrix} h_j^{rs}, \forall j, rs \\ \phi^{rs}, \forall rs \\ p_{l_m}^m, \forall l_m, m \\ \pi^m, \forall m \end{bmatrix} \text{ and} \quad (3.17)$$

$$\mathbf{F}(\mathbf{x}) = \begin{bmatrix} \sum_k g_{jk}^{rs}(\mathbf{h}) q_k - \phi^{rs} \\ \sum_j \delta_j^{rs} h_j^{rs} - N^{rs} \\ \pi^m - \sum_{l_1=1}^n \dots \sum_{l_q=1, q \neq m}^n \dots \sum_{l_M=1}^n \left(\prod_{i=1, i \neq m}^M p_{l_i}^i \right) \left[\sum_j g_{j, l_1, \dots, l_m, \dots, l_M}^{rs}(\mathbf{h}) h_j^{rs} \right] \\ \sum_{l_m} p_{l_m}^m - 1 \end{bmatrix}. \quad (3.18)$$

It should be clear that when $\mathbf{x}$ satisfies the NCP (3.15)-(3.18), the h_j^{rs} and $p_{l_m}^m$ in $\mathbf{x}$ must satisfy the risk-averse traffic assignment (3.1)-(3.12). Moreover, the NCP (3.15)-(3.18) must have at least one solution. The proof can be found in section 3.2.4. Furthermore, this NCP can have more than one solution. This will be shown in the numerical studies.

The NCP (3.15)-(3.18) can be transformed to a minimisation problem via a gap function. Let Ω be the solution set to the NCP formulation and $\psi = \{\mathbf{x} \geq \mathbf{0}, \mathbf{F}(\mathbf{x}) \geq \mathbf{0}\}$. A function $G(\mathbf{x})$ is the Gap function for the NCP if the following three conditions are satisfied:

$$(i) G(\mathbf{x}) \geq 0;$$

$$(ii) G(\mathbf{x}) = 0 \Leftrightarrow \mathbf{x} \in \Omega, \text{ and};$$

(iii) $\min_{\mathbf{x} \in \mathcal{V}} G(\mathbf{x}) = 0$ is a global minimum.

These properties mean that if $G(\mathbf{x})$ can be found to be zero, then one has a solution to the NCP. On the other hand, if the NCP has no solution, then $G(\mathbf{x})$ must have global solutions with positive objective values (Theorem 3.1 in Kanzow and Fukushima, 1996).

In this chapter, the gap function in Lo and Chen (2000) is adopted:

$$G(\mathbf{x}) = \sum_i \varphi(x_i, F_i(\mathbf{x})), \quad (3.19)$$

where

$$\varphi(a, b) = \frac{1}{2} \phi^2(a, b), \text{ and} \quad (3.20)$$

$\phi(a, b)$ is the *Fischer Function* defined as follows:

$$\phi(a, b) = \sqrt{(a^2 + b^2)} - (a + b). \quad (3.21)$$

By adopting the gap function (3.19), the NCP (3.15)-(3.18) can be reformulated into a minimisation program:

$$\min_{\mathbf{x}} G(\mathbf{x}), \quad (3.22)$$

where $G(\mathbf{x})$ is defined by (3.19)-(3.21). This minimisation program can then be solved by a number of existing algorithms. In this chapter, this program is solved by the Generalized Reduced Gradient (GRG) algorithm (Abadie and Carpentier, 1969). Further details of this algorithm are given in Appendix A.

3.2.4 The Proof of the Existence of Solutions

The existence of solutions of the NCP (3.15)-(3.18) is proven by using two important results in Nagurney (1993):

1. A NCP can be reformulated into a Variational Inequality Problem (VIP) as the VIP contains the NCP as a special case (Proposition 1.4), and

2. A VIP must have at least one solution if the mapping function $\mathbf{F}(\mathbf{x})$ is continuous with respect to $\mathbf{x}$ and if the solution set is the compact convex set and bounded (Theorems 1.4 and 1.5).

Based on Proposition 1.4 in Nagurney (1993), the following proposition is described below:

Proposition 1: The NCP (3.15)-(3.18) can be reformulated into the following VIP.

Find $\mathbf{x}^$ such that:*

$$(\mathbf{x} - \mathbf{x}^*)^T \mathbf{F}(\mathbf{x}^*) \geq 0, \forall \mathbf{x} \in \Omega, \quad (3.23)$$

where $\mathbf{x}$ and $\mathbf{F}(\mathbf{x})$ are defined as in (3.17) and (3.18), respectively; $\Omega = R_+^u$ is the solution set; u is the dimension of $\mathbf{x}$, which is equal to the sum of the number of paths in the network, the number of OD pairs, the product of the number of demons and the number of links, and the number of demons.

This proposition implies that if the VIP (3.23) has a solution, the NCP (3.15)-(3.18) will also have the same solution. The following proves that the VIP (3.23) indeed has a solution using the result in Theorems 1.4 and 1.5 in Nagurney (1993).

Proposition 2: A solution to the VIP (3.23) exists.

Proof:

It is first shown that the solution set Ω can be reduced to a bounded compact convex set. By definition, the link selection probability $p_{l_m}^m$ must not exceed one and the route flow h_j^{rs} must be less than or equal to the corresponding demand:

$$p_{l_m}^m \leq 1 \text{ and} \quad (3.24)$$

$$h_j^{rs} \leq N^{rs}. \quad (3.25)$$

Using (3.24) and (3.25), one has:

$$\sum_k g_{jk}^{rs}(\mathbf{h}) q_k \leq \sum_k g_{jk}^{rs}(\mathbf{N}) q_k \leq \sum_k g_{jk}^{rs}(\mathbf{N}), \quad (3.26)$$

where $\mathbf{N} = [h_j^{rs}, \forall j, rs \mid h_j^{rs} = N^{rs}]$.

Since by definitions $N^{rs} \leq \max_{rs} N^{rs}$ and $\sum_k g_{jk}^{rs}(\mathbf{N}) \leq \max_j \sum_k g_{jk}^{rs}(\mathbf{N})$, (3.25) and (3.26)

can respectively be rewritten as:

$$h_j^{rs} \leq \max_{rs} N^{rs} \quad \text{and} \quad (3.27)$$

$$\sum_k g_{jk}^{rs}(\mathbf{h}) q_k \leq \max_j \sum_k g_{jk}^{rs}(\mathbf{N}). \quad (3.28)$$

By (3.28) and the fact that the minimum expected OD travel cost must be less than or equal to the expected route travel cost between that OD pair (i.e.,

$\phi^{rs} = \min_d \left(\sum_k g_{dk}^{rs}(\mathbf{h}) q_k \right) \leq \sum_k g_{jk}^{rs}(\mathbf{h}) q_k$), it is known that the minimum expected OD travel cost is bounded:

$$\phi^{rs} \leq \max_j \sum_k g_{jk}^{rs}(\mathbf{N}) \quad (3.29)$$

Moreover, by (3.14), (3.24) and (3.25), one has:

$$\begin{aligned} \pi^m &= \max_{w_m} \sum_{l_1=1}^n \dots \sum_{l_q=1, q \neq m}^n \dots \sum_{l_M=1}^n \left(\prod_{i=1, i \neq m}^M p_{l_i}^i \right) \left[\sum_j g_{j, l_1, \dots, w_m, \dots, l_M}^{rs}(\mathbf{h}) h_j \right] \\ &\leq \max_{w_m} \sum_{l_1=1}^n \dots \sum_{l_q=1, q \neq m}^n \dots \sum_{l_M=1}^n \left[\sum_j g_{j, l_1, \dots, w_m, \dots, l_M}^{rs}(\mathbf{N}) N^{rs} \right] \\ &\leq \max_m \left\{ \max_{w_m} \sum_{l_1=1}^n \dots \sum_{l_q=1, q \neq m}^n \dots \sum_{l_M=1}^n \left[\sum_j g_{j, l_1, \dots, w_m, \dots, l_M}^{rs}(\mathbf{N}) N^{rs} \right] \right\}. \end{aligned} \quad (3.30)$$

From (3.24), (3.27), (3.29), (3.30), one can define an upper bound $\tilde{x}$ for each element

x_i of vector $\mathbf{x}$ to be: $\tilde{x} = \max \left\{ 1, \max_{rs} N^{rs}, \max_j \sum_k g_{jk}^{rs}(\mathbf{N}), \Phi \right\}$,

where $\Phi = \max_m \left\{ \max_{w_m} \sum_{l_1=1}^n \dots \sum_{l_q=1, q \neq m}^n \dots \sum_{l_M=1}^n \left[\sum_j g_{j, l_1, \dots, w_m, \dots, l_M}^{rs}(\mathbf{N}) N^{rs} \right] \right\}$.

With the upper bound $\tilde{x}$, one can define a closed ball $B_r(\mathbf{0})$ centred at $\mathbf{0}$ with the

radius $r = \bar{x}$ so that the solution set $\Omega = R_+^u$ to be a bounded compact convex set.

It is now shown that $\mathbf{F}(\mathbf{x})$ is a continuous function of $\mathbf{x}$. From (3.5)-(3.8), it is not difficult for to see that $g_{jk}^{rs}(\mathbf{h})$ is a continuous function of h_j^{rs} and q_k is a continuous function of $p_{l_m}^m$. Consequently, have $\sum_k g_{jk}^{rs}(\mathbf{h})q_k - \phi^{rs}$ to be a continuous function of ϕ^{rs} , $p_{l_m}^m$, and h_j^{rs} . Moreover, it is not difficult for us to see the following:

1. $\pi^m - \sum_{l_1=1}^n \dots \sum_{l_q=1, q \neq m}^n \dots \sum_{l_M=1}^n \left(\prod_{i=1, i \neq m}^M p_{l_i}^i \right) \left[\sum_j g_{j, j_1, \dots, j_m, \dots, j_M}^{rs}(\mathbf{h}) h_j^{rs} \right]$ is a continuous function of π^m , $p_{l_m}^m$, and h_j^{rs} ;
2. $\sum_j \delta_j^{rs} h_j^{rs} - N^{rs}$ is a continuous function of h_j^{rs} , and;
3. $\sum_{l_m} p_{l_m}^m - 1$ is a continuous function of $p_{l_m}^m$.

It can therefore be said that all elements of $\mathbf{F}(\mathbf{x})$ are continuous functions of $\mathbf{x}$ and hence $\mathbf{F}(\mathbf{x})$ is a continuous function of $\mathbf{x}$.

Based on propositions 1 and 2, it can be concluded that the NCP (3.15)-(3.18) must have at least one solution.

3.3 NUMERICAL STUDIES

In this section numerical studies are carried out to: examine the effects of relaxing the assumptions on the demon type and capacity reduction on expected network cost, give an insight into the network demon behaviour in selecting links to damage, demonstrate the existence of multiple solutions to the proposed game, and compare the link selection behaviour of the OD-specific and network demons and their impacts on expected network cost. Overall the results indicate the importance of the assumptions used to expected network cost and reliability measures, and provide some further insights into the nature of the route choice game.

3.3.1 The Effect of Introducing Multiple Network-Specific Demons

In this study, the effect of introducing multiple network-specific demons into the game is considered. The simple test network used to generate results is shown in figure 3.1 below and consists of one OD pair with three parallel links. Since this network has only one OD pair, the demon(s) here can be considered as OD-specific demon(s), and the results here can also be used to explain the effect of introducing more than one demon per OD pair. So, in this study, the term ‘demons’ is used instead of ‘network demons’. The following network parameters also apply:

Free-flow travel times: $t_1^0 = t_2^0 = t_3^0 = 10$ minutes.

Link capacities: $c_{1k} = c_{2k} = c_{3k} = 4000$ Vehicles/hour, if a link is not selected for damage by any demon;
 $c_{1k} = c_{2k} = c_{3k} = 2000$ Vehicles/hour, if a link is selected for damage by one or more demons.

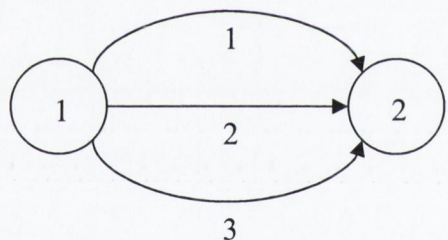


Figure 3.1: Test network 1

While the network chosen need not be symmetric, it is used here for illustrative purposes. In this example, it is assumed that the capacity is always reduced by 50% regardless of how many (network or OD) demons select the same link to damage simultaneously. Again the reason for this is to simplify the analysis. Later, this assumption will be challenged by studying the effects of relaxing it, that is, by allowing for a greater capacity reduction when multiple demons select the same link to damage. It is also assumed here that there is as much freedom as possible for demons to damage links so that on the one hand demons are forced to damage at least one link per day, while on the other hand they are not restricted from selecting the same links to damage that other demons have selected. Therefore the number of possible scenarios of link damage increases with the number of demons in the network. In this study the total expected network cost is calculated for the cases when there

are one, two and then three demons in the network giving, respectively, 3, 9 and 27 possible scenarios of link damage. For comparison purposes, this study also considers the case without demons (i.e., there is no capacity uncertainty).

Figure 3.2 shows the total expected network cost for each case under various levels of network congestion. The level of network congestion is represented by the congestion index (C.I.):

$$\text{Congestion Index} = \frac{\text{Total Number of Network Users}}{\text{Total Network Capacity}},$$

where the total network capacity can be determined in general using the following rules:

- a) For networks with links in series only, the total network capacity is equal to the minimum link capacity of all links.
- b) For networks with links in parallel only, the total network capacity is simply equal to the sum of all the individual link capacities.
- c) For other networks, the total network capacity can be obtained by first dividing the networks into many (strongly connected) sub-networks with only links in series or links in parallel and then by considering each sub-network as an ‘artificial’ link and applying rules a) and b) repeatedly until the total network capacity is obtained. Replacing some links in the original network by many ‘artificial’ links with equal capacity may be necessary before dividing the network into many sub-networks.

In general, the higher the value of C.I., the more congested the network is.

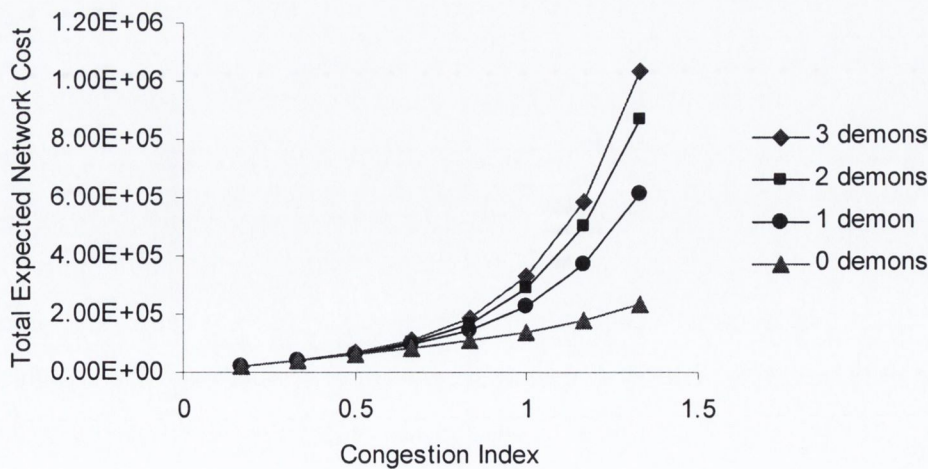


Figure 3.2: The total expected network cost for a range of congested conditions

According to figure 3.2, it can be seen that when the number of demons is fixed, the total expected network cost increases with rising levels of congestion in the network, which is expected. Moreover, the total expected network cost tends to increase with the number of demons present in the network. Again, this is expected. The reasons are as follows: First, the network cost under capacity degradation must be higher than the one without as a lower capacity leads to a higher travel cost. Therefore, the total expected network cost in the multiple-demon case must be higher than the one in the no-demon case. Second, more demons allows more links to be degraded at the same time as shown in table 3.1, and hence a lower network capacity and a higher total expected network cost. This is why two or more demons can certainly impose a higher network cost on the users than one demon.

Table 3.1a: Link capacities for the one demon case

	Scenario		
	1	2	3
Link 1	2000	4000	4000
Link 2	4000	2000	4000
Link 3	4000	4000	2000

Table 3.1b: Link capacities for the two demon case

	Scenario								
	1	2	3	4	5	6	7	8	9
Link 1	2000	4000	4000	2000	2000	2000	4000	2000	4000
Link 2	4000	2000	4000	2000	4000	2000	2000	4000	2000
Link 3	4000	4000	2000	4000	2000	4000	2000	2000	2000

It is also clear from figure 3.2 that the differences in the total expected network cost produced by different numbers of demons become very significant as the number of users (the travel demand) tends to the capacity of the network, i.e. as the network becomes more congested. At low levels of demand relative to the capacity (un-congested conditions), figure 3.2 shows that the number of demons assumed to be active here has little or no impact on the total expected network cost. This result makes sense when one considers some damage to a link such as, for example, a lane closure due to road repairs. This would have the effect of halving the capacity of the link but would not cause that much delay to users when the travel demand is low. However, if the travel demand were high then a lane closure would have a serious impact on the congestion levels. In fact, in this example at the highest demand level used (i.e., 16000 users, C.I. =1.33), it was found that the differences between the total expected network costs is very significant ranging from about 41% when going from assuming one demon to two demons in the network, up to about 68.5% when

going from assuming one demon to three demons. According to Bell (2000), the total expected network cost provides a useful measure of the network reliability since it reflects the expectations of the pessimistic users. Therefore, it can be concluded that the number of demons assumed will have a significant impact on the total expected network cost that is used to measure the network reliability.

The obvious question at this point is – how to determine exactly how many demons one should choose to be active in the network (or in an OD pair)? This is a difficult question to answer since it is also linked to the other question of how much one should assume that each demon contributes to the link capacity reduction when the demon selects a link to damage. However, one possible method may be to choose the number of demons that are suitable based on some available data on incidents that have led to past link failures in the network. Very often, however, this kind of data may be unavailable.

Actually, in this study an even more uncertain network is considered than the ones considered in Bell and Cassir (2002) or in chapter 2/Szeto et al. (2006), since more demons are considered, which results in greater uncertainty in link capacity. If only one demon is active then there are three scenarios of link damage (see table 3.1a) whereas if two demons are active there are nine scenarios of link damage (see table 3.1b). These two tables illustrate how greater uncertainty arises. It is clear from the two-demon case that there are many more scenarios to choose from in which the capacities of links 1, 2, or 3 (or all) are reduced by 50%. Take scenarios 3 to 9 for example - here two link capacities are reduced by 50% since the two demons select different links to damage.

To conclude, in this study the effect of ignoring multiple demons which really means ignoring greater uncertainty in link capacity (or ignoring greater network capacity uncertainty) has been ascertained and the implications to network reliability pointed out, however, the question of exactly how many demons to assume in the model of risk-averse traffic assignment is left to future studies and further discussion on this issue is provided in chapter 7.

3.3.2 The Effect of Introducing Capacity Reduction in some Proportion to the Number of Demons

In chapter 2 and in previous work on risk-averse traffic assignment (Bell and Cassir, 2002) the conventional approach to modelling the damaged capacity was to assume that if a demon selects a link to damage, then the capacity of that particular link is half the value of its initial capacity. The rationale for such an assumption is based on a link having two operational states – congested and un-congested. Even when a number of OD-specific demons select the same link to damage at the same time, the capacity is still reduced by half. Because of this, a similar assumption was used in the first numerical study above: when more than one network-specific demon selects the same link to damage at the same time, the capacity is still reduced by 50%. Indeed, it is reasonable to assume that if many (network or OD-specific) demons select the same link to damage at the same time, the resultant link capacity reduction is in some proportion to the number of active demons selecting this particular link to damage, since there are many more than just two operational states in real road networks. Therefore in this study, the assumptions on link capacity degradation are relaxed due to the presence of multiple demons and the effects of introducing the capacity reduction of a particular link on the maximum expected pay-off to the demon(s) are examined.

The same network as shown in figure 3.1 above is used and therefore the results here can be used to explain the effects of ‘greater capacity degradation’ due to either multiple network-specific demons or multiple OD-specific demons. The free-flow travel times and undamaged link capacities are the same as in the first study (above) but the damaged link capacities are set based on the relaxed assumptions. The values for the damaged link capacities are given in table 3.2. In this example, two different cases of this are considered: Case one (two) capacity reduction is linear (non-linear) with respect to the number of demons. The calculations that were carried out in the first numerical study are repeated here using the same travel demand range, and from this obtain the maximum expected pay-offs to the demon(s) for different congestion levels in the network represented again by the congestion index defined earlier.

Table 3.2: The damaged link capacities

	No. of demons selecting a link to damage			
	0	1	2	3
Link Capacity (vph) for Case 1	4000	3000	2000	1000
Link Capacity (vph) for Case 2	4000	2000	1000	500

3.3.2.1 Case 1: Linear Capacity Reduction

Figure 3.3 shows the total expected network cost in the case of linear capacity reduction. The trend is similar to the one in the first numerical study (i.e., base case). The total expected network cost tends to increase with the number of demons present in the network. Moreover, the total expected network cost increases as the congestion levels rise for any number of demons. Furthermore, like before, the differences in the total expected network costs produced by different numbers of demons are negligible when the congestion levels are low, regardless of the number of demons present. However, when the network is congested, the demons have a much greater impact on the total expected network cost and these differences are quite substantial as can be seen from the graph in figure 3.3. For the highest level of demand shown here of 16,000 users ($C.I=1.33$), the total expected network cost varies greatly depending on how many demons are assumed to be active. In fact, going from assuming one demon to assuming three demons produces a massive 339.9% increase in the total expected network cost. Contrast this to the first numerical study in which a cruder assumption on capacity reduction was used, one finds only a corresponding 68.5% difference in the total expected network cost when going from having one to three active demons in the network. This has important implications for network reliability defined based on expected network cost. The total expected network cost resulting from multiple demons in the network will be quite different (higher) than the case with one demon and will therefore result in a different value that is used to measure the network reliability.

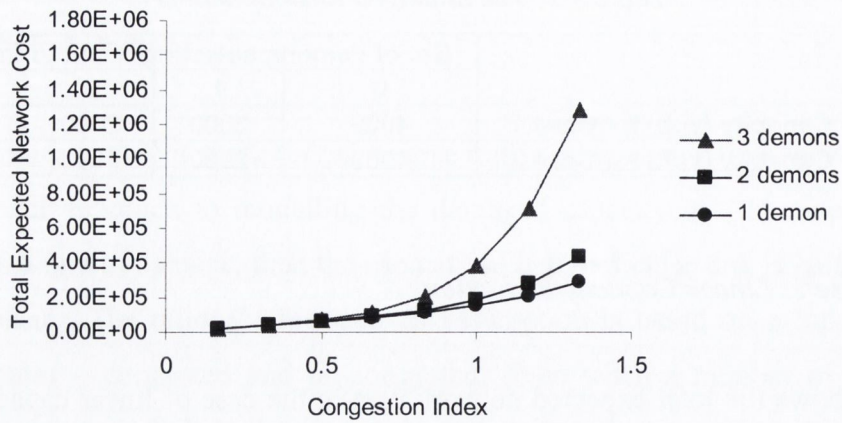


Figure 3.3: The total expected network cost for a range of demons from free-flow to congested conditions when a linear capacity reduction is assumed

3.3.2.2 Case 2: Non-linear Capacity Reduction

Figure 3.4 reveals the total expected network cost in the case of non-linear capacity reduction. The observation is similar but the manner in which the total expected network cost increases with increasing levels of congestion differs and the differences in total expected network cost produced by different numbers of demons are even more substantial when compared with those in the base and linear capacity reduction cases previously examined in this chapter.

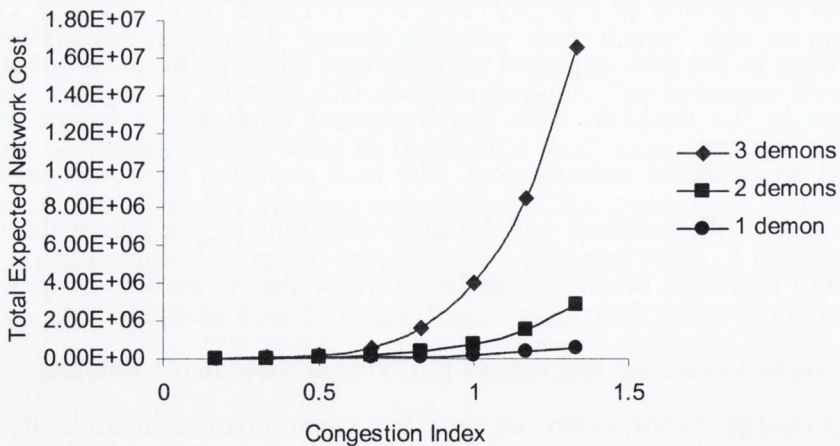


Figure 3.4: The total expected network cost for a range of demons from free-flow to congested conditions when a non-linear capacity reduction is assumed

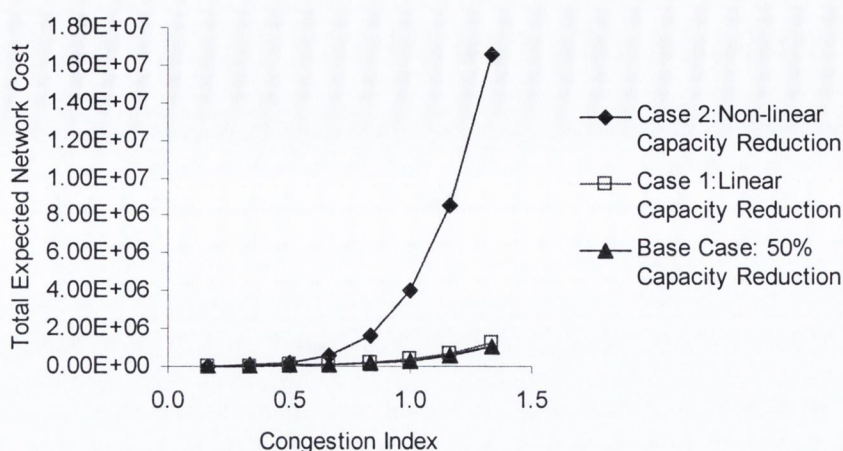


Figure 3.5: The total expected network cost vs. congestion levels for the three different assumptions on link capacity degradation for the three demon case

To illustrate this clearly, figure 3.5 is plotted, which clearly shows the total expected network cost against congestion levels for the three cases on link capacity degradation for the three demon case. From this figure, it can be seen that the total expected network cost in the non-linear capacity reduction case is much larger than those in the base and linear capacity reduction cases when C.I. is greater than or equal to 1. At the highest level of demand (i.e., C.I. = 1.33), the total expected network cost in the non-linear capacity reduction case is 16 times larger than that in the base case. This large difference not only is significant but also poses more questions than it necessarily answers regarding how much uncertainty one should consider (i.e., how many demons) and further how much link degradation should be assumed. From both this numerical study and the first one, it is quite clear that the assumptions regarding capacity degradation may have more significance than those regarding how many demons to select. However, when both of these assumptions are relaxed, they produce significantly higher network cost for users to face and are worthy of further study due to the implications for the network reliability.

3.3.3 The Demon Probabilities – the Demons’ Evil Behaviour

It is worthwhile to have a closer look at the probabilities of demons selecting particular links to damage so as to gain a better understanding of the demons’ evil behaviour (evil in the sense of how they maximise their expected pay-offs). This is of importance since it gives us a sense for which links get damaged with higher probabilities, what scenarios are more likely to occur and also provides us with some insight into the problem. For these

purposes a simple test network is chosen consisting of just one OD pair with two parallel links. This sample network is shown in figure 3.6.

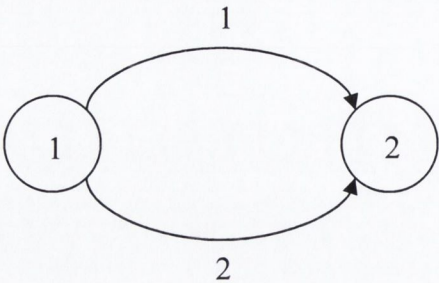


Figure 3.6: Test network 2

The following network parameters also apply:

- Free-flow travel times: $t_1^0 = 5$ mins; $t_2^0 = 10$ mins.
- Link capacities: $c_{1k} = c_{2k} = 4000$ vehicles/hour if there is no damage and 2000 vehicles per hour if demon(s) select these links to damage.

3.3.3.1 Case 1: The One Demon Case

Firstly, a simple case is considered where there is only one demon active in the network, because this case will provide us with some insight into the ‘evil behaviour’ and increases the understanding of what happens when there are two or more demons involved. In this simple case, there are only two resulting scenarios. These two scenarios are respectively: demon 1 selects link 1 to damage; demon 1 selects link 2 to damage. The links that the demon selects to damage varies in their probabilities over a travel demand range (here the demand from 2,000-10,000 users is examined). It is expected that the sole demon will damage the link with the most flow with the highest probability. In this way the demon can impose the highest network cost on most users thereby maximising its maximum expected pay-off.

Figure 3.7 shows the risk-averse flows in the network. Interpreting figure 3.7 is relatively straightforward since all flow will choose link 1 initially because it is the minimum cost route. As link 1 becomes congested (volume/capacity =0.75, or congestion index = 0.375), i.e. when it tends towards its capacity of 4000 vph, the users begin to slowly choose link 2. From a demand of about 6,000 users onwards (or congestion index= 0.75), the users begin

to choose both routes fairly evenly although there is of course still more flow on link 1 since it is the lowest cost route.

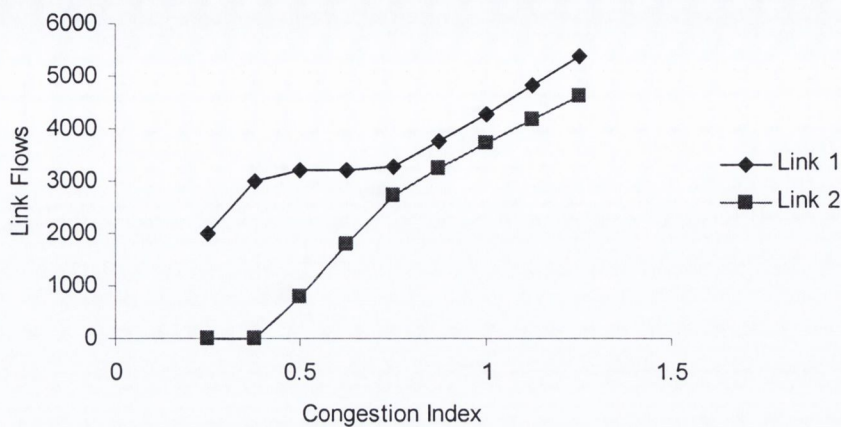


Figure 3.7: The link flow patterns for the case of one active demon

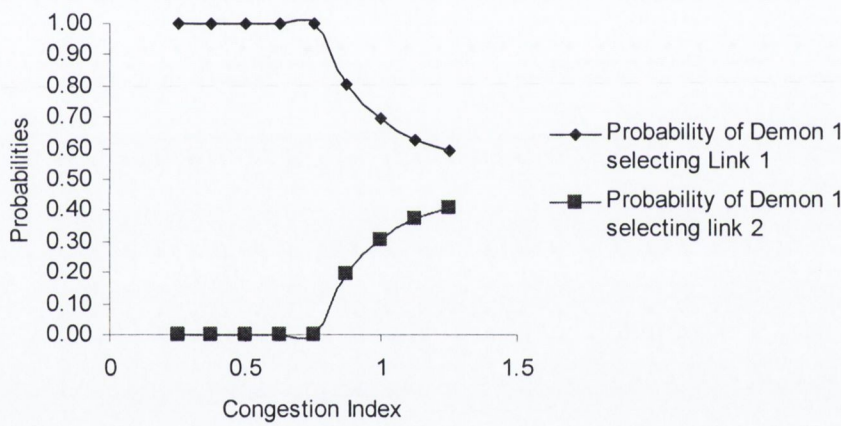


Figure 3.8: The link damage probabilities for the one demon case

The probabilities of demon 1 selecting links 1 and 2 to damage closely mirror the flow patterns. These probabilities are shown in a plot against the rising congestion levels in figure 3.8. The sole demon will choose link 1 to damage until the demand reaches a level approaching capacity of that facility. At this point the demand is high enough to force some users to choose link 2. However, at this point there is still not that much flow on link 2 so the demon continues to damage link 1 almost exclusively. At about 6,000 users (or congestion index = 0.75), the flow on link 2 becomes significant enough to attract the demon to choose link 2 to damage with a somewhat lower probability and link 1 still with a higher probability. However, the values of the probabilities of the demon selecting links 1 and 2 to damage tend towards approx 0.55 and 0.45 respectively once congestion begins to occur. These values seem reasonable as they reflect how the slightly higher probability

of the demon selecting link 1 to damage is due to the fact that there are more users on this link. In this way the demon can maximise the network cost faced by the user and hence, in the process, maximise its own expected pay-off, as it was expected here.

3.3.3.2 Case 2: The Two Demon Case

Secondly, two demons are considered here, which is beneficial as no one has presented and discussed the results about the case of two or more (network-specific or OD-specific) demons. This example adopts the same network and parameters as in the case just above where one demon was considered. As before, the demons are not restricted from damaging the same links at the same time. Table 3.3 sets out the possible scenarios of link damage.

Table 3.3: Scenarios for two demons in network 2

Scenario	Link 1	Link 2
1	Demons 1 & 2	0
2	0	Demons 1 & 2
3	Demon 1	Demon 2
4	Demon 2	Demon 1

In figure 3.9, the link flows for the users who face the maximum (total expected) network cost imposed by two demons are plotted for increasing levels of congestion in the network. These link flow patterns follow a broadly similar trend to flows in the network with one demon (see figure 3.7), and that is, when the network is un-congested the users tend to choose the minimum cost route until the demand rises to a level that approaches the capacity of that facility. After this point the users begin to choose route 2 but do so slightly more readily than in the previous case with just one demon. In general, however, the flow patterns are very similar between the two cases. In some ways like before, the demon behaviour in selecting links to damage reflects the popularity of the routes chosen. When only one active demon was assumed in the network, this demon damaged link 1 with a probability of one at low demand levels, since link 1 was the minimum cost route and it attracted all of the users.

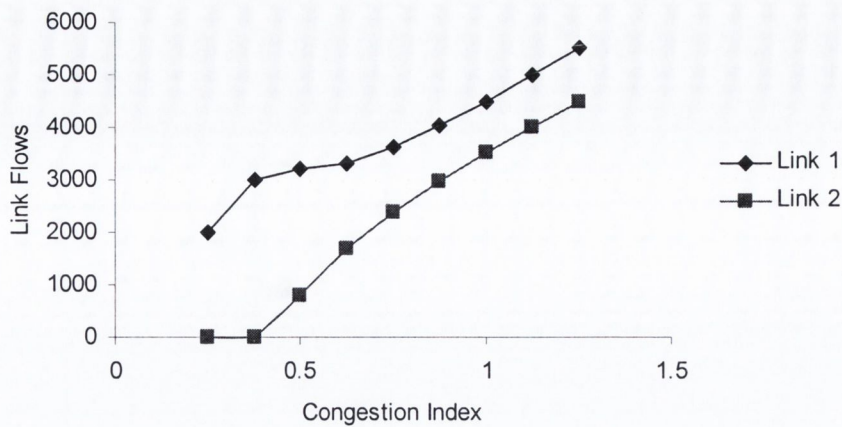


Figure 3.9: The link flow patterns for the case of two active demons

Similarly, in the case of two demons, according to table 3.4, one of the demons damages link 1 with a probability of one when the demand is low (but both routes carry flows) for the same reason just mentioned in the one demon case, while the other demon will damage the other link since there is no further benefit or pay-off to be gained from damaging the same link as the first demon. The explanation for this is relatively simple. If demon 1 damages link 1 with a probability of one which in un-congested conditions imposes maximum (total expected) network cost on all users then demon 2 receives a higher pay-off in damaging link 2 with a probability of one so as to impose the network cost on the remaining users and in the process maximise its expected pay-off. There is no benefit to demon 2 to further damage link 1, since under the capacity assumptions here, the capacity of the link will not reduce any further by doing so. For rising congestion levels when users begin to choose route 2, the pay-off to the demons remain at a maximum since they each individually damage the two links and therefore impose the highest network cost possible on all users.

Table 3.4 also reveals that the demon selection strategies change with the congestion level. For example, demon 1 selects link 1 to damage with probability of one when the congestion indices are 0.25, 0.375, 0.5, 0.875, and 1.25 but selects link 1 to damage with zero probability when the congestion indices are 0.625, 0.75, 1.0, and 1.125. However, as long as one demon selects a different link to damage from the other demon, their pay-offs must be maximised. This suggests that the ‘flip-flop’ solution as shown in table 3.4 is only one of the optimal solutions, which will be further discussed in the next subsection.

Table 3.4: The demons' link selection probabilities for the two demon case

Congestion Index	Link selection probabilities			
	Demon 1 Selects		Demon 2 Selects	
	Link 1	Link 2	Link 1	Link 2
0.25	1	0	0	1
0.375	1	0	0	1
0.5	1	0	0	1
0.625	0	1	1	0
0.75	0	1	1	0
0.875	1	0	0	1
1	0	1	1	0
1.125	0	1	1	0
1.25	1	0	0	1

One point is worthwhile to mention. In a sense the overall trend is similar to the case with one demon where the sole demon begins to damage the link with less flow with a lower probability so as to impose costs on all users. However, significantly, the sole demon only begins to damage link 2 when the number of users on that link is greater than about 3000 vph (congestion index = 0.375). Compare this to the evil behaviour of the two demons in this case, where the other demon always damages link 2 with a probability of one no matter how many users are on that link. These slightly different strategies may help to explain why the maximum expected pay-off is higher when there are two demons instead of one since they quite literally have the resources to impose maximum damage on the users by damaging more links and therefore in the process maximising their expected pay-offs. In fact the percentage increase in the maximum expected pay-off when two demons are assumed here instead of one is approximately 63%. These results follow a similar trend to the earlier studies in this chapter involving multiple demons on network 1 (see figure 3.1). Figure 3.10 below shows the maximum expected pay-offs to the demon(s) for rising congestion levels. As before the significant differences occur when the network is very congested. In general it can be concluded that whenever both links are damaged this implies the optimal solution.

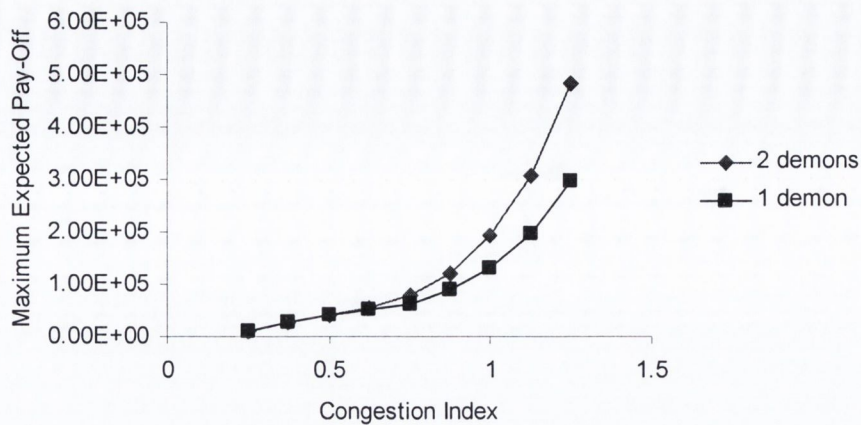


Figure 3.10: The maximum expected pay-off vs. level of congestion in the network

Similar effects and trends in the demon probabilities can be observed for a three links network with 3 network-specific demons. The simpler network used here with two network-specific demons makes it easier to visualise the results. In general, the trends that are shown here hold well but it must be mentioned however that due to the nature of the problem, there can be multiple solutions as shown in the following example.

3.3.4 The Existence of Multiple Solutions

Table 3.5 shows another set of the optimal link selection probabilities for the demons when link 2 carries flow. The overall effect in terms of the total expected network cost is the same, since as the table shows that both demons still select different links to damage for each level of congestion shown. In other words it is demonstrated here that the demons choose exactly the opposite links to damage between the solution shown here and those shown in the previous numerical study (see section 3.3.3).

To summarise the results shown in tables 3.4 and 3.5, when link 2 carries flow then both demons have a probability of one in different links. However, with one small exception, which occurs when the network is very un-congested. In this scenario there are other possible solutions. These other possible probabilities occur when there is no flow on link 2 (i.e., when the demand is lower than the capacity of the minimum cost route). In this case it is possible for both demons to select link 1 (the minimum cost route) to damage with a probability of one. This does not affect the maximum expected pay-off since there is no real benefit in terms of getting a higher network cost for the other demon to damage link 2

when there is no flow on this link so the overall expected pay-offs to the demons remains the same. It can be said, therefore, that when link 2 has no flow, at least one of the demons has a probability of 1 in selecting link 1 to damage.

Table3.5: Another set of the optimal demons' link selection probabilities

Congestion Index	Link selection probabilities			
	Demon 1 Selects		Demon 2 Selects	
	Link 1	Link 2	Link 1	Link 2
0.25	0	1	1	0
0.375	0	1	1	0
0.5	0	1	1	0
0.625	1	0	0	1
0.75	1	0	0	1
0.875	0	1	1	0
1	1	0	0	1
1.125	1	0	0	1
1.25	0	1	1	0

3.3.5 OD-Specific vs. Network- Specific Demons in a Multi-Commodity Network

The purpose of this study is to highlight the impacts of the OD-specific demons as formulated in Bell and Cassir (2002) and the network demons as described by this chapter since in previous examples only one OD pair was considered which does not allow us to consider more than one OD-specific demon and does not allow for other possibilities such as all network demons damaging one OD pair only. To have a fair comparison between the two types of demons, it is assumed that the link capacity is reduced by 50% when it is selected for damage by one or more demons. This assumption is necessary here since there is no possibility of more than one OD-specific demon selecting the same link to damage in the same OD pair. Furthermore this example considers only two network-specific demons to be active in the network since there are only two OD-specific demons present (one for each OD pair). The network used here is shown in figure 3.11 and consists of two OD pairs: OD pairs (1,2) and (2,3). OD pair (1,2) connects origin 1 to destination 2 via two parallel and symmetric links. OD pair (2,3) connects origin 2 to destination 3 via one link only – link 3.

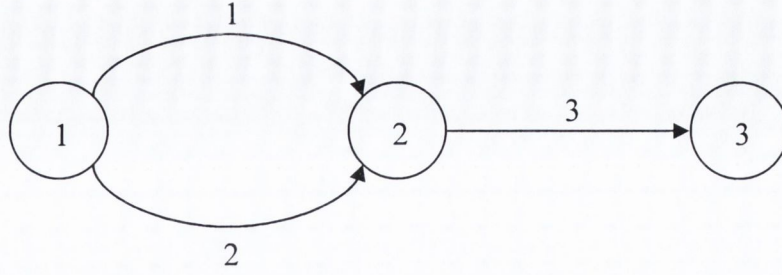


Figure 3.11: Test network 3

The following network parameters chosen for illustrative purposes also apply:

Free-flow travel times: $t_1^0 = t_2^0 = t_3^0 = 10$ mins.

Link capacities: $c_{1k} = c_{2k} = c_{3k} = 4000$ Vehicles/hour, if a link is not selected for damage by any demon;

$c_{1k} = c_{2k} = c_{3k} = 2000$ Vehicles/hour, if a link is selected for damage by one or more demons.

OD travel demand: $N^{12} = 6000$ users,

Two demand levels for OD pair (2,3) are considered: 2000 users (case 1) and 4000 users (case 2).

3.3.5.1 Case 1: $N^{23} = 2000$ Users

The link selection strategies of the demons for this case are shown in table 3.6. Due to the two parallel and symmetric links, the OD-specific demon for OD pair (1,2) will always have an equal probability of selecting links 1 and 2 to damage to maximise the total expected network cost to the users. The OD-specific demon for OD pair (2,3) will always damage link 3 with a probability of one since under the assumptions in Bell and Cassir (2002), the O-D specific demon for OD pair (2,3) is forced to damage at least one link per day and only link 3 connects the OD pair (2,3). Regardless of the demand for travel and network parameters, the OD-specific demons will always behave in this way. However, the same cannot be said for the network demons. The formulation proposed in this chapter allows the network demons to have the freedom to ‘roam’ the network choosing whatever links they wish to damage so as to maximise the total expected network cost to the users.

For the fixed demand level of 2000 users here, both network-specific demons concentrate entirely on damaging links between OD pair (1,2) and they do so in such a way that both demons select different links to damage. Whether demon 1 selects link 1 to damage and demon 2 selects link 2 to damage or vice-versa does not matter as the total expected network cost to the users remains the same.

Table 3.6: Link selection probabilities of the demons for case 1

OD Pair	Link	OD-Specific Demons		Network Demons	
		Demon 1	Demon 2	Demon 1	Demon 2
1-3	1	0.5	0	1	0
	2	0.5	0	0	1
2-3	3	0	1	0	0

It is clear from this example that the network demons are more flexible and can choose to concentrate on one OD pair only, in this case OD pair (1,2), which seems appropriate since this OD pair has a higher demand and therefore more users can be affected by damaging these links. The OD-specific demons do not have this freedom however, and are *forced* to damage different OD pairs. This example shows that this may not necessarily be the best strategy since there are not as many users affected by link 3 being damaged, and therefore the total expected network cost imposed on the users by the OD-specific demons is consequently lower than those imposed by the network demons in this case. In this particular example, the total expected network cost imposed on the users by the network-specific demons is 14.75% higher than that of the OD-specific demons. This is a significant difference and underlines how the network demons give the worse case scenario of link damage in this case. This has important implications for the network reliability since it is defined based on these values.

3.3.5.2 Case 2: $N^{23} = 4000$ Users

In this case, the example considers the same network and network parameters as before. However, the demand parameters are changed somewhat with OD pair (2,3) having an increased travel demand of 4000 users. The resulting demons' link selection strategies are shown in table 3.7. As before, the link selection strategies of the OD-specific demons remains the same (as one would expect) but the network demons select different links to damage than before. In this case the two links in OD pair (1,2) are selected for damage with equal probability by one network demon and the sole link in OD pair (2,3) is selected

for damage with a probability of one by the second network demon. This means that in this case the link damage probabilities under the two assumptions are the same and hence the OD-specific and network-specific demons impose the exact same network cost on the users.

Table 3.7: Link selection probabilities of the demons for case 2

OD Pair	Link	OD-Specific Demons		Network Demons	
		Demon 1	Demon 2	Demon 1	Demon 2
1-3	1	0.5	0	0.5	0
	2	0.5	0	0.5	0
2-3	3	0	1	0	1

By comparing the two cases, it can be seen that the network demons can give the same link damage probabilities as the OD-specific demons. This implies that the network demon approach is more general in the sense that this approach can capture the results obtained by the OD-specific demon approach. This is reasonable as network demons can behave as if they are OD-specific demons. Moreover, the network demons can at least yield the same total expected network cost as the OD-specific demons. This means that the network demon approach can always give the worst case scenario in network reliability analysis. This is again reasonable as network demons have freedom to choose links anywhere in the network to damage, regardless of what OD pair those links connect.

3.4 CONCLUDING REMARKS

This chapter develops a formulation of the risk-averse traffic assignment that allows for the presence of multiple network-specific demons and therefore allows us to consider the effect of greater network uncertainty in terms of capacity degradation. Through numerical studies the importance of the assumption on how many demons to use is demonstrated. A numerical study is also performed to consider the effect of relaxing the assumption on capacity reduction when links are selected for damage by one or more demons. The results show just how important it is to accurately capture the link capacity degradations. Relaxing this assumption also allows for more operational states to be modelled. Apart from these two contributions, this chapter also provides some further insights into the demon behaviour in selecting links to damage and what their optimal damage strategies are, demonstrates the existence of multiple solutions to this route choice game, and compares

the network demon approach proposed here and the traditional OD-specific demon approach.

CHAPTER 4

Measuring Network Reliability Considering Paradoxes: The Multiple Network Demon Approach

4.1 INTRODUCTION

Road networks are now required to be highly reliable to ensure that users can experience smooth travel under both normal and abnormal traffic conditions so that users can avoid unexpected delays due to disruptions such as adverse weather, traffic accidents, breakdowns, signal failures, road-works, and landslides. This requirement has led many researchers to develop various reliability measures or indicators to assess the network reliability, including *encountered reliability* and *travel time reliability*. Bell and Schmöcker (2002) define encountered reliability as the probability of not encountering link degradations on the path with least perceived costs whereas according to Asakura and Kashiwadani (1991), travel time reliability is commonly defined as the probability that a trip can be made successfully within a given acceptable time interval. Travel time reliability receives much more attention in the literature than encountered reliability possibly because the notion of the former was proposed much earlier. A number of previous studies relating to these measures were mentioned in chapter 1.

The above two reliabilities can be studied through the game theoretic approach (e.g., Bell, 2000; Bell and Cassir, 2002; Szeto et al., 2006). One important advantage of this approach (mentioned earlier in chapter one of thesis) is that it can provide normative measures of network reliability. Such normative measures are useful in pointing out the most vulnerable/critical nodes (either origins or destinations), paths, and OD pairs in terms of encountering disruptions and travel time. Different designs may be compared on the basis of these measures.

The main idea behind the game theoretic approach was outlined in some detail in the preceding chapters. However, in chapter 2 and in traditional approaches the formulation developed based on this approach only allows exactly one demon per OD pair in which

each OD-specific demon is not free to damage links outside of its own specific OD pair. This one demon per OD pair assumption is not general enough. For example, it does not allow two links to be damaged on one OD pair but no link to be damaged on other OD pairs. It also restricts the number of demons or links to be damaged to be equal to the number of OD pairs and does not allow the number of links to be damaged in a network to be greater or smaller than the number of OD pairs. Furthermore, the capacity reduction when link damage occurs is typically 50%. In other words only two operational states are accounted for – congested and un-congested. This clearly does not model the wide range of capacity levels. In chapter 3 of this thesis, a more general game theoretic formulation via the non-linear complementarity problem (NCP) approach was proposed by relaxing the assumptions on the number/type of demons and the form of capacity reduction. This proposed formulation extended the existing game theoretic formulation to allow demons (called network-specific demons or simply network demons) to be free to select any links to damage and to impose no restriction on the number of demons or number of links to be damaged. However, the formulation is a path based one that requires path enumeration. This is not attractive from the computational point of view.

Another assumption of the game is that the users selfishly and non-cooperatively seek to maximise their own utilities by seeking the expected minimum cost routes without regard to the overall effect this may have on other network users. This assumption in the traditional user equilibrium problem can give rise to the seemingly counter-intuitive result known as “Braess’ Paradox”, (Braess, 1968). This paradox describes how adding a new link to a transportation network might not improve the operation of the system in the sense of reducing the total system travel time in the network. This phenomenon has received considerable attention in the transportation and more general scientific literature and most of the past efforts can be found in Pas and Principio (1997) or Bazzan and Klugl (2005). However, to date no study has considered the existence of analogues of Braess’ paradox in risk-averse user equilibrium assignment or the resulting implications for travel time and encountered reliability measurements.

In this chapter, a *link-based* NCP multi-demon formulation is proposed which can avoid path enumeration, and is proved to be equivalent to the path-based one proposed in chapter 3. The link-based NCP formulation is also proved to have at least one solution, and can be reformulated into a variational inequality problem and an unconstrained minimisation problem that can be solved by many existing solution algorithms. The impact of this

extension on the travel time and encountered reliabilities is then assessed. Moreover the effect of the assumption of 50% capacity reduction is challenged, especially when there are multiple demons involved, and when more than one demon selects a particular link to damage. In addition, the effect of using or relaxing this assumption with particular reference to the implications for encountered and travel time reliabilities is examined. A brief study is also conducted to demonstrate the existence of two analogues of Braess' paradox for expected network travel time/cost and travel time reliability in risk-averse user equilibrium assignment, namely stochastic Braess' paradox and reliability-paradox, that if one adds a road to a network then all the travellers may be worse off in terms of expected network travel time/cost and travel time reliabilities respectively.

Finally, the formulations and results provided in this chapter are based upon a paper by this author accepted for publication in the journal 'Transportation Research Record' following its earlier inclusion in the 'Proceedings of the 88th Transportation Research Board Meeting', (see Appendix B for further details). The remainder of this chapter is organised as follows: Section 4.2 proposes a link-based multiple network demon formulation. Section 4.3 depicts the reliability measures. Section 4.4 details the numerical studies while section 4.5 provides some concluding remarks.

4.2 FORMULATION

A general transportation network with multiple links, routes, and Origin-Destination (OD) pairs is considered. There are N^{rs} homogenous players between each OD pair rs , each of whom it is assumed seeks minimum cost routes among all other alternatives (hereafter they are referred to in this chapter as the 'users'). There are also M demons with the freedom to roam the network without restriction and with the capability to select any link to damage, including links already selected for damage by other demon(s). It is assumed that the network-specific demons are intent on causing maximum delays (and hence increased costs) to the users by damaging link(s) in the network. By doing so the demons maximise their expected pay-off, which is the total network cost faced by the users. Moreover, it is assumed that no user knows which link(s) the demon(s) will choose to damage, and none of the demons know in advance which link a user may decide to use. In general it is assumed that each link is assumed to have two costs depending on whether it is damaged or not. However, for the purposes of some of the studies in this chapter, this assumption is

relaxed so that the link costs increase with an increasing number of demons that select a particular link to damage. This is due to the assumption that the link capacity reduces linearly or non-linearly with respect to the number of demons choosing that particular link. Furthermore, both N^{rs} and M are assumed to be fixed. In addition, travel time is assumed to be equivalent to travel cost for simplicity although the framework presented here can easily incorporate the value of time to deal with the relationship between travel time and travel cost. With this setting, the proposed risk-averse user equilibrium traffic assignment problem with multiple network demons can be formulated by applying the results in Bell and Cassir (2002) and Nash (1951).

The proposed problem can be formulated through the path-based approach where this approach employs path flows as decision variables. This approach is easier to understand but normally requires that all paths are known in advance or path enumeration. This requirement results in a formulation that is computationally intensive for applications to large networks since there are many paths and hence many variables in the formulation. To develop a formulation that does not have this requirement, the proposed problem is formulated through the link-based approach that adopts link flows as decision variables. The number of variables required in this approach is therefore much fewer because there are considerably fewer links than routes in large networks.

Similar to chapters 2, 3 and Szeto et al. (2006, 2007), the proposed link-based problem comprises two main sub-problems: the demon problem and the user problem.

4.2.1 The Demon Problem

The demon problem describes the objective of each demon. Each demon seeks its mixed strategy to maximise its expected pay-offs. Its pay-off can be mathematically formulated as:

$$\theta_{l_m}^m = \sum_{l_1=1}^n \cdots \sum_{l_q=1, q \neq m}^n \cdots \sum_{l_M=1}^n \left\{ \left(\prod_{i=1, i \neq m}^M P_{l_i}^i \right) \left[\sum_a t_{a, l_1, \dots, l_m, \dots, l_M}(\mathbf{v}) v_a \right] \right\}, \quad (4.1)$$

where $\theta_{l_m}^m$ is the expected pay-off to demon m when demon m selects link l_m to damage (or the total expected network cost to the users when demon m selects link l_m to damage)

and considers the link selection strategies of all other demons; $p_{l_m}^m$ is the probability of demon m selecting l_m to damage; M is the number of network-specific demons; n is the number of links in the network; $\mathbf{v}$ is the vector of link flow; v_a is the flow on link a ; i , m , and q are indices for demons; $t_{a,l_1,\dots,l_m,\dots,l_M}(\mathbf{v})$ denotes the flow-dependent cost on link a in scenario $(l_1,\dots,l_m,\dots,l_M)$ where demon m ($m=1,\dots,M$) selects link l_m to damage. For simplicity, this travel cost is modelled by the separable function below:

$$t_{ak}(v_a) = \alpha_a + \beta_a \left(\frac{v_a}{c_{ak}} \right), \quad (4.2)$$

where α_a is the free-flow travel time of link a ; β_a is the delay parameter for link a ; c_{ak} is the capacity of link a in scenario $k=(l_1,\dots,l_m,\dots,l_M)$.

Let $\pi^m = \max_{l_m} \theta_{l_m}^m$ be the maximum expected pay-off to demon m . The problem for multiple network demons can then be written as the following non-linear complementarity conditions:

$$p_{l_m}^m \{ \pi^m - \theta_{l_m}^m \} = 0, m = 1, \dots, M, l_m = 1, \dots, n, \quad (4.3)$$

$$\pi^m - \theta_{l_m}^m \geq 0, m = 1, \dots, M, l_m = 1, \dots, n, \quad (4.4)$$

$$p_{l_m}^m \geq 0, m = 1, \dots, M, l_m = 1, \dots, n. \quad (4.5)$$

Conditions (4.3)-(4.5) represent the necessary and sufficient conditions of the Nash equilibrium of a non-cooperative, mixed strategy game for the demons (see Nash 1951). According to (4.5), $p_{l_m}^m$ can be positive or zero, depending on whether demon m selects a particular link to damage. When demon m selects a particular link to damage with some probability ($p_{l_m}^m > 0$) then the term inside the braces must be zero and the maximum expected pay-off to that demon, π^m , must be equal to the total expected network cost $\theta_{l_m}^m$ when the demon selects that link to damage. If a particular demon does not choose a particular link l_m to damage ($p_{l_m}^m = 0$), then according to (4.4), the maximum expected pay-off to the demon must be greater than or equal to the expected pay-off of that demon

when that demon selects link l_m to damage. Note that by definition, the probability $p_{l_m}^m$ in (4.3)-(4.5) must satisfy:

$$\sum_{l_m} p_{l_m}^m - 1 = 0. \quad (4.6)$$

4.2.2 The User Problem

The user problem can be viewed as a non-cooperative game in which each homogenous player tries to select the route with minimum expected cost. The user problem can be approximated to deterministic traffic assignment when the number of homogenous users is large (Bell and Cassir, 2002). Unlike the formulations in Szeto et al. (2006, 2007) and chapters 2 and 3 of this thesis, the user problem in this chapter is formulated through the link-based approach. To develop a link-based user problem, let

$$\varphi_a^r = \sum_k t_{ak}(\mathbf{v}) q_k - [\phi^{ry} - \phi^{rx}], \forall r, a = (x, y), \quad (4.7)$$

where q_k is the probability of scenario k occurring; x and y are the entry and exit nodes of link a respectively. ϕ^{ry} is the minimal expected travel time from r to y . The first term in (4.7) is the expected travel time on link a , $\sum_k t_{ak}(\mathbf{v}) q_k$, and the term $[\phi^{ry} - \phi^{rx}]$ is the minimal expected travel time from the entry node of a to the exit node of a for flows departing from origin r . φ_a^r can then be interpreted as the difference between the expected travel time on link a (directly connecting x and y) and the minimal expected travel time from x to y for flows departing from origin r . Alternatively, equation (4.7) can be interpreted as the difference of the minimal expected travel time from r to y , ϕ^{ry} , and the expected travel time from r to y via the minimal expected travel time route from r to x and link a (i.e., $\phi^{rx} + \sum_k t_{ak}(\mathbf{v}) q_k$).

The scenario probability q_k in (4.7) depends on the link selection probability of all demons.

Let $k = (l_1, \dots, l_m, \dots, l_M)$, l_m be the link selected for damage by demon m , $p_{l_m}^m$ be the

probability of demon m selecting l_m to damage, and M be the number of network-specific demons. The scenario probability q_k is then defined by:

$$q_k = \prod_{m=1}^M p_{l_m}^m . \quad (4.8)$$

With (4.7) and (4.8), the first-order condition of deterministic traffic assignment can be expressed as the following link-based non-linear complementarity conditions:

$$v_a^{rs} \varphi_a^r = 0, \forall r, s, a , \quad (4.9)$$

$$\varphi_a^r \geq 0, \forall r, a , \quad (4.10)$$

$$v_a^{rs} \geq 0, \forall r, s, a , \quad (4.11)$$

where v_a^{rs} denotes the flows on link a departed from origin r going to destination s . This disaggregated link flow defines the aggregated link flows v_a :

$$v_a = \sum_{rs} v_a^{rs}, \forall a , \quad (4.12)$$

and must satisfy the flow conservation condition at each node:

$$\sum_{a \in O(z)} v_a^{rs} - \sum_{b \in I(z)} v_b^{rs} - Q_z^{rs} = 0, \forall r, s, z , \quad (4.13)$$

where $Q_z^{rs} = \begin{cases} -N^{rs}, & \text{if } z = s, \\ 0, & \text{otherwise,} \end{cases}$ N^{rs} is the travel demand of OD pair rs ; $O(z)$ is the set

of outbound links emanating from node z ; $I(z)$ is the set of inbound links feeding into node z . Note that the flow conservation equations at origin nodes are not required to be considered as they are redundant.

Conditions (4.9)-(4.11) represent the link-based user choice conditions of the deterministic traffic assignment. When $v_a^{rs} > 0$, link a must be used by the flows departing from r (i.e.

$v_a^r = \sum_s v_a^{rs} > 0, \forall r, a$) and the expected travel time on link a , $\sum_k t_{ak}(\mathbf{v}) q_k$, is equal to the minimal expected travel time $[\phi^{ry} - \phi^{rx}]$ from the entry node of link a to the exit node of link a for flows departing from origin r through $\phi_a^r = 0$ according to (4.9) (or link a is on the minimal travel time route from r to the exit node of link a , y). When $v_a^{rs} = 0$, the expected travel time on link a is greater than or equal to the minimal expected travel time from the entry node of link a to the exit node of link a for flows departing from origin r based on (4.10).

The link-based user choice conditions (4.9)-(4.11) are equivalent to the path-based user choice conditions developed in chapter 3 and re-stated below:

$$h_j^{rs} \left[\sum_k g_{jk}^{rs}(\mathbf{h}) q_k - \min_d \left(\sum_k g_{dk}^{rs}(\mathbf{h}) q_k \right) \right] = 0, \forall rs, j, \quad (4.14)$$

$$\sum_k g_{jk}^{rs}(\mathbf{h}) q_k - \min_d \left(\sum_k g_{dk}^{rs}(\mathbf{h}) q_k \right) \geq 0, \forall rs, j, \quad (4.15)$$

$$h_j^{rs} \geq 0, \forall rs, j, \quad (4.16)$$

where h_j^{rs} is the flow on route j between OD pair rs ; $\mathbf{h}$ is the route flow vector; d is the minimum expected travel cost route between OD pair rs ; $g_{jk}^{rs}(\mathbf{h})$ represents the travel time of route j between OD pair rs in scenario k based on the flow vector $\mathbf{h}$ and is the total travel time on every link that is on this route:

$$g_{jk}^{rs}(\mathbf{h}) = \sum_a \delta_{ja} t_{ak}(\mathbf{v}), \quad (4.17)$$

where δ_{ja} is 1 if link a is on path j , and 0 otherwise; The element of $\mathbf{v}$, v_a , in (4.17) is calculated from v_a^{rs} based on (4.12) and v_a^{rs} is obtained from h_j^{rs} by:

$$v_a^{rs} = \sum_j \delta_{ja} h_j^{rs}. \quad (4.18)$$

The proof of the equivalence is done by six steps as follows:

Step 1: (4.14) implies (4.9).

(4.14) implies that if route j connecting OD pair rs carries flow ($h_j^{rs} > 0$), the corresponding expected route travel time $\sum_k g_{jk}^{rs}(\mathbf{h})q_k$ must be equal to the minimum expected travel time $\min_d \left(\sum_k g_{dk}^{rs}(\mathbf{h})q_k \right)$. Without loss of generality, let $J = (r = z_1 - z_2 - \dots - z_l = s)$ be the minimum expected travel time path connecting origin r and destination s where $z_i, i=1, \dots, l$ is the node on path J and $h_j^{rs} > 0$. (4.14) implies that for every $i=2, 3, \dots, l-1$, sub-path $J_1 = (r = z_1 - z_2 - \dots - z_i)$ is the minimum expected travel time path from r to z_i . This sub-path optimality condition can be proved by contradiction. Suppose J is the minimum expected travel time path from origin r to destination s but sub-path J_1 is not the minimum expected travel time path from r to intermediate node z_i . This implies that there is a shorter expected travel time path J_2 from r to intermediate node z_i compared with J_1 . As a result, there is a path formed by J_2 and subpath $z_i - z_{i+1} - \dots - z_l = s$ connecting r and s that is shorter than J , contradicting the assumption. Therefore, it is concluded that sub-path J_1 is the minimum expected travel time path from r to z_i . The expected travel time on sub-path J_1 is:

$$\phi^{ri} = \sum_{b=1}^{i-1} \sum_k t_{(z_b, z_{b+1})k}(\mathbf{v})q_k. \quad (4.19)$$

Similarly, the minimal expected travel time from r to z_{i+1} is:

$$\phi^{r,i+1} = \sum_{b=1}^i \sum_k t_{(z_b, z_{b+1})k}(\mathbf{v})q_k. \quad (4.20)$$

(4.20) minus (4.19) gives $\phi^{r,i+1} - \phi^{ri} = \sum_k t_{(z_i, z_{i+1})k}(\mathbf{v})q_k$ and

$$\sum_k t_{(z_i, z_{i+1})k}(\mathbf{v})q_k - (\phi^{r,i+1} - \phi^{ri}) = 0. \quad (4.21)$$

As the flow on J is greater than zero ($h_j^{rs} > 0$), the flow on each link on this path must be greater zero (i.e., $v_a^{rs} > 0$), which must occur with condition (4.21)

simultaneously. Therefore $v_a^{rs} \left(\sum_k t_{(z_i, z_{i+1})k}(\mathbf{v}) q_k - (\phi^{r, i+1} - \phi^{ri}) \right) = 0$. Since the above proof can apply to any arbitrarily minimum expected travel time path of any OD pair, it can be concluded that (4.14) implies (4.9).

Step 2: (4.15) implies (4.10).

Consider two paths $J = (r = z_1 - z_2 - \dots - z_i - \dots - z_{i'} - \dots - z_l = s)$ and $J_3 = (r = z_1 - z_2 - \dots - z_i - z_{i'} - \dots - z_l = s)$. Path J is the minimum expected travel time path between r and s whereas J_3 consists of $a = (z_i, z_{i'})$ and two sub-paths of $J : (r = z_1 - z_2 - \dots - z_i)$ and $(z_{i'} - \dots - z_l)$. These two sub-paths are connected by link $a = (z_i, z_{i'})$. From (4.15), $\sum_k g_{J_3k}^{rs}(\mathbf{h}) q_k - \sum_k g_{Jk}^{rs}(\mathbf{h}) q_k \geq 0$. By substituting (4.17) into this inequality and simplifying the resulting inequality, gives $\sum_k t_{ak}(\mathbf{v}) q_k - \sum_{b=i}^{i'-1} \sum_k t_{(z_b, z_{b+1})k}(\mathbf{v}) q_k \geq 0$. The second term on the left hand side is equal to $\phi^{ri'} - \phi^{ri}$ according to (4.19). As a result, $\sum_k t_{ak}(\mathbf{v}) q_k - (\phi^{ri'} - \phi^{ri}) \geq 0, a = (i, i')$. This consideration can be applied to any links in a network, and hence one can have (4.15) implying (4.10).

Step 3: (4.16) implies (4.11).

The non-negativity condition of route flows (4.16) implies the non-negativity condition of link flows (4.11) as link flow is the sum of route flows on that link according to (4.18) and the sum of non-negative real numbers must be non-negative.

Step 4: (4.9) implies (4.14).

Consider a path $j = (r = z_1 - z_2 - \dots - z_{i'} - \dots - z_i - \dots - z_l = s)$ in which all links a on this path carry flows (i.e., $v_a^{rs} > 0$). The complementary condition (4.9) and definition (4.7) require $\sum_k t_{(z_i, z_{i+1})k}(\mathbf{v}) q_k - [\phi^{rz_{i+1}} - \phi^{rz_i}] = 0$. Rearranging this equation gives $\phi^{rz_{i+1}} = \sum_k t_{(z_i, z_{i+1})k}(\mathbf{v}) q_k + \phi^{rz_i}$. This recursive relationship implies

$\phi^{rs} = \sum_{i=1}^{l-1} \sum_k t_{(z_i, z_{i+1})k}(\mathbf{v}) q_k + \phi^{rr}$, where the first term is the expected path travel time $\sum_k g_{jk}^{rs}(\mathbf{h}) q_k$ according to (4.17) and the second term equals zero by definition. It can therefore be concluded that condition (4.9) implies the travel time of a used path j between an OD pair to be equal to the minimal expected travel time for that OD. This analysis can be repeated for other used paths between any OD pairs. It can therefore be concluded that (4.9) implies (4.14).

Step 5: (4.10) implies (4.15).

Without loss of generality, consider all the links on the path $j = (r = z_1 - z_2 - \dots - z_{l'} - \dots - z_l = s)$ between OD pair rs . From (4.10), $\sum_k t_{ak}(\mathbf{v}) q_k - [\phi^{rz_{i+1}} - \phi^{rz_i}] \geq 0, \forall a = (z_i, z_{i+1}), i = 1, \dots, l-1$. Adding all these inequalities together, gives $\sum_{i=1}^{l-1} \sum_k t_{(z_i, z_{i+1})k}(\mathbf{v}) q_k - [\phi^{rs} - \phi^{rr}] \geq 0$. The first term is the expected path travel time $\sum_k g_{jk}^{rs}(\mathbf{h}) q_k$ according to (4.17) whereas the square bracket term $\phi^{rs} - \phi^{rr}$ can be simplified to ϕ^{rs} as $\phi^{rr} = 0$ by definition. This gives $\sum_k g_{jk}^{rs}(\mathbf{h}) q_k - \phi^{rs} \geq 0$. By repeating the same analysis to other paths, one can obtain (4.10) implying (4.15).

Step 6: (4.11) implies (4.16).

Suppose (4.11) holds but $h_j^{rs} < 0$ for some paths j . Consider a link a where there is only one path using it. Since the flow on this path can be negative, the flow on link a can be negative, contradicting the assumption. Hence (4.11) implies (4.16). This completes the proof.

4.2.3 Link-Based Non-linear Complementarity Problem Formulation

The risk-averse traffic assignment with multiple network-specific demons is to find v_a^{rs} and $p_{l_m}^m$ such that (4.1)-(4.13) are satisfied. This problem can be reformulated as a link-based non-linear complementarity problem (NCP): to find $\mathbf{y}$ such that

$$\mathbf{y} \geq \mathbf{0}, \mathbf{H}(\mathbf{y}) \geq \mathbf{0}, \text{ and } \mathbf{y}^T \cdot \mathbf{H}(\mathbf{y}) = 0, \quad (4.22)$$

where

$$\mathbf{y} = \begin{bmatrix} \mathbf{x} \\ \hat{\phi}^{rsz}, \forall rs, z/r \\ \pi^m, \forall m \end{bmatrix}, \mathbf{H}(\mathbf{y}) = \begin{bmatrix} \mathbf{F}(\mathbf{x}) \\ \sum_{a \in O(z)} v_a^{rs} - \sum_{b \in I(z)} v_b^{rs} - Q_z^{rs}, \forall rs, z/r \\ \sum_{l_m} p_{l_m}^m - 1, \forall m \end{bmatrix}, \quad (4.23)$$

$$\mathbf{x} = \begin{bmatrix} v_a^{rs}, \forall rs, a \\ p_{l_m}^m, \forall m, l_m \end{bmatrix}, \quad (4.24)$$

$$\mathbf{F}(\mathbf{x}) = \begin{bmatrix} \psi_a^{rs}, \forall rs, a \\ \pi^m - \theta_{l_m}^m, \forall m, l_m \end{bmatrix}. \quad (4.25)$$

$$\hat{\phi}^{rsz} = \phi^{rz}, \forall r, s, z, \text{ and} \quad (4.26)$$

$$\psi_a^{rs} = \phi_a^r, \forall r, s, a. \quad (4.27)$$

This is because of the following: first, conditions (4.3)-(4.5) and (4.9)-(4.11) can be rewritten as $\mathbf{x} \geq \mathbf{0}$, $\mathbf{F}(\mathbf{x}) \geq \mathbf{0}$, and $\mathbf{x}^T \cdot \mathbf{F}(\mathbf{x}) = 0$. Second, at optimality, the minimal expected travel time to node z from r , $\hat{\phi}^{rsz}$ and the pay-off of demon m , π^m must be positive, and hence flow conservation (4.13) and the constraint for link selection probabilities (4.6) must hold at optimality. Third, including $\sum_{a \in O(z)} v_a^{rs} - \sum_{b \in I(z)} v_b^{rs} - Q_z^{rs} \geq 0$ and $\sum_{l_m} p_{l_m}^m - 1 \geq 0$ do not change the optimality conditions of the original problem.

The NCP (4.22) is equivalent to the following variational inequality problem (VIP) (Proposition 1.4 in Nagurney (1993)):

Find $\mathbf{y}^*$ such that:

$$(\mathbf{y} - \mathbf{y}^*)^T \mathbf{H}(\mathbf{y}^*) \geq 0, \forall \mathbf{y} \in \Omega, \quad (4.28)$$

where $\mathbf{y}$ and $\mathbf{H}(\mathbf{y})$ are defined by (4.23); $\Omega = R_+^u$ is the solution set; $u = (|OD| + M)N + |OD|(|Z| - 1) + M$, and $|OD|$ and $|Z|$ are the number of OD pairs and

nodes respectively. The mapping function $\mathbf{H}(\mathbf{y})$ in (4.28) is continuous with respect to $\mathbf{y}$. Moreover, the solution set can be reduced to a closed ball $B_r(\mathbf{0})$ centred at $\mathbf{0}$ with the radius $r = \max\{1, \Phi_1, \Phi_2, \Phi_3\}$, where $\Phi_1 = \max_{rs} \sum_{rs} N^{rs}$, $\Phi_2 = \max_{rs} \max_{j \in P^{rs}} \left(\sum_k \delta_{ja} t_{ak}(\Phi_1) \right)$, $\Phi_3 = \max_m \left\{ \max_{w_m} \sum_{l_1=1}^n \dots \sum_{l_q=1, q \neq m}^n \dots \sum_{l_M=1}^n \left[\sum_a t_{a, l_1, \dots, w_m, \dots, l_M}(\Phi_1) \Phi_1 \right] \right\}$ and P^{rs} is the path set for OD pair rs . By theorems 1.4 and 1.5 in Nagurney (1993), the VIP (4.28) must have at least one solution. Therefore, the NCP (4.22), which is equivalent to the VIP (4.28), must have at least one solution.

The NCP (4.22) can be also transformed to an unconstrained minimisation problem via the gap function proposed in Lo and Chen (2000):

$$\min_{\mathbf{y}} G(\mathbf{y}) = \frac{1}{2} \sum_i \left[\sqrt{(y_i^2 + H_i(\mathbf{y})^2)} - (y_i + H_i(\mathbf{y})) \right]^2. \quad (4.29)$$

This reformulation is based on the fact that when $G(\mathbf{y})$ is minimised, $G(\mathbf{y}) = 0$ and condition (4.22) is satisfied. This minimisation program can then be solved by a number of existing optimisation algorithms. In this chapter, this program is solved by the Generalized Reduced Gradient Algorithm (Abadie and Carpentier, 1969), although the problem can also be solved by solving (4.28) via existing VIP solution methods. Further details of this algorithm are provided in Appendix A.

4.3 THE RELIABILITY MEASURES

This chapter considers two types of reliabilities: encountered reliability and travel time reliability.

4.3.1 Encountered Reliability

Encountered reliability is the probability of not encountering link degradations on the path with least perceived costs (Bell and Schmöcker, 2002). It measures the likelihood of users encountering a disruption on their *preferred routes* (or *selected routes*). The level of user

information about the possible and actual locations of disruptions is important to consider when calculating encountered reliability since users will often try to avoid degraded links and links which may be degraded. Other than the level of the information to the users, the extent of their risk-averseness is also important to consider when determining encountered reliability because risk-averse users may be willing to travel longer to avoid disruptions whereas risk-neutral users will still travel on their preferred routes based on perceived cost considerations regardless of the probability of encountering disruptions along their preferred routes. In the following sections the encountered reliability measures are defined on the path and OD levels.

4.3.1.1 Encountered Reliability of a Path

It is defined as the probability of a trip encountering no link failure (link degradation) of all links on the path or the probability of no demon damaging the least cost path:

$$R_j^E = \sum_k \left(\prod_a \left[S_a(c_{ak} = c_a^0) \right]^{\delta_{ja}} \right) q_k, \quad (4.30)$$

where R_j^E is the encountered probability of path j and δ_{ja} is the route link incidence indicator; δ_{ja} is 1 if link a is on path j , and 0 otherwise. c_a^0 is the initial capacity of link a ; $S_{ak}(c_{ak} = c_a^0)$ is the 0-1 state variable of link a in scenario k . It is equal to 1 if the capacity of link a in scenario k , c_{ak} , equals its initial capacity c_a^0 and 0 otherwise.

4.3.1.2 Encountered Reliability of an OD Pair

It is the probability of a trip not encountering a failure on any link used by a trip, or the weighted average of the encountered reliability of each path between the OD pair. The weighted average of each path is represented by the proportion of demand on that path. Mathematically, the encountered reliability of a trip between OD pair rs :

$$R_{rs}^E = \sum_{j \in P^{rs}} \frac{h_j^{rs}}{N^{rs}} R_j^E, \quad (4.31)$$

where R_j^E is the encountered probability of a trip between OD pair rs ; h_j^{rs} is the flow on

path j and P^{rs} is the path set of OD pair rs .

4.3.2 Travel Time Reliability

Travel time reliability is commonly defined as the probability that a trip on a particular path can be made successfully within a given acceptable time threshold (Asakura and Kashiwadani, 1991). This measure is a direct indicator of the impact of disruptions on network users, and is a good indicator if the travel time is likely to increase significantly due to the disruptions or the events that lead to link degradations (Nicholson et al., 2003). This measure is influenced by the precision of the information about disruptions available to the users as the information affects their evaluation of the travel time on each route, their route choice, and hence their actual travel times. Moreover, this measure is affected by the level of travel demand. In general, the higher the level of travel demand, the higher the travel time of each user because of the congestion effect.

Travel time can also be defined on the OD and network levels in addition to the path level discussed before. OD (network) travel time reliability is defined as the probability that the weighted average of all path travel times between a given OD pair (the sum of the travel times of all trips in the network) is less than or equal to an acceptable threshold. Unlike the path travel time reliability that measures the performance of the path and gives useful information to travellers, the OD and network travel time reliabilities measure the performances of OD pairs and networks respectively and can be used by network planners for network planning purposes. In the following sections the travel time reliability measures are defined on the path, OD and network levels.

4.3.2.1 Path Travel Time Reliability

It is defined as the probability that the travel time of a given path j, T_j^P , is within an acceptable threshold ε_j :

$$R_j^P = P(T_j^P \leq \varepsilon_j) = \sum_k S_{jk}^{rs} (g_{jk}^{rs} \leq \varepsilon_j) q_k, \quad (4.32)$$

where R_j^P is the reliability of path j ; $P(\cdot)$ denotes the probability function; $S_{jk}^{rs} (g_{jk}^{rs} \leq \varepsilon_j)$ is the 0-1 state variable for path j between OD pair rs in scenario k , which is equal to 1 if the travel time of path j between OD pair rs in scenario k , g_{jk}^{rs} , is less than or equal to the acceptable threshold ε_j and 0 otherwise. The path travel time g_{jk}^{rs} is the total travel time on every link that is on this path and is defined in (4.17).

4.3.2.2 OD Travel Time Reliability

It is defined as the probability that the travel time of a trip between a given OD pair rs , T_{rs}^{OD} , is within an acceptable threshold ε_{rs} :

$$R_{rs}^T = P(T_{rs}^{OD} \leq \varepsilon_{rs}) = \sum_k S_k^{rs} \left(\sum_{j \in P^{rs}} \frac{h_j}{N^{rs}} g_{jk}^{rs} \leq \varepsilon_{rs} \right) q_k, \quad (4.33)$$

where R_{rs}^T is the OD travel time reliability of OD pair rs , $S_k^{rs} \left(\sum_{j \in P^{rs}} \frac{h_j}{N^{rs}} g_{jk}^{rs} \leq \varepsilon_{rs} \right)$ is the 0-1 state variable for OD pair rs in scenario k , which is equal to 1 if the travel time between OD pair rs in scenario k , $\sum_{j \in P^{rs}} \frac{h_j}{N^{rs}} g_{jk}^{rs}$, is within the acceptable threshold ε_{rs} .

4.3.2.3 Network Travel Time Reliability

Network travel time reliability R_N^T is defined as the probability of the sum of the travel times of all trips, TSTT, is less than or equal to an acceptable threshold ε_{TSTT} :

$$R_N^T = P(TSTT \leq \varepsilon_{TSTT}) = \sum_k S_k \left(\sum_j g_{jk}^{rs} h_j^{rs} \leq \varepsilon_{TSTT} \right) q_k, \quad (4.34)$$

where R_N^T is the network travel time reliability. $S_k \left(\sum_j g_{jk}^{rs} h_j^{rs} \leq \varepsilon_{TSTT} \right)$ is the 0-1 state variable for the network, which is equal to 1 if TSTT in scenario k , $\sum_j g_{jk}^{rs} h_j^{rs}$, is within the acceptable threshold ε_{TSTT} .

4.4 NUMERICAL STUDIES

4.4.1 The Effect of Multiple Network-Specific Demons and Multiple States of Link Damage to Reliability Measurement

The numerical studies in this example consider the effect of (i) introducing multiple network-specific demons into the game, and (ii) relaxing the capacity reduction assumption, on the network reliability measures defined earlier. The simple test network and parameters used to generate results is the classical Braess network (see figure 4.1). Since this network has only one OD pair, the demon(s) here can be considered as OD-specific demon(s), and the results here can also be used to explain the effect of introducing more than one demon per OD pair. So, in this study, the term ‘demons’ is used instead of ‘network demons’. The following parameters are also adopted here:

Normal link capacities:

- $c_{1,k} = c_{2,k} = c_{3,k} = c_{4,k} = c_{5,k} = 1$ if a link is not selected for damage by any demon

Damaged link capacities:

Case 1: 50% Capacity Reduction:

- $c_{1,k} = c_{2,k} = c_{3,k} = c_{4,k} = c_{5,k} = 0.5$ if a link is selected for damage by one or more demons

Case 2: Non-linear Capacity Reduction:

- $c_{1,k} = c_{2,k} = c_{3,k} = c_{4,k} = c_{5,k} = 0.5$ if a link is selected for damage by one demon;
- $c_{1,k} = c_{2,k} = c_{3,k} = c_{4,k} = c_{5,k} = 0.25$ if a link is selected for damage by two demons simultaneously;
- $c_{1,k} = c_{2,k} = c_{3,k} = c_{4,k} = c_{5,k} = 0.125$ if a link is selected for damage by three demons simultaneously.

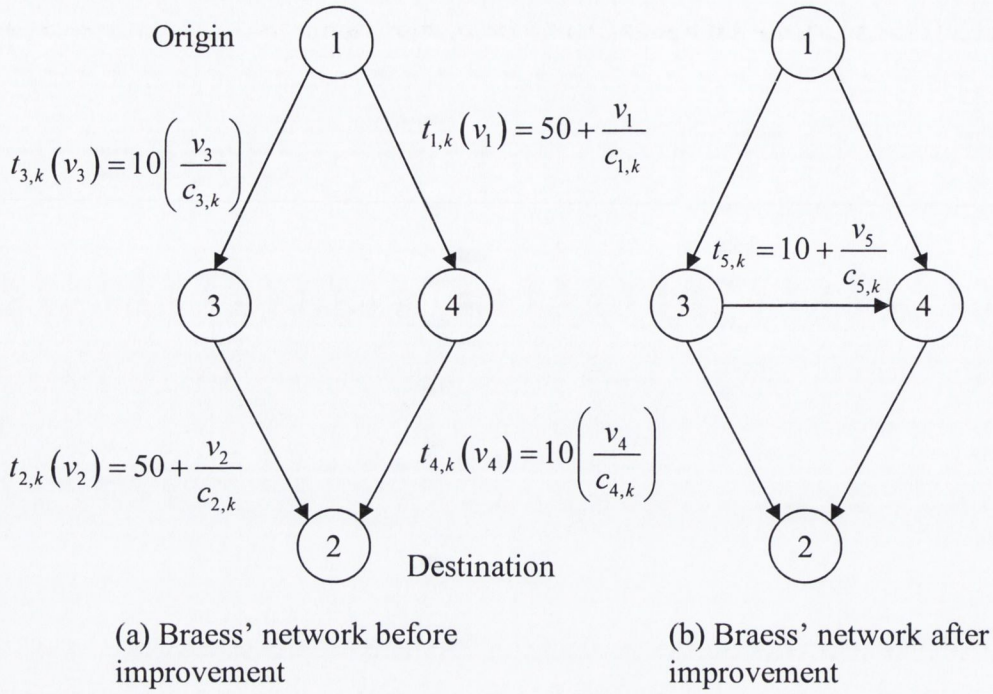


Figure 4.1: Braess' network

The first numerical study considers the Braess network before the addition of the new link (link 5). The resulting four-links network is symmetric and is chosen here to illustrate the effects of introducing multiple demons without being complicated by other factors. Furthermore, the results presented here will facilitate an easy comparison with the ones obtained in the next study when the fifth link is introduced into the Braess network and the existence of paradoxical phenomena is demonstrated in the route choice game. In this example, it is assumed first that the capacity is always reduced by 50% regardless of how many (network or OD) demons select the same link to damage simultaneously. Then this assumption is challenged by assuming as the second case of capacity degradation that the link capacity reduces non-linearly with respect to an increasing number of demons present. As much freedom as possible for demons to damage links is also assumed, so that on the one hand demons are forced to damage at least one link per day, while on the other hand they are not restricted from selecting the same links to damage that other demons have selected. Therefore the number of possible scenarios of link damage increases with the number of demons in the network.

Travel Time Reliability The travel time reliabilities are calculated for the case of one, two and three demons under the two capacity reduction assumptions: 50% capacity reduction (case 1) and non-linear capacity reduction (case 2). The acceptable threshold for each travel time reliability measure is defined as the product of the corresponding user

equilibrium travel time and the safety margin factor. In this study, the safety margin factor is set to be 1.2.

The results are presented in table 4.1. Since the network consists of only one OD pair, the OD travel time reliability and network travel time reliability are the same. Furthermore since the network is symmetric, the path travel time reliabilities of paths one and two are equal. One other point worth noting is that this example does not establish the exact demand levels at which the reliabilities jump from one value to the next. However, this study is not concerned with establishing this here but rather focuses on the overall effect of assuming the presence of multiple demons and relaxing the capacity assumptions on the reliability measures.

For the case of 50% reduction, at the lowest levels of demand (one and two users) shown in table 4.1 (case 1), the number of demons to assume is not altogether that important since there are not that many users on the degraded routes. Thereafter there is a significant drop in the path travel time reliabilities for every other rising level of demand shown here. The impact of the multiple demons is clear – the greater the number of demons the greater is the drop in path travel time reliability since the network capacity is lower on average. In more congested conditions the additional demons have a far greater impact since there are many more users travelling on degraded paths.

For the case of non-linear capacity reduction, the path travel time reliabilities decrease even more quickly though of course this is to be expected since the demons can combine to inflict even higher expected network costs on the users. In fact it seems that the capacity reduction assumption may be even more important than the number of demons to assume due to the additive effect of the demons in significantly lowering the travel time reliabilities. It also seems more appropriate for the capacity to degrade in this way with respect to the increasing number of demons if one considers multiple capacity states in reality and real-life events which can combine to lower the capacity of a road.

The overall trend of the network travel time reliabilities shown for both cases can also be explained in a similar way to the path travel time reliabilities. From table 4.1 (case 2), the effect on the network travel time reliabilities when a non-linear capacity assumption is used can be clearly seen. The network is considered to be unreliable at much lower demand levels since the multiple network demon formulation has the capability to model many

more operational states such as the ones shown in case two, and therefore can be said to better capture the worse-case scenario of the network reliability.

Table 4.1: *The effect of multiple network-specific demons and multiple states of link damage on travel time reliability measurement*

Travel Time Reliabilities for a Range of Demand Levels under Different Levels of Capacity Uncertainty for Braess' Four-Links Network						
Case 1: 50% Capacity Reduction						
	Path Travel Time Reliability			Network Travel Time Reliability		
Demand	1 Demon	2 Demons	3 Demons	1 Demon	2 Demons	3 Demons
1	1	1	1	1	1	1
2	1	1	1	1	1	1
3	0.5	0.25	0.125	1	0.5	0.25
4	0.5	0.25	0.125	1	0.5	0.25
5	0.5	0.25	0.125	1	0.5	0.25
6	0.5	0.25	0.125	1	0.5	0.25
7	0.5	0.25	0.125	1	0.5	0.25
8	0.5	0.25	0.125	0	0	0
9	0.5	0.25	0.125	0	0	0
10	0.5	0.25	0.125	0	0	0
Case 2: Non-linear Capacity Reduction						
	Path Travel Time Reliability			Network Travel Time Reliability		
Demand	1 Demon	2 Demons	3 Demons	1 Demon	2 Demons	3 Demons
1	1	0.75	0.50	1	1	0.750
2	1	0.75	0.50	1	0.50	0
3	0.50	0.250	0.125	1	0.50	0
4	0.50	0.250	0.125	1	0.50	0
5	0.50	0.250	0.125	1	0.50	0
6	0.50	0.250	0.125	1	0.50	0
7	0.50	0.250	0.125	1	0.50	0
8	0.50	0.250	0.125	0	0	0
9	0.50	0.250	0.125	0	0	0
10	0.50	0.250	0.125	0	0	0

The effects of relaxing the number of demons are more easily understood when a comparison of the total expected network costs is made between assuming one demon and three demons in this scenario. For example at a high demand level of 10 users, going from assuming three demons as opposed to just one produces a massive 77% increase in the total expected network costs faced by the users for the case of non-linear capacity reduction. According to Bell (2000) this value is considered as a measure of the network reliability. So while the results are not surprising they do raise some interesting questions

such as – how to determine exactly how many demons should be chosen to be active in the network (or in an OD pair)? This is a difficult question to answer since it is also linked to the other question of how much one should assume that each demon contributes to the link capacity reduction when the demon selects a link to damage. However, one possible method may be to choose the number of demons that are suitable based on some available data on incidents that have led to past link failures in the network. Very often, however, this kind of data may be unavailable. This issue is further discussed in chapter 7.

Encountered Reliability Overall, the effect of assuming the presence of one demon results in a constant path encountered reliability value of 0.5, two demons results in a value of 0.25 while three demons results in a value of 0.125, for the two cases of capacity reduction. The results show that when the number of demons is fixed, the path encountered reliability remains the same with increasing levels of travel demand. This is because the network is symmetric, leading to the outcome that both paths have the same flow and path selection probabilities for damage. For asymmetric networks, the path encountered reliability changes with demand levels. Moreover, the path encountered reliability decreases with the increasing number of demons present in the network. This result is expected in any network since more demons allow more links to be degraded in the network at the same time which means that the network capacity is lower on average. As a result, the users then have a higher probability in encountering a link failure in their journey and consequently the encountered reliability is lower for multiple demons than one demon. In fact in the special case where the network is symmetric like here, any assumption on capacity reduction will produce the same set of results, since by definition the encountered reliability is the probability of a trip-maker encountering failure of *any* magnitude on their path or journey.

4.4.2 Stochastic Braess' Paradox and Reliability Paradox in Risk-Averse User Equilibrium Traffic Assignment

In this section, the existence of stochastic Braess' paradox and a reliability paradox in risk-averse user equilibrium traffic assignment is demonstrated through a numerical example. This example adopts the improved Braess network (i.e., after the fifth link has been added) and for simplicity considers the capacity reduction to be 50% regardless of the numbers of demons present. By comparing the Expected Total System Travel Times (ETSTT) with those obtained from the 'un-improved' Braess network (i.e., four links network) in the first

example, the occurrence of stochastic Braess' paradox can be demonstrated. Furthermore, by comparing the network travel time reliabilities between the two cases, this example demonstrates the existence of a paradoxical phenomenon in reliability measurement.

The results are summarized in table 4.2. In the presence of one or many demons and with a demand of three users, the paradox occurs since the ETSTT is always higher after the addition of the new link. This result demonstrates that stochastic Braess' paradox exists in risk-averse assignment. When the demand approximately equals three users the difference in ETSTT is greatest (and hence the paradoxical situation is at its worst). Given that the ETSTT is in fact also the total expected network costs faced by the users and since these costs can be regarded as a measure of the network reliability, it is not difficult to see just how important it is to take account of stochastic Braess' paradox in risk-averse assignment. It is worth mentioning here that failure to appreciate the effects of stochastic Braess' paradox would result in a misleading expectation that the network costs would be lower after the new link is added. For example, in the one demon case the ETSTT is approximately 20% higher after the new link has been added. Since a higher value for the total expected network costs implies a less reliable network, underestimating it by this amount would have serious implications for designing reliable transport networks.

Table 4.2: *The effect of Braess' paradox on network reliability measurement*

Braess' Network Before and After the Network is Improved when the Demand Level is 3 Users						
Number of Demons Present	Before/After Improvement	Travel Time Reliabilities				ETSTT
		Path 1-3-2	Path 2-4-2	Path 1-3-4-2	Network	
1	Before	0.50	0.50	-----	1.00	222.00
	After	0.50	0.50	0.00	0.00	266.31
2	Before	0.25	0.25	-----	0.50	233.25
	After	1.00	1.00	0.50	0.50	269.99
3	Before	0.125	0.125	-----	0.25	238.88
	After	1.00	1.00	0.25	0.25	271.54

Table 4.2 also shows the effect of ignoring stochastic Braess' paradox to the travel time reliability measures that were defined earlier. In fact, this gives rise to another type of paradox, namely the reliability paradox. For all cases of the number of demons present, the reliabilities of paths one and two are equal due to the fact that these two paths are symmetric. When link (3,4) is added to the network an additional path is created (path 1-3-4-2) and the reliability of this path varies depending on how many demons are assumed in

the model. If only one demon is present then this path has a reliability value of zero since most of the flow uses this path with the remaining flow being split equally between the other two paths. The network travel time reliability in this case is also zero which strongly suggests that the new link should not have been added at all since not alone is the new path highly unreliable but the network as whole becomes highly unreliable. This concurs with the general observation that addition of a new link can leave all users worse off than before.

For the two and three demon cases not as much flow is attracted to the new path while like before the remaining flow is split equally between paths one and two. Therefore, for these cases, the reliabilities of paths one and two actually increase when the new link is added while the new path itself has a travel time reliability value of 0.5. In these two cases the users are no worse off than before the addition of the new link in terms of the reliabilities measured although they are still worse off in terms of the overall increase in the total expected network costs. However, this situation is far from ideal since the provision of a new link should do more than merely leave users no worse off – it should at least reduce the overall network travel time reliability but it may not, as this example clearly shows. This implies that the new link should not be constructed at all.

This example illustrates the importance of accounting for stochastic Braess' paradox and the reliability paradox in reliable/robust network design. Ignoring stochastic Braess' paradox can result in a far lower total expected network cost which suggests that the network is more reliable than it actually is. In terms of reliability measures, such as the ones considered in this chapter, it is also clear that the network users can be worse off after the new link is added to the network.

4.5 CONCLUDING REMARKS

This chapter develops a link-based formulation of the risk-averse traffic assignment via the NCP approach that allows for the presence of multiple network-specific demons and therefore allows for the effect of greater network uncertainty in terms of capacity degradation to be considered. This model can avoid path enumeration which is important particularly when dealing with multiple network demons which already increases computation times due to the large number of scenarios of link damage to be considered.

This becomes even more significant when applying the model to larger real-life networks. Under this framework it was proved that there is a least one solution to this link-based problem and this problem is equivalent to the path-based one. This formulation can be reformulated into a VIP and an unconstrained minimisation program for solutions. The studies in this chapter also demonstrated the effect of assuming multiple demons and relaxing the capacity assumption to network reliability measurement through numerical studies. The results show just how important it is to accurately capture the link capacity degradations particularly when the presence of more than one demon is assumed. Relaxing this assumption also allows for more operational states to be modelled. Apart from these contributions, the studies in this chapter also demonstrated the existence of stochastic Braess' paradox and reliability paradox (analogues of Braess' paradox for expected network travel cost and reliability) in risk-averse assignment, and provided some further insights into the implications for network design if these paradoxes are ignored.

CHAPTER 5

Designing Reliable Transport Networks: The Multiple Network Demon Approach

5.1 INTRODUCTION

The ability of society to function efficiently depends to a large degree on networks of many kinds and in particular the transportation network. In previous chapters, the range and type of disruptive events have been thoroughly discussed. It is well accepted that road networks are now required to be highly reliable to ensure that users can experience smooth travel under both normal and abnormal traffic conditions. This requirement has meant that many early research efforts in the area have focused on developing measures for evaluating the network reliability. However, very few studies have progressed to the next step that is, according to the general framework proposed in Du and Nicholson (1997), that the final goal of network reliability studies must be to provide useful information on network improvement schemes to transport policy makers. Therefore, progress in the area lies in the ability to use reliability studies to determine an optimal modification of the network. This problem (without uncertainty) is commonly known as the transport Network Design Problem (NDP) and developments in the area have been described in some detail in chapter 1 of this thesis.

Several researchers have tackled the NDP under uncertainty. A brief review of these studies was provided in chapter 1. Many of these studies account for uncertain demand, uncertain supply and/or uncertain travel times. However, very few studies consider the users' route choice behaviour in a *damaged network*. This is due in part to the difficulty in collecting reliable behavioural data in a degraded network. One such approach which can model route choice behaviour in a damaged network without relying on previous data relating to capacity reducing incidents is the game theoretic approach (Bell, 2000; Bell and Cassir, 2002; and Szeto et al., 2006). The main idea and advantages of this approach have already been described in some detail in preceding chapters of this thesis.

In this chapter, game theoretic formulations of risk-averse traffic assignment are proposed to model the lower level problem (i.e., the users' route choice behaviour) of the bi-level NDP. This is achieved by formulating the lower level problem using the generalised (multiple network demon) risk-averse traffic assignment proposed in chapter 3 of this thesis. For comparison purposes, the lower level problem of the NDP is also formulated here using the Bell and Cassir (2002) OD-specific demon formulation of risk-averse traffic assignment. This chapter is the first such study to formulate the NDP under either approach. The multiple network demon approach to NDP allows the network demons to be free to select any links to damage and to impose no restriction on the number of demons or number of links to be damaged. The formulations proposed in this chapter assume a fixed travel demand. This allows for a better comparison between the proposed multiple network demon approach to NDP and the Bell and Cassir (2002) OD-specific demon approach to NDP. Furthermore, the formulations proposed here adopt the path-based approach (in contrast to chapter 4). This makes them easier to understand and allows them to be consistent with the format used in chapter 3 of this thesis and that used in Bell and Cassir (2002). Moreover, the simple test networks used in this chapter contain the same number of links as paths and therefore avoiding path enumeration is not necessarily advantageous in this case.

Similar to Bell (2000) and Bell and Cassir (2002) the total expected network costs resulting from this game act as a useful measure of network reliability. For this particular case a similar definition for network reliability to the one in Bell (2000) is given below:

Definition: A network is reliable if the total expected network costs (or Expected Total System Travel Times) are acceptable even when users are extremely pessimistic about the state of the network.

The impact of the extension is assessed on the total expected network cost and hence on the location and amount of network improvements which occur. Moreover, the effect of the assumption of 50% capacity reduction is challenged, especially when there are multiple demons involved, and when more than one demon selects a particular link to damage. In addition, the effect of using or relaxing this assumption is examined, with particular reference to the implications for reliable network design. The studies conducted provide a better insight into the demon behaviour in selecting links to damage and also compares the results obtained from OD-specific and network-specific demons.

The resulting model allows network uncertainty to be explicitly considered in the network design problem. For the upper level problem of this proposed NDP it is then necessary to decide on the appropriate form of the decision variables. The upper level problem of the NDP has been formulated with different decision variables in the past. A brief discussion of this was provided in chapter 1. In this chapter, the Continuous Network Design Problem (CNDP) is adopted. For simplicity, only link widening is considered and therefore this chapter proposes a reliable CNDP (RCNDP) model by modelling the users' route choice behaviour in a damaged network through the proposed multiple network demon approach. For comparison purposes, this chapter also proposes a RCNDP model by modelling the users' route choice behaviour in a damaged network through the OD-specific demon approach.

With the extension to the RCNDP one can consider reliable capacity improvement schemes. Two variations of the proposed RCNDP model with multiple network demons are formulated in this chapter. The first formulation considers that the network improvements are funded by a budget provided by a government while the second formulation considers that the cost of network improvements are exactly recovered by tolling links in the network. The RCNDP is formulated as a single-level mathematical program by setting the lower level problem as a set of constraints. This allows the use of existing non-linear programming techniques for solutions. A number of examples are given to: illustrate the reliable network improvement strategies of the proposed model under a number of different cases of link capacity reduction; demonstrate how only the most vulnerable links receive improvement; and show how the proposed approach can prioritise improvements to links most vulnerable to failure in a network. Moreover, the examples show the importance of using network demons as opposed to OD-specific demons particularly in relation to designing adequate capacity for the vulnerable links in the network. In all the numerical studies, the Generalized Reduced Gradient (GRG) algorithm (Abadie and Carpentier, 1969) is adopted to obtain solutions. Further details of this algorithm are given in Appendix A.

Finally, the formulation and results provided in this chapter are based upon an award winning paper (arising from an international paper competition) published by this author in the journal 'HKIE Transactions' (see Appendix B for further details). The chapter is organised as follows: Section 5.2 formulates the RCNDP model; section 5.3 describes the numerical studies, while section 5.4 provides some concluding remarks.

5.2 FORMULATION

In this section, two different variations of the proposed RCNDP model are formulated. Both models are broadly similar but differ in one key aspect. The capacity improvements determined by one model are funded by tolls imposed on the network users and furthermore all improvement costs must be exactly recovered. The capacity improvements in the other model are funded by a budget provided by the government as a lump sum.

The proposed RCNDP is in essence a bi-level problem consisting of an upper level problem and a lower level problem. The upper level problem broadly describes the objectives of a planning authority who decide how to improve the network under a certain available budget or using revenue raised through tolling links in the network. The lower level problem describes the users travel behaviour in route choice in response to improvements to links in the network made by the planning authority. In this chapter, the proposed approach used to model the lower level problem is the game theory approach leading to what is known here as a generalised risk-averse traffic assignment problem. For the purposes of obtaining solutions, the problem can also be formulated as a single level minimisation program by setting the lower level problem as a set of constraints.

5.2.1 Generalised Risk-Averse Traffic Assignment Constraints

The risk-averse traffic assignment constraints presented in this section are broadly similar to those proposed in chapter 3. However, they are restated here for clarity since a number of modifications are proposed to capture capacity improvements and link tolls.

A general transportation network with multiple links, routes, and Origin-Destination (OD) pairs is considered. There are N^{rs} homogenous players between each OD pair rs , each of whom it is assumed seeks minimum cost routes among all other alternatives (hereafter they are referred to in this chapter as the ‘users’). There are also M demons with the freedom to roam the network without restriction and with the capability to select any link to damage, including links already selected for damage by other demon(s). It is assumed that the network-specific demons are intent on causing maximum delays (and hence increased costs) to the users by damaging link(s) in the network. By so doing the demons maximise their expected pay-off which is the total network cost faced by the users. Moreover, it is

assumed that no user knows which link(s) the demon(s) will choose to damage, and none of the demons know in advance which link a user may decide to use. In general, it is assumed that each link has two costs depending on whether it is damaged or not. However, for the purposes of some of the studies this assumption is relaxed so that the link costs increase with an increasing number of demons that select a particular link to damage. Furthermore, both N^{rs} and M are assumed to be fixed. With this setting, the risk-averse user equilibrium traffic assignment problem with multiple network demons can be formulated by applying the results in Nash (1951), Bell and Cassir (2002), and Szeto et al. (2006).

Similar to Bell and Cassir (2002), Szeto et al. (2006) and previous chapters, the multiple network demon formulation of risk-averse traffic assignment comprises two main sub-problems: the user problem and the demon problem.

5.2.1.1 The User Problem

The user problem can be viewed as a non-cooperative game in which each homogenous player tries to select the route with minimum expected trip cost. The user problem can be approximated to deterministic traffic assignment when the number of homogenous users is large (Bell and Cassir, 2002). The first order condition of deterministic traffic assignment can be expressed as the following non-linear complementarity conditions:

$$h_j^{rs} \left[\sum_k g_{jk}^{rs}(\mathbf{h}) q_k - \min_d \left(\sum_k g_{dk}^{rs}(\mathbf{h}) q_k \right) \right] = 0, \quad (5.1)$$

$$\sum_k g_{jk}^{rs}(\mathbf{h}) q_k - \min_d \left(\sum_k g_{dk}^{rs}(\mathbf{h}) q_k \right) \geq 0, \quad (5.2)$$

$$h_j^{rs} \geq 0, \quad (5.3)$$

where h_j^{rs} is the flow on route j between OD pair rs ; $\mathbf{h}$ is the route flow vector; $g_{jk}^{rs}(\mathbf{h})$ represents the cost of route j between OD pair rs in scenario k based on the flow vector $\mathbf{h}$; d is the minimum expected travel cost route between OD pair rs ; q_k is the probability of scenario k occurring.

According to (5.1), if route j connecting OD pair rs carries flow ($h_j^{rs} > 0$), the corresponding expected route travel cost $\sum_k g_{jk}^{rs}(\mathbf{h})q_k$ must be equal to the minimum expected travel cost $\min_d \left(\sum_k g_{dk}^{rs}(\mathbf{h})q_k \right)$ between OD pair rs , as the term in the square brackets in (5.1) must equal zero. If route j carries no flow ($h_j^{rs} = 0$), the corresponding expected route travel cost must be greater than or equal to the minimum expected travel cost based on (5.2). Condition (5.3) is the non-negativity condition of route flows.

The route flow h_j^{rs} in (5.1)-(5.3) must satisfy flow conservation, expressed as:

$$N^{rs} = \sum_j \delta_j^{rs} h_j^{rs}, \quad (5.4)$$

where N^{rs} is the travel demand of OD pair rs ; δ_j^{rs} is the route-OD incidence indicator. $\delta_j^{rs} = 1$ if j connects OD pair rs , and $\delta_j^{rs} = 0$ otherwise.

The cost of route j between OD pair rs in scenario k , $g_{jk}^{rs}(\mathbf{h})$, in (5.1) and (5.2) is the total cost on every link that is on this route:

$$g_{jk}^{rs}(\mathbf{h}) = \sum_a \delta_{ja} (\psi t_{ak}(v_a) + \rho_a), \quad (5.5)$$

where δ_{ja} is 1 if link a is on route j , and 0 otherwise; ρ_a is the toll charged (if any) for using link a ; $t_{ak}(v_a)$ denotes the flow-dependent travel time on link a in scenario k and when multiplied by the value of time parameter ψ is converted to a monetary value. $t_{ak}(v_a)$ is defined by the Bureau of Public Roads (BPR) function in (5.6) below:

$$t_{ak}(v_a) = t_a^0 \left(1 + 0.15 \left(\frac{v_a}{c_{ak}} \right)^4 \right), \quad (5.6)$$

where t_a^0 is the free-flow travel time of link a ; $\overline{c_{ak}}$ is the capacity of link a in scenario k ; v_a is the link flow. Both $\overline{c_{ak}}$ and v_a are defined in (5.7) and (5.8) respectively:

$$v_a = \sum_{rs} \sum_j \delta_{ja} h_j^{rs}. \quad (5.7)$$

$$\overline{c_{ak}} = \lambda_{ak} (c_a + y_a) \quad (5.8)$$

Equation (5.7) states that the link flow is simply the sum of the route flows using that link. In equation (5.8) c_a is the undamaged capacity of link a , y_a is the undamaged capacity improvement and λ_{ak} is a link damage indicator which, in general, equals one if link a is not selected by any demon and equals 0.5 if a link is selected by one or more demons. In general this study considers that the damage caused by the demons results in 50% link capacity reduction (i.e., $\lambda_{ak} = 0.5$), however, the advantage of the proposed formulation is the flexibility it offers to relax this assumption when a number of demons combine together and select the same link to damage resulting in further link capacity reduction. Therefore, for the purposes of some of the studies here, λ_{ak} may equal some other values.

The scenario probability in (5.1)-(5.2), q_k , depends on the link selection probability of all demons. Let $k = (l_1, \dots, l_m, \dots, l_M)$, l_m be the link selected for damage by demon m , $p_{l_m}^m$ be the probability of demon m selecting l_m to damage, and M be the number of network-specific demons. The scenario probability q_k is defined by:

$$q_k = \prod_{m=1}^M p_{l_m}^m. \quad (5.9)$$

5.2.1.2 The Demon Problem

The demon problem describes the objectives of the demons. All demons seek the mixed strategy to maximise their individual expected pay-offs or the total expected network cost to the users. The problem for multiple network demons can be written as the following non-linear complementarity conditions:

$$p_{l_m}^m \left\{ \begin{aligned} & \max_{w_m} \sum_{l_1=1}^n \dots \sum_{l_q=1, q \neq m}^n \dots \sum_{l_M=1}^n \left(\prod_{i=1, i \neq m}^M p_{l_i}^i \right) \left[\sum_j g_{j, l_1, \dots, w_m, \dots, l_M}^{rs}(\mathbf{h}) h_j^{rs} \right] \\ & - \sum_{l_1=1}^n \dots \sum_{l_q=1, q \neq m}^n \dots \sum_{l_M=1}^n \left(\prod_{i=1, i \neq m}^M p_{l_i}^i \right) \left[\sum_j g_{j, l_1, \dots, l_m, \dots, l_M}^{rs}(\mathbf{h}) h_j^{rs} \right] \end{aligned} \right\} = 0, m=1, \dots, M, l_m=1, \dots, n, \quad (5.10)$$

$$\left\{ \begin{aligned} & \max_{w_m} \sum_{l_1=1}^n \dots \sum_{l_q=1, q \neq m}^n \dots \sum_{l_M=1}^n \left(\prod_{i=1, i \neq m}^M p_{l_i}^i \right) \left[\sum_j g_{j, l_1, \dots, w_m, \dots, l_M}^{rs}(\mathbf{h}) h_j^{rs} \right] \\ & - \sum_{l_1=1}^n \dots \sum_{l_q=1, q \neq m}^n \dots \sum_{l_M=1}^n \left(\prod_{i=1, i \neq m}^M p_{l_i}^i \right) \left[\sum_j g_{j, l_1, \dots, l_m, \dots, l_M}^{rs}(\mathbf{h}) h_j^{rs} \right] \end{aligned} \right\} \geq 0, m=1, \dots, M, l_m=1, \dots, n, \quad (5.11)$$

$$p_{l_m}^m \geq 0, \quad (5.12)$$

where n is the number of links in the network; $g_{j, l_1, \dots, l_m, \dots, l_M}^{rs}(\mathbf{h})$ is the cost of route j between OD pair rs in scenario $k = (l_1, \dots, l_m, \dots, l_M)$; l_m and M follow earlier definitions; w_m is the link selected for damage by demon m that gives the demon maximum expected pay-off; i and q are indices for demons. $p_{l_m}^m$ is the probability of demon m selecting l_m to damage, must satisfy:

$$\sum_{l_m} p_{l_m}^m - 1 = 0, \quad (5.13)$$

and be non-negative by definition as shown in (5.12).

Conditions (5.10)-(5.11) represent the necessary and sufficient conditions of the Nash equilibrium of a non-cooperative, mixed strategy game for the demons (see Nash 1951). The second square bracket term in (5.10) is the total network cost in scenario $(l_1, \dots, l_m, \dots, l_M)$. The second term is the total expected network cost when demon m selects link l_m to damage (or the expected pay-off to demon m when it selects link l_m to damage) and considers the link selection strategies of all other demons. The first term is the maximum expected pay-off to demon m or the maximum total expected network cost. When demon m selects a particular link to damage with some probability ($p_{l_m}^m > 0$) then the term inside the braces must be zero and the maximum expected pay-off to that demon (the first term inside the braces of (5.10)) must be equal to the total expected network cost (the second term inside the braces of (5.10)) when the demon selects that link to damage. If

a particular demon does not choose a particular link l_m to damage ($p_{l_m}^m = 0$), then the maximum expected pay-off to the demon must be greater than or equal to the total expected network cost when the demon selects link l_m to damage according to (5.11).

5.2.2 Design Constraints

For practical reasons such as, for example, the unavailability of development land in urban areas, there is often an upper bound on the amount of capacity expansion that a highway link can undergo. The model is further constrained to return only non-negative values of capacity expansion since capacity reduction is not considered here and in general it is rare in practice. Mathematically these constraints may be expressed as:

$$\overline{c_{ak}} \leq u_a, \quad (5.14)$$

$$y_a \geq 0. \quad (5.15)$$

Equation (5.14) limits the total capacity of a link in any scenario to be not greater than the maximum allowable capacity of that link, u_a , while condition (5.15) is the non-negativity constraint of capacity enhancements.

5.2.3 Improvement Cost Functions and Budgetary Constraint

The total improvement cost K_t is computed by summing the corresponding improvement costs of all links.

$$K_t = \sum_a m_a(y_a), \quad (5.16)$$

where:

$$m_a(y_a) = b_{a,0} t_a^0 (y_a)^{b_{a,1}}, \forall a, \quad (5.17)$$

$m_a(\cdot)$ is the improvement cost function of link a with parameters $b_{a,0}$ and $b_{a,1}$. Equation (5.17) describes how the improvement cost of a link is proportional to the capacity enhancement y_a and its length $b_{a,0} t_a^0$.

The budgetary constraint can be expressed as:

$$B_t - K_t = 0, \quad (5.18)$$

where B_t is the network improvement budget allocated by the government to fund the cost of the network improvements. Equation (5.18) requires that the budget provided by the government is entirely spent to fund link capacity improvements to the network.

5.2.4 Tolling Conditions, Revenue Functions and the Cost-Recovery Constraint

For practical or political reasons a planning authority may decide to set an upper limit on the maximum allowable toll which can be charged to the users:

$$\rho_a \leq \rho_{\max}, \quad (5.19)$$

$$\rho_a \geq 0. \quad (5.20)$$

Condition (5.19) constrains the toll set to be less than a certain amount while condition (5.20) is added for completeness to ensure that tolls charged are non-negative.

The total toll revenue, R_t , can be expressed as a function of the equilibrium link flows and tolls:

$$R_t = \sum_a \rho_a (nv_a), \quad (5.21)$$

where n is a parameter to convert hourly link flows to annual link flows.

Exact cost recovery is considered in one of the proposed models to ensure that all link capacity improvements are funded by tolls charged to the network users. Exact cost recovery occurs when the total revenue generated through tolling exactly covers the total link capacity improvement cost. The exact cost recovery (ECR) constraint may be expressed as:

$$R_t - K_t = 0, \quad (5.22)$$

5.2.5 The Objective Function

For simplicity, an objective function with a straightforward interpretation is adopted. This study considers the minimisation of expected total system travel time (ETSTT). ETSTT describes the expected total travel time spent by users in the network and is very straightforward to compute given the link flows and damage probabilities. Mathematically, it can be expressed as:

$$ETSTT = \sum_a \sum_k \left(v_a (t_{ak}(v_a)) q_k \right), \quad (5.23)$$

where v_a is the link flow as given in equation (5.7); $t_{ak}(v_a)$ is the link travel time in scenario k given in equation (5.6); q_k is the probability of scenario k occurring. The product $\left(v_a (t_{ak}(v_a)) q_k \right)$ is evaluated in each scenario for every link in the network.

5.2.6 The Multiple Network Demon RCNDP Formulations

5.2.6.1 Model A1: RCNDP with Budget

The RCNDP model with budget can then be formulated as a single level minimisation program by setting the lower level problem as a set of constraints:

$$\min_{\mathbf{h}, \mathbf{y}, \mathbf{p}_{l_m}^m} ETSTT = \sum_a \sum_k \left(v_a (t_{ak}(v_a)) q_k \right)$$

subject to: network demon risk-averse traffic assignment constraints (5.1)-(5.13),
 design constraints (5.14)-(5.15),
 improvement cost and budgetary constraints (5.16)-(5.18),

where $\mathbf{h}, \mathbf{y}$ and $\mathbf{p}_{l_m}^m$ are the vectors of route flows, capacity improvements and the probabilities of demon m selecting l_m to damage.

5.2.6.2 Model A2: RCNDP under Exact Cost Recovery

The RCNDP model under exact cost recovery can be formulated as a single level minimisation program by setting the lower level problem as a set of constraints:

$$\min_{\mathbf{h}, \mathbf{y}, \mathbf{p}_{l_m}^m, \boldsymbol{\rho}} ETSTT = \sum_a \sum_k \left(v_a \left(t_{ak} \left(v_a \right) \right) q_k \right)$$

subject to: network demon risk-averse traffic assignment constraints (5.1)-(5.13),
 design constraints (5.14)-(5.15),
 improvement cost constraints (5.16)-(5.17),
 tolling, revenue and exact cost recovery constraints (5.19)-(5.22),

where $\mathbf{h}, \mathbf{y}, \boldsymbol{\rho}$ and $\mathbf{p}_{l_m}^m$ are the vectors of route flows, capacity improvements, link tolls and the probabilities of demon m selecting l_m to damage.

5.2.7 The OD-Specific Demon RCNDP Formulation

For completeness, and to provide a comparison with the multiple network-specific demon formulations of RCNDP proposed in section 5.2.6, this chapter also proposes the following RCNDP model under the OD-specific formulation of risk-averse assignment.

The RCNDP model with OD-specific demons with a budget can then be formulated as a single level minimisation program by setting the lower level problem as a set of constraints:

$$\min_{\mathbf{h}, \mathbf{y}, \mathbf{p}_{l_m}^m} ETSTT = \sum_a \sum_k \left(v_a \left(t_{ak} \left(v_a \right) \right) q_k \right)$$

subject to: OD-specific demon risk-averse traffic assignment constraints (as proposed in Bell and Cassir (2002)),
 design constraints (5.14)-(5.15),
 improvement cost and budgetary constraints (5.16)-(5.18),

where $\mathbf{h}, \mathbf{y}$ and $\mathbf{p}_{l_m}^m$ are the vectors of route flows, capacity improvements and the probabilities of demon m selecting l_m to damage.

The proposed formulations described in sections 5.2.6 and 5.2.7 are single-level minimisation programs, which can be solved by many existing optimisation algorithms. In the numerical sections which follow, the RCNDP formulations are solved using the Generalized Reduced Gradient (GRG) algorithm proposed by Abadie and Carpentier (1969). The detailed algorithmic steps are described in Lo and Szeto (2004) and in Appendix A of this thesis. It is acknowledged here that the GRG algorithm is not a global optimisation algorithm. Therefore, it is not easy to know whether the objective value attains its global minimum when the algorithm stops. As part of a future study it may be necessary to develop some global optimisation techniques tailored for this problem.

5.3 NUMERICAL STUDIES

In the following two examples, a number of important aspects of the proposed RCNDP model are illustrated including the design approach when many operational states of a network link are considered and the impacts on the design of assuming network demons instead of OD-specific demons.

5.3.1 The Effect of Multiple Network Demons in Reliable Transport Network Design

This study considers the effect of using multiple network demons in the reliable network design problem. In previous chapters of this thesis, a number of disadvantages associated with assuming OD-specific demons in reliability analysis/reliable network design were given. One such drawback in assuming only one OD-specific demon to be present in each OD pair relates to the amount of capacity reduction which can occur when a link is selected for damage by the demon. The capacity reduction usually assumed when one OD-specific demon selects a link to damage is half of its initial value. The rationale for such an assumption is based on a link having two operational states – congested and un-congested. Even when a number of OD-specific demons select the same link to damage at the same time, the capacity is still reduced by half. As a result, fewer operational states of the network links can be modelled. It can be possible for more than one OD demon to be in a particular OD pair when the OD pair concerned shares a common link or links with another OD pair. Even when this scenario does occur, the OD demons cannot combine together to degrade the link any further, i.e., the capacity of the link does not degrade any

further. This is clearly not realistic since it is very likely that, in practice, two different capacity degrading incidents occurring on the same link will contribute further to the capacity reduction and, if the capacity reduction is severe enough, result in partial or complete closure of the link. The proposed formulation poses no such restrictions and allows one to consider multiple network demons which are free to roam the network selecting links to damage including links selected for damage by other demon(s). Therefore, when two capacity reducing incidents occur on the same link in the network (for example, when two demons select the same link to damage) then it is possible under the proposed approach to assume that the capacity reduces even further.

In this study it is shown that many more operational states can be modelled under the proposed network demon approach to reliable network design. This is achieved by assuming that the capacity can decrease linearly and non-linearly with respect to an increasing number of demons selecting the same link to damage. Of course other assumptions are also possible but these ones are adopted just to illustrate the effect. This example then considers the effect of these different capacity reducing assumptions on the link selection probabilities of the demons and hence on the expected network costs faced by the network users which are also the maximum pay-offs to the network demons. By enabling the model to consider higher or lower values of capacity reduction, this approach can also consider higher and lower maximum expected network costs that can be faced by the users and hence the resulting network design (link improvement strategy) will be different than that of the conventional approach.

Furthermore, the link selection strategies of the demons may be different resulting in possible different locations for the vulnerable or critical links in the network. The reason for this is as follows: the network demons' objective is to maximise the network costs faced by the user and so if both demons can combine to cause greater capacity reduction on the same link then this is the most likely strategy to be chosen. However, the users' objective is to minimise their expected travel costs and so they seek to offset this effect by switching routes. The resulting equilibrium route flows of the users and link selection probabilities of the demons determine which network links are critical (i.e., which links when damaged cause the greatest disruption to the users) and by how much each link should be expanded.

This particular study adopts the proposed RCNDP model (Model A2: RCNDP under Exact Cost Recovery) to obtain results. The simple test network used is shown in figure 5.1 and consists of one OD pair with two parallel links. For this study, the presence of two network demons is assumed. This example deliberately avoids using a network with greater than one OD pair so that the results are not complicated by other factors such as the network demons choosing to damage links on only one particular OD pair. Furthermore, the focus here is to highlight the capability of the proposed formulation to model more operational states and to show how this has an impact on the amount and location of the link capacity improvements.

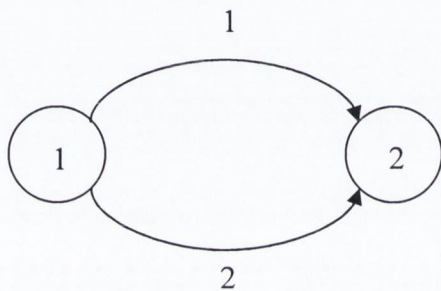


Figure 5.1: Test network 1

The following parameters chosen for illustrative purposes also apply:

Free-flow travel times: $t_1^0 = 5$ minutes; $t_2^0 = 10$ minutes

Initial undamaged link capacities (without improvements): $c_1^0 = c_2^0 = 4000$ vph

Maximum allowable link capacity: $u_1 = u_2 = 10000$ vph

Value of time: $\psi = \text{€}30/\text{hour}$

Factor to convert hourly flows to annual flows: $n = 8760$

Improvement cost parameters: $b_{1,0} = \text{€}43,000/\text{veh.}$; $b_{1,1} = \text{€}1/\text{veh}$

Demand parameter: $N_1^{12} = 6000$ vph

Maximum allowable toll: $\rho_{\max} = \text{€}10$

The above parameters can be broadly classified into two different types: network parameters and financial parameters. Estimation of the network parameters such as the free-flow travel time and link capacity are relatively straightforward since they are related to the geometry of the network and the speed limit among other factors. For the networks used in these examples no data is available as the network parameters are chosen purely for illustrative purposes. The value of time parameter is generally estimated from surveys of

sections of the population in which respondents are generally asked to put a monetary value on how much an hour of their time is worth. In general this value is related to the income of the person being interviewed. Therefore the final figure calculated for the ‘value of time’ is generally an average value since it will of course be higher for some sections of the population and lower for others. Other parameters such as the improvement cost parameter can be estimated from previous construction projects that have taken place in the particular country of interest. The types of link capacity reduction which occurs when a link is selected for damage by the network demons are given in table 5.1. Two different cases of this are considered: Case one (two) capacity reduction is linear (non-linear) with respect to the number of demons.

Table 5.1: The damaged link capacities

	No. of demons selecting a link to damage	
	1	2
Link Capacity (vph) Case 1: Linear Capacity Reduction	$\frac{3}{4}(c_{lk}^0 + y_a)$	$\frac{1}{2}(c_{lk}^0 + y_a)$
Link Capacity (vph) Case 2: Non-linear Capacity Reduction	$\frac{(c_{lk}^0 + y_a)}{2}$	$\frac{(c_{lk}^0 + y_a)}{4}$

The proposed Model A2: RCNDP under Exact Cost Recovery is solved by using the Generalized Reduced Gradient (GRG) algorithm proposed by Abadie and Carpentier (1969) and values are obtained for the decision variables which are the equilibrium route flows, link selection probabilities of the demons, optimal capacity improvements and tolls. This procedure is repeated for two different cases of capacity reduction: Case one (two) capacity reduction is linear (non-linear) with respect to the number of demons and these results are compared to the base case, which is, that the capacity of a link is always half of its initial value regardless of whether it is selected for damage by one or more demons at the same time. The results are discussed below.

Table 5.2 and table 5.3 show the equilibrium route flows and the demons’ link selection probabilities.

Table 5.2: Route Flows for the different cases of capacity reduction

Route Flows (vph)			
Route	Base Case	Case 1	Case 2
1	4100.50	3640.99	3511.31
2	1899.50	2359.01	2488.59

Table 5.3: Demon link selection probabilities for the different cases of capacity reduction

Demon Link Selection Probability						
	Base Case		Case 1		Case 2	
	Demon A	Demon B	Demon A	Demon B	Demon A	Demon B
Link 1	1.00	0.00	1.00	1.00	0.81	0.83
Link 2	0.00	1.00	0.00	0.00	0.19	0.17

In table 5.2 it can be observed that the different cases of capacity reduction result in quite different route flow patterns. Case 2 (non-linear capacity reduction) results in the users choosing routes somewhat more evenly than the base case although there is still a higher flow on route 1 since it is the minimum cost route. These different flow patterns arise due to the different ‘behaviour’ of the network demons under the different capacity reduction assumptions. For example, from table 5.3, under case 2 (non-linear capacity reduction) the demons combine to damage link 1 (with probability of approximately 0.8 each) since this link attracts the greatest number of network users and because they can further damage this link by their combined effect (since the capacity decreases non-linearly with respect to an increasing number of network demons). If this is compared to the base case, quite a different result can be seen. In the base case the link capacity reduction will always be 50% regardless of the number of demons choosing a particular link to damage. Hence, the demons select both link 1 and link 2 with a probability of 1 since by doing so they can maximise their pay-offs which are the maximum expected network costs faced by the users. Case 1 (linear capacity reduction) is broadly similar to that of case 2 for the same reasons as mentioned above. However, for case 1 both demons select link 1 with a probability of 1 since their combined effect is not as great as that of the non-linear case and so to maximise their pay-offs they need to concentrate on selecting link 1 to damage. The link selection probabilities of the demons can be used to generate overall scenarios of link damage. For clarity these scenarios are described in table 5.4 and the probability of them occurring in table 5.5.

Table 5.4: Description of the scenarios of link damage

Description of the Scenarios of Link Damage		
Scenario	Link 1	Link 2
1	Demon A, Demon B	0
2	0	Demon A, Demon B
3	Demon A	Demon B
4	Demon B	Demon A

Table 5.5: Scenario probabilities for the different cases of capacity reduction

Scenario Probability			
Scenario	Base Case	Case 1	Case 2
1	0.00	1.00	0.67
2	0.00	0.00	0.03
3	1.00	0.00	0.14
4	0.00	0.00	0.16

The link selection probabilities of the demons and the resulting link flow travel patterns of the users result in different values of the ETSTT for all three cases of link capacity reduction. It is worth noting that the ETSTT corresponds to the total expected network costs faced by the users or alternatively the maximum expected pay-offs of the network demons. It is these values which have a major impact on the amount of capacity improvements that the model determines. Table 5.6 shows the ETSTT values for the different cases of link capacity reduction. The highest value of ETSTT results from case 2 (non-linear capacity reduction). Compared to the base case, the ETSTT is approximately 16% higher for case 2. This is to be expected since the demons can combine together to cause the greatest amount of damage under this assumption. Since the purpose of the game theoretic approach to reliable network design is to produce the most pessimistic outcome then this assumption seems to be the most appropriate. This is particularly relevant when one considers that in a real life scenario two capacity reducing incidents occurring on the same link would result in higher travel times due to their likely combined effect in causing further disruption. The base case produces approximately a 10% higher value for ETSTT than case 1 (linear capacity reduction). However, this is to be expected since only two network demons are considered here. For the cases where there are many more network demons present one would expect that the linear capacity reduction assumption would in general produce a higher value for ETSTT than the base case.

Table 5.6: ETSTT values for the different cases of capacity reduction

ETSTT		
Base Case	Case 1	Case 2
67322.39	60724.65	78240.96

Table 5.7 and table 5.8 show the link capacity improvements and corresponding link tolls for all cases of link capacity reduction. Referring to table 5.7, it can be observed that case 2 (non-linear capacity reduction) results in a far higher capacity improvement for link 1 and a significant capacity improvement for link 2. This result is to be expected since the demons combine under this case and degrade the capacity of link 1 by the greatest amount.

In fact the proposed approach clearly identifies that link 1 (the critical link in this case since it attracts the most flow and has a high probability of being selected for damage by both demons) should receive the greatest priority under the improvement scheme. In fact when non-linear capacity reduction is assumed, the resulting network improvements to link 1 are almost six times greater than those recommended by case 2 (linear capacity reduction). Furthermore, under the assumption of non-linear capacity reduction, link 2 is also improved by 827vph. It would seem from these results that the assumption of non-linear capacity reduction results in the most reliable and realistic network improvement scheme. This emphasises the importance of the proposed RCNDP model where the presence of network demons is possible since a non-linear capacity reduction is clearly not possible under the OD-specific demon approach to reliable network design.

Table 5.7: Link capacity improvements for the different cases of capacity reduction

Link Capacity Improvements (vph)			
Link	Base Case	Case 1	Case 2
1	949.75	609.73	3840.75
2	0.00	0.00	827.50

Table 5.8: Link tolls for the different cases of capacity reduction

Link Toll (€)			
Link	Base Case	Case 1	Case 2
1	1.14	0.82	5.37
2	0.00	0.00	3.26

Table 5.8 shows the link tolls which are charged to the users to fund the capacity improvements. This example considers exact cost recovery and assumes that the tolls will be charged for a five year period to fund the improvements. From table 5.8, it can be observed that under case 2 the highest tolls are charged and also a toll is charged on link 2 to fund the major capacity expansion on link 1 and the improvements to link 2. The other two cases of capacity reduction result in far lower tolls being charged but as can be seen from the results the capacity expansions are lower and possibly result in a less reliable network design than that produced by case 2 (non-linear capacity reduction).

From this study it is clear that the number of demons which are assumed and the type of capacity reduction they cause have important implications for network reliability and reliable network design. In fact it is quite clear that the assumptions regarding capacity degradation may have more significance than those regarding how many demons to select.

However, when both of these assumptions are relaxed, they produce significantly higher network cost for users to face and are worthy of further study due to the implications for the design of reliable networks. Under the proposed approach it is at least possible to define a greater number of demons in the network than before where the number of demons was dictated by the number of OD pairs in the network. Moreover, it is possible under the proposed approach for the demons to combine together and cause greater disruption to the users. This is important considering that the original purpose or motivation of the game theory approach is to produce a cautious and reliable network design (based on the maximum expected network costs faced by users who are pessimistic in the estimate of their travel costs).

5.3.2 OD-Specific vs. Network-Specific Demons in the Design of a Multi-Commodity Network

The purpose of this study is to compare the effects of the network demons approach and the OD demons approach to the RCNDP in a multi-commodity (multi-OD pair) network. In the first example only one OD pair was considered which does not allow one to consider more than one OD-specific demon and does not allow for other possibilities such as all network demons damaging one OD pair in the network only. This example considers these effects by comparing the two approaches on a multi-OD pair network. To have a fair comparison between the two types of demons, it is assumed that the link capacity is always reduced by 50% when a link is selected for damage by one or more of the demons. This assumption is necessary here since the OD-specific demons cannot always combine to cause greater damage to a link in terms of capacity reduction. Furthermore, only two network-specific demons are considered to be active in the network since there are only two OD-specific demons present (one for each OD pair of the test network). The network used here is shown in figure 5.2 and consists of two OD pairs: OD pairs (1,3) and (2,3). OD pair (1,3) connects origin 1 to destination 3 via two routes: route 1 which uses link 1 only and route 2 which uses links 2 and 3. OD pair (2,3) connects origin 2 to destination 3 via one link only – link 3.

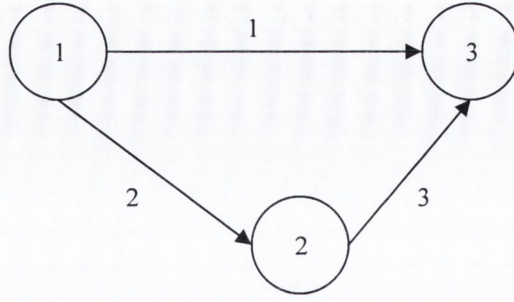


Figure 5.2: Test network 2

The following parameters chosen for illustrative purposes also apply:

Free-flow travel times: $t_1^0 = 10$ minutes; $t_2^0 = 3$ minutes; $t_3^0 = 5$ minutes

Initial undamaged link capacities (without improvements): $c_1^0 = c_2^0 = c_3^0 = 4000$ vph

Damaged link capacities (with improvements):

- $\lambda_{ak} = 1$ if link a is not selected by any demon(s) in scenario k ,
- $\lambda_{ak} = 0.5$ if link a is selected by any number of demon(s) in scenario k ,

where λ_{ak} is a link-damage incidence indicator.

Maximum allowable link capacity: $u_1 = u_2 = u_3 = 10000$ vph

Improvement cost parameters: $b_{1,0} = \text{€}43,000/\text{veh.}$; $b_{1,1} = \text{€}1/\text{veh}$

Demand parameters: $N_1^{13} = 5000$ vph, $N_1^{23} = 1000$ vph

Budget: $B_i = \text{€}946\text{m}$

In this study, a certain scenario is considered where the demand on OD pair 1-3 is high and the demand on OD 2-3 is relatively low, relative to the initial capacities of the links. To obtain results, this example first solves the proposed Model A1: RCNDP with a budget (with network demons) and compares this to the RCNDP model with a budget (with OD-specific demons). Both models are solved using the Generalized Reduced Gradient (GRG) algorithm proposed by Abadie and Carpentier (1969). The results are discussed below.

The route flow patterns, link flow patterns and link selection probabilities of the demons for case 1 (RCDNP with network demons) and case 2 (RCDNP with OD-specific demons) are shown in tables 5.9, 5.10 and 5.11 respectively. The network demons select links 1 and 3 to damage with a probability of 1 each since these are the links which carry the most flow (table 5.10) and by damaging these links the demons maximise the expected network costs faced by the users. Neither of the network demons damage link 2 because this link

forms part of route 2 in the network and in any event the users have already faced maximum disruption on this route since link 3 has been selected for damage. Moreover, in this case the network demons do not combine to damage any link in the network since it is assumed in this example that the capacity reduction is always 50% regardless of the number of demons present (to allow for a fair comparison to the OD-specific demons model). Therefore, the network demons cannot achieve any higher pay-off by damaging the same links. The OD-specific demons do not have the same flexibility as the network demons. They are confined to damaging links contained in their own OD pairs (see table 5.11). Therefore the demon on OD 2-3 is confined to always damage only link 3. The OD demon on OD 1-3 also chooses to damage link 3 since this link carries the highest flow on this OD pair (see table 5.10). Unlike the network demons, the OD-specific demons cannot roam the network freely. As a result, the vulnerable or critical links in this network are defined differently depending on whether one assumes network-specific demons or OD-specific demons in the formulation. The resulting ETSTT values are also quite different. The network-specific demons approach produces an ETSTT of 66450 approximately while the OD-specific demons approach produces a lower value of 57086 approximately. This difference is quite significant (approximately 16%) and shows that the network demons approach produces higher maximum expected network costs to the users. Since the network design is determined based on this value (and the available budget), a capacity improvement scheme results which is more capable of withstanding the worst case scenarios of network reliability. The locations of the improvements will also differ due to the different link selection strategies of the different types of demons.

Table 5.9: Route flows for the network and OD-specific demons

Route Flows (vph)		
Route	Network Demon	OD-Specific Demon
1	3047.30	1480.26
2	1952.70	3519.74
3	1000.00	1000.00

Table 5.10: Link flows for the network and OD-specific demons

Link Flows (vph)		
Link	Network Demon	OD-Specific Demon
1	3047.30	1480.26
2	1952.70	3519.74
3	2952.70	4519.74

Table 5.11: Link selection probabilities of the network and OD-specific demons

Link Selection Probabilities of the Demons				
Link	Network Demon		OD-Specific Demon	
	A	B	OD 1-3	OD 2-3
1	0	1	0	n/a
2	0	0	0	n/a
3	1	0	1	1

Table 5.12 shows the link capacity improvements for case 1 (RCDNP with network demons) and case 2 (RCDNP with OD-specific demons). The results show contrasting network improvement strategies.

Table 5.12: Link capacity improvements for the RCNDP with different types of demon

Link Capacity Improvements (vph)		
Link	Network Demon	OD-Specific Demon
1	2018.35	54.64
2	605.51	1988.23
3	0.00	3097.78

In both cases the full budget is spent on the network improvements – this is due to the constraint included in this model which requires that the budget is fully used to fund the improvements. Case 2 shows how the OD-specific demon model concentrates the improvements on links 2 and 3. Since these links are the shortest links in the network, the improvement budget allows for significant capacity expansions. On the other hand, for case 1, the network demons approach is to concentrate primarily on link 1. By expanding the capacity of this longer link the budget does not allow for the same scale of capacity expansion but on the other hand widening this link lowers the travel time on route 1 and makes it a more attractive option for users when choosing the best route travelling from origin 1 to destination 3. Overall, quite different improvement schemes can be observed between the two cases and this underlines the importance of the type of demons to assume. The network demons RCNDP model offers a more flexible approach for a number of reasons. Firstly, the model allows for any number of demons per OD pair; secondly, they have the ability to roam the network freely; thirdly, they can combine together to reduce the capacity of a particular link, and finally they allow many more operational states to be modelled. Overall this study demonstrates the superiority of the network demon approach to reliable network design.

5.3.3 Discussion

The results from the two numerical studies carried out using the sample networks demonstrate the advantages of the network demon approach to reliable transport network design. The results presented here are somewhat preliminary since this is the first study to consider such an approach. It is expected that similar findings will be observed when this approach is applied in larger more realistic networks. In fact the advantage of the network demon approach is that it allows us the flexibility to choose as many demons as is appropriate. This contrasts to the OD demon approach which limits the number of OD demons to be exactly equal to the number of OD pairs. This restriction of the OD demon approach could produce absurd results if applied to, for example, the Chicago Regional Network which consists of 3,202,310 OD pairs and 39,018 links. Clearly this very large number of demons to damage considerably fewer links than OD pairs would be inappropriate here. For this reason it is this author's view that the network demon approach is more appropriate for application to realistic networks since it allows the decision maker the discretion to specify the number of demons to assume. Furthermore in certain other networks there may exist many more links than OD pairs in which case the proposed approach can allow for more demons if necessary thereby resulting in higher expected network costs and hence a better 'worse-case scenario' design.

5.4 CONCLUDING REMARKS

This chapter proposes a formulation of the reliable continuous network design problem by modelling the lower level problem of the NDP via the multiple network demon approach. The resulting model allows one to explicitly consider uncertainty in the transport network design problem and does not rely on data on user behaviour in a damaged network. Furthermore, the proposed network demon approach to modelling the users' route choice behaviour in reliable network design allows for the presence of multiple network-specific demons and therefore allows for the effect of greater network uncertainty in terms of capacity degradation to be considered. The framework of the RCNDP permits modelling of many more operational states than could be previously considered and determining the links most vulnerable to failure and allowing these links to receive priority in the improvement scheme thereby resulting in a reliable network design. By basing the network design on the expected travel costs of the users when confronted by supply uncertainty,

this approach introduces improvements which cater for the worst case scenarios of network reliability. This chapter formulated the RCNDP as a single-level optimisation program, solved it through the GRG solution algorithm, and through numerical examples showed the importance of the type of demon to use and the amount of capacity reduction to assume. Overall, the results show just how important it is to accurately capture the link capacity degradations in reliable transport network design.

CHAPTER 6

Conventional and Reliable Formulations of the Discrete Network Design Problem over Time

6.1 INTRODUCTION

Over time there will be inevitable changes in the travel demand patterns in a transport network. To capture these changing demands and design a gradually upgraded network, it is necessary to introduce the time dimension into the conventional (i.e., without uncertainty) transport Network Design Problem (NDP). Previous efforts have extended the continuous network design problem to cater for such considerations. However, the continuous problem does not always indicate improvements that are practical. Therefore, in first part of this chapter, a model of the time-dependent discrete network design problem (DNDP-T) is proposed. The importance of the model is illustrated by two numerical studies, which demonstrate the importance of introducing the time dimension to the NDP and the importance of introducing discrete decision variables to the NDP over time. An example of the NDP under two cost-recovery principles is also provided to illustrate the application of the model.

Chapter 5 of this thesis proposed a formulation of the reliable continuous network design problem (RCNDP). However, this approach modelled the capacity improvements as a continuous variable and furthermore considered the improvements to be static or one-off events without taking into account changing travel demands or changing link damage probabilities over a planning period. Therefore, to design a reliable transport network properly it is necessary to take into account users reactions to incidents which cause disruption in different parts of the network over time. In order to design a reliable network to deal with such time-varying events it is necessary to formulate the Reliable Network Design Problem (RNDP) with the time dimension explicitly considered. Furthermore, for practical reasons, it is imperative that any improvement scheme be possible to implement. Therefore, in second part of this chapter a time-dependent reliable discrete network design problem (RDNDP-T) is formulated, where the capacity improvements are modelled as

discrete variables and where the link flow is assigned to paths in the network under risk-averse traffic assignment. For simplicity, in this particular case, the model of risk-averse traffic assignment with OD-specific demands and fixed demands is extended to consider time-varying travel demands and gradual network upgrades and is used to assign the link flow to paths in the network in the proposed RDNDP-T. Later in this chapter some aspects of the proposed RDNDP-T model are demonstrated through a numerical study.

Finally, the formulations and results provided in this chapter, relating to the time-dependent conventional and reliable formulations of the discrete network design problem, are broadly based upon two papers published by this author in the journals 'HKIE Transactions' and 'Journal of the Eastern Asia Society of Transportation Studies'. In addition, a conference version of the paper relating to the conventional DNDP-T model received an award locally and was short-listed for an international paper award competition (see Appendix B for further details). The remainder of this chapter is organised as follows. Section 6.2 outlines the importance of the network design problem over time and in particular the discrete case proposed here. Section 6.3 describes the NDP over time. Section 6.4 formulates the conventional DNDP-T model. Section 6.5 briefly depicts the solution method. Section 6.6 describes the numerical studies to illustrate three aspects of the proposed DNDP-T model: the importance of introducing discrete capacity constraints to NDP, the importance of introducing the time dimension to NDP, and cost recovery over time. Section 6.7 describes the proposed reliable discrete network design problem over time. Section 6.8 formulates the RDNDP-T model. Section 6.9 describes a numerical study to demonstrate the key aspects of this approach. Finally Section 6.10 provides some concluding remarks.

6.2 BACKGROUND TO DNDP-T

The planning, construction and operation of transport infrastructure projects is very often an event that takes place over time subject to increasing travel demands and available resources (funding, land-use etc). Indeed it is common for new roads to be constructed and gradually upgraded over time. Very often these new additions to the network are financed through imposing tolls on the network users over the planning horizon. The question of timing and phasing of transport project implementation are therefore important considerations. When is the correct time to improve the network and who should pay for

it? Should the full-blown network be built now and future generations allowed to benefit from it or should its upgrade be progressively implemented? How about the financial arrangement over time? It is important to consider these issues in the context of planning, construction and operation of transport infrastructure projects. This analysis belongs to the transport Network Design Problem (NDP).

The transport NDP is to improve a transport network in a manner that makes efficient use of resources, while also taking account of the travel choice behaviour of potential network users or trip makers. Comprehensive reviews of the NDP have been carried out in the past, for example: Yang and Bell (1998). Further details on the NDP have been provided in chapter 1. Traditionally the NDP is formulated as a bi-level problem with the upper level modelling the supply decisions undertaken by a planning authority to satisfy some objective, while the lower level models the travel choice decision of potential trip makers. Unlike the network supply, the potential trip makers are not directly under the influence of the planning authority but they react to the changes in transport networks by altering their travel decisions.

The NDP can be briefly classified into two catalogues: the Continuous Network Design (CNDP) and the Discrete Network Design Problem (DNDP). The former only includes continuous decision variables whereas the latter can contain *both continuous and discrete decision variables*. Moreover, the former takes the network topology as given and is concerned with optimizing the network parameters whereas the latter is concerned with the network topology in addition to optimizing network parameters. Indeed, the CNDP can be considered an approximation to the DNDP when the network topology is fixed, and has received considerably more attention in the literature (see chapter 1 for further details) possibly due to the computational benefits arising from using a continuous approximation to the discrete problem especially for large transport networks. Abdulaal and LeBlanc (1979) concluded from their computational results that a continuous approximate solution could be obtained at about 25% of the cost of obtaining an exact discrete solution. Despite the computational difficulty associated with the DNDP, there are still some examples in the literature (see chapter 1 for further details).

However, these efforts consider the road network improvements to be one-off events and do not consider gradual improvements to the network or changing travel demands over a planning period. More recently, Lo and Szeto (2004) posed the CNDP with the time

dimension explicitly considered so as to allow for gradual improvements to the network and changing travel demand patterns over a planning period as well as analyzing financial arrangements and determining time-dependent tolls. Their extension to include the time dimension is known as Time-dependent CNDP (CNDP-T). The framework of the CNDP-T has allowed many other possibilities and issues to be explored that were previously impossible under the static or time-independent framework. A number of examples of previous work on CNDP-T were given in chapter 1.

As mentioned already, the computational advantages of using a continuous approximation to the discrete problem can be substantial. However, the disadvantage is that it may be difficult in practice to translate the continuous solution into capacity improvements which can be implemented. In many cases this problem is exacerbated when the continuous solution determines relatively small improvements on a large number of links or many small improvements on the same link over many years. These small improvements are not realistic to implement due to the fact that road capacity is mainly determined by the number of lanes and lanes are indivisible. This chapter avoids these problems by proposing a time-dependent formulation of the DNDP known here as the DNDP-T.

With the extension to the DNDP-T one can not only consider time-dependent travel demands and the gradually upgraded network over the planning period but also more realistic capacity improvement schemes. Furthermore, in this chapter these considerations are extended to include tolling and improvement strategies under two different principles of cost recovery, namely annual cost recovery and overall cost recovery. The annual cost recovery requires that the cost associated with the network design in each year is exactly recovered by the toll revenue generated in the same year, while the overall cost recovery only requires that the total cost of the network design is exactly recovered by the toll revenue generated throughout the design period.

The DNDP-T under cost recovery is formulated as a single-level mathematical program which allows the use of existing non-linear programming techniques for solutions. This study adopts a combination of the Generalized Reduced Gradient (GRG) algorithm and the Branch and Bound (B & B) algorithm to obtain solutions. Three examples are also given to highlight different aspects of the DNDP-T model, namely the importance of introducing discrete variables to NDP over time, the importance of introducing the time dimension into the analysis, and comparison between the two different principles of cost recovery.

6.3 DESCRIPTION OF THE NDP-T

A general transportation network, which is made up of nodes and links is considered. Links represents roadways whereas nodes represent zone centroids or intersections. The planning period is divided up into N equal intervals. The time-space nature of the problem can then be described by N networks graphically. Figure 6.1 shows an example of the NDP-T over three future times A, B, and C; the time scale between them is in the order of years. Over time, the network links are gradually improved as shown by the thickening lines, in response to the changing demand pattern (and land-use pattern if any). The demand patterns may change non-uniformly over time; new towns or suburbs can accommodate a higher future population than already built-up areas, for example.

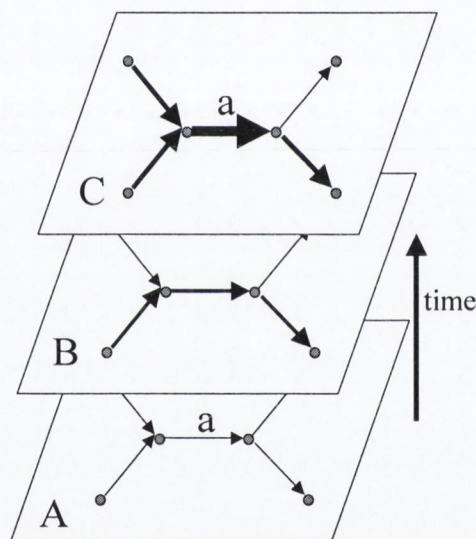


Figure 6.1: The NDP-T

Instead of biasing the network design to be entirely determined by the ultimate demand pattern C far away in the future, the objective is to optimise the network design over the entire planning horizon, so that the network will not produce sub-optimal results in-between now and C, which could be many years apart. For example, one may not want to widen link a in figure 6.1 at time A in anticipation of a large demand at time C far away in the future. This is especially important if the demand projections are subject to large variations. Over-committing resources too early would limit the chances of recourse in the future. On the other hand, the economy of scale, the associated maintenance costs, and the trade-off between inflation, interest rate, and discount rate of benefits must be duly considered. All these considerations can be considered in this NDP-T framework.

The time scale considered in the NDP-T is typically in years, as compared with the second-to-second scale of traffic dynamics, or the day-to-day scale of route choice dynamics. For ease of exposition, let the planning horizon be discretized, as illustrated in figure 6.1. For each discretized time, the traffic assignment pattern is considered as following some sort of user-equilibrium pattern. And for simplicity this study follows the static equilibrium approach. The demands, whether considered as given or elastic, vary with time along the planning horizon. One can conceptualize the NDP-T as a bi-level mathematical program. The upper level problem specifies the objective to be achieved across the entire planning horizon by determining, for example, the network capacity enhancements and tolls over time subject to the budget, economic, and demand growth constraints. The lower level problem contains copies of the user-equilibrium conditions, one for each discretized time along the planning horizon, with respect to the time-dependent design variables specified in the upper level problem.

The idea of DNDP-T basically is the same as NDP-T except the DNDP-T considers lane additions to links and discrete capacity values. This leads to an optimisation problem with discrete variables.

6.4 FORMULATION OF THE DNDP-T MODEL

6.4.1 Framework and Assumptions

A general transportation network with multiple links, routes and origin-destination (OD) flows over a planning period $[0, T]$ is considered. The planning period is divided up into L equal intervals. The length of each interval considered in this study is one year, but the formulation indeed does not pose any restriction on the length of each interval and the number of intervals used, provided that their product is T . In this way the proposed model has the flexibility to cater for situations where the construction period is longer than that assumed in this chapter. This may be particularly relevant when, for example, a planner wishes to consider the addition of a new link or a large number of new lanes to an existing link in the network. This study adopts a similar set of assumptions as those given in Lo and Szeto (2004). These assumptions are as follows:

1. The potential demand growth over time is predicted perfectly.

2. Traffic assignment follows Wardrop's first principle (Wardrop, 1952).
3. This study only considers the addition of indivisible facilities to the network topology and takes the topology as given.
4. The functional form of the travel demand function remains unchanged over time although the parameters of this function do change over time.
5. The link costs and travel demand functions are separable.
6. The interest rates and inflation rates remain constant over time (for simplicity).
7. Only improvements costs are captured in the analysis, again for simplicity.

Lo and Szeto (2004) argue that the second assumption is reasonable since the length of each planning period is reasonably long so that the equilibrium condition is achieved within each period, and that each user has perfect information about the network state as is typically assumed in deterministic traffic assignment. Furthermore, in planning, the time duration between each design period is usually long enough for equilibrium to occur. The effect of the construction transition period, which is relatively short, is ignored. Most of the above assumptions are adopted to simplify the analysis and can be relaxed for future studies.

6.4.2 Time-Dependent Traffic Assignment Constraints

The planning horizon of the DNDP-T is composed of a number of different time 'slices'. For each time slice, the route choice behaviour of the network users follows Wardrop's Principle (Wardrop, 1952). Therefore the Wardrop's conditions hold at each discrete time (time slice) in the planning period. The Wardrop's condition for each time interval can be expressed as the following non-linear complementarity conditions:

$$h_{jr}^{rs} [\phi_{jr}^{rs} - \pi_{\tau}^{rs}] = 0, \forall rs, j, \tau, \quad (6.1)$$

$$\phi_{jr}^{rs} - \pi_{\tau}^{rs} \geq 0, \forall rs, j, \tau, \quad (6.2)$$

where h_{jr}^{rs} and ϕ_{jr}^{rs} are, respectively, the hourly flow and route travel cost for route j between Origin-Destination (OD) pair rs in year τ ; π_{τ}^{rs} is the minimum travel cost between OD pair rs in year τ . According to (6.1), if route j carries flow between OD pair rs in year τ ($h_{jr}^{rs} > 0$), the corresponding route travel cost between OD pair rs in year

τ , $\phi_{j\tau}^{rs}$, must be equal to the minimum travel cost π_{τ}^{rs} between OD pair rs in year τ , as the term in the square bracket in (6.1) must equal zero. If route j carries no flow (i.e., $f_{j\tau}^{rs} = 0$), the corresponding route travel cost between OD pair rs , $\phi_{j\tau}^{rs}$, must be greater than or equal to the minimum travel cost π_{τ}^{rs} based on (6.2).

Each time slice includes flow conservation and non-negativity conditions, expressed as:

$$\sum_j h_{j\tau}^{rs} = N_{\tau}^{rs}, \forall rs, \tau, \quad (6.3)$$

$$h_{j\tau}^{rs} \geq 0, \forall rs, j, \tau, \quad (6.4)$$

where N_{τ}^{rs} is the travel demand of OD pair rs and is defined by an elastic travel demand function in (6.9) below. Equation (6.3) states that the sum of the route flows between an OD pair must be equal to the demand of that OD pair. Inequality (6.4) ensures that route flow must be non-negative.

6.4.3 The Definitional Constraints

The following definitional constraints define the relationships among the hourly path flow $h_{j\tau}^{rs}$, link tolls $\rho_{a\tau}$, the link capacity enhancements $\gamma_{a\tau}$, the route travel costs $\phi_{j\tau}^{rs}$, the hourly link flow $v_{a\tau}$, and the link travel time $t_{a\tau}$. The relationship between path flows and link flows is defined by:

$$v_{a\tau} = \sum_{rs} \sum_j h_{j\tau}^{rs} \delta_a^j, \forall a, \tau, \quad (6.5)$$

where δ_a^j is a link-route incidence indicator, which equals one if link a is on route j , and zero otherwise. Equation (6.5) states that the flow on link a is equal to the sum of the corresponding route flows on that link. The link travel time based on the link flow in (6.5) is given by the following Bureau of Public Roads (BPR) link performance function:

$$t_{a\tau}(v_{a\tau}) = t_a^0 \left[1 + 0.15 \left(\frac{v_{a\tau}}{c_a^0 + \sum_{i=1}^{\tau} y_{ai}} \right)^4 \right], \forall a, \tau, \quad (6.6)$$

where t_a^0 and c_a^0 denote the free-flow travel time and initial capacity of link a , respectively. The summation term in (6.6) represents the total capacity enhancements of link a up to the period τ and when added to the initial capacity (as in the denominator of the round brackets of (6.6)) denotes the link capacity in period τ after implementing the capacity enhancements and inclusive of period τ .

The route travel cost is defined as the sum of the link travel times converted to a monetary value and link tolls charged on those links using that route. Mathematically it is expressed as:

$$\phi_{j\tau}^{rs} = \sum_a (\psi t_{a\tau}(v_{a\tau}) + \rho_{a\tau}) \delta_a^j, \forall rs, j, \tau, \quad (6.7)$$

where ψ is the value of time and $\rho_{a\tau}$ is the toll charged on link a in year τ . Equation (6.7) computes the route travel cost based on the corresponding link travel costs.

6.4.4 Travel Demand Functions

The travel demand N_τ^{rs} of each OD pair rs in year τ is modelled as an elastic function of the potential demand $\widetilde{N}_\tau^{rs}$ and its lowest travel cost π_τ^{rs} of that year:

$$N_\tau^{rs} = D_\tau^{rs}(\pi_\tau^{rs}, \widetilde{N}_\tau^{rs}), \forall rs, \tau, \quad (6.8)$$

where $D_\tau^{rs}(\cdot)$ is the continuous travel demand function for OD pair rs in year τ . The elastic travel demand function is generally decreasing, implying that higher travel costs lead to lower travel demands. In this particular study the following non-linear travel demand function is adopted:

$$N_\tau^{rs} = \widetilde{N}_\tau^{rs} e^{-\gamma \pi_\tau^{rs}}, \quad (6.9)$$

where γ^{rs} is the parameter of the travel demand function of OD pair rs .

The potential demand of each OD pair in each year describes the relationships between the potential demands in successive periods. It represents the potential travel growth due to population growth, expressed as:

$$\widetilde{N}_{\tau}^{rs} = G^{rs}(\widetilde{N}_{\tau-1}^{rs}), \tau > 1, \forall rs, \quad (6.10)$$

where $G^{rs}(\cdot)$ is the potential demand function of OD pair rs . The potential demand in year τ is modelled here to depend only on the potential demand in the previous year and not on the traffic conditions. This assumption is adopted here merely to simplify the analysis and follows similar assumptions used in Lo and Szeto (2009). The potential demand function can be easily extended to include changes in land-use patterns or other local environment factors. For simplicity, the DNDP-T model in this chapter employs a linear potential demand function, defined as:

$$\widetilde{N}_{\tau}^{rs} = \widetilde{N}_{\tau-1}^{rs} [1 + p^{rs}], \quad (6.11)$$

where p^{rs} is the growth rate of potential demand between OD pair rs .

6.4.5 Design and Toll Constraints

For practical reasons such as, for example, the unavailability of development land in urban areas, there is often an upper bound on the amount of capacity expansion that a highway link can undergo. Furthermore, in practice capacity depends on the number of lanes and link widening is a result of the addition of indivisible lanes to the existing network topology. Therefore this study models the capacity expansions as the multiple of the capacity of one lane. The model is further constrained to return only non-negative values of capacity expansion since capacity reduction is not considered here and in general it is rare in practice. Mathematically these constraints may be expressed as:

$$c_a^0 + \sum_{i=1}^{\tau} y_{ai} \leq u_a, \forall a, \tau, \quad (6.12)$$

$$y_{ar} \geq 0, \forall a, \tau, \quad (6.13)$$

$$y_{ar} = 0, l, 2l, \dots, \forall a, \tau, \quad (6.14)$$

where l is the capacity of a standard lane. Equation (6.12) constrains the design to be less than its maximum allowable capacity by limiting the total capacity to be less than its maximum allowable capacity u_a . Condition (6.13) is the non-negativity constraint of capacity enhancements while equation (6.14) constrains capacity enhancements to be the multiple of the capacity of a standard lane. For simplicity, in the numerical studies in this chapter, l is taken as one without loss of generality. Equation (6.14) constrains the capacity improvements to be discrete in nature and only allows for capacity increases which are practical to implement in reality, such as the addition of one, two or three new lanes of traffic etc., as opposed to the many small capacity increases to the network very often yielded by the continuous approximation to the discrete model.

For political and other reasons it is not always feasible to set the toll charge too high while in practice only non-negative tolls are possible and charged at certain links. Mathematically these conditions are expressed as:

$$\rho_{ar} \leq \rho_{a,\max}, \quad (6.15)$$

$$\rho_{ar} \geq 0, \forall a, \tau, \quad (6.16)$$

Condition (6.15) requires that the toll charged on link a in year τ be less than the maximum allowable toll $\rho_{a,\max}$ for that link. If the link is not allowed to charge any toll, $\rho_{a,\max}$ is set to zero. Condition (6.16) ensures that all tolls are greater than or equal to zero.

6.4.6 Cost and Revenue Functions

The annual toll revenue R_τ shown in equation (6.17) below is simply equal to the sum of the product of the equilibrium link flows and link tolls multiplied by a parameter n to convert hourly flows to annual flows. The annual improvement cost K_τ is computed by summing the corresponding improvement costs of all links.

$$R_\tau = \sum_a \rho_{ar} (nv_{ar}), \forall \tau, \quad (6.17)$$

$$K_{\tau} = \sum_a m_{a\tau}(y_{a\tau}), \forall \tau, \quad (6.18)$$

where:

$$m_{a\tau}(y_{a\tau}) = b_{a,0} t_a^0 (1+j)^{\tau-1} (y_{a\tau})^{b_{a,1}}, \forall a, \tau, \quad (6.19)$$

$m_{a\tau}(\cdot)$ is the time-dependent improvement cost function of link a in year τ with parameters $b_{a,0}, b_{a,1}$ and inflation factor $(1+j)^{\tau-1}$. Equation (6.19) describes how the improvement cost of a link is proportional to the capacity enhancement $y_{a\tau}$ and its length $b_{a,0} t_a^0$.

6.4.7 Cost Recovery Constraints

In this study the DNDP-T model is also subject to exact cost recovery constraints. For simplicity this study does not consider subsidies or the case of full cost recovery (where there is a profit/surplus at the end of the planning period). Exact cost recovery (in this context) occurs when the revenue obtained through tolling exactly covers the construction cost. Exact cost recovery is further classified here into two types: Overall exact Cost Recovery (OCR) and Annual exact Cost Recovery (ACR), respectively expressed as:

$$\sum_{\tau} \frac{R_{\tau}}{(1+i)^{\tau-1}} - \sum_{\tau} \frac{K_{\tau}}{(1+i)^{\tau-1}} = 0, \quad (6.20)$$

$$R_{\tau} - K_{\tau} = 0, \forall \tau, \quad (6.21)$$

where i is the rate of interest.

The OCR equation (6.20) requires that, in present value terms, the total toll revenue collected over the planning period must exactly equal the total cost of all improvements carried out in the same period. Under the OCR principle the total cost and revenue must be balanced at the end of the planning period (but not necessarily at intervals during the planning period). On the other hand the ACR principle described in equation (6.21) requires that the revenue and improvement costs must be balanced annually. Unlike OCR, there is no possibility of revenue saved from previous years being used under ACR.

6.4.8 The Objective Function

This study adopts the maximisation of Consumer Surplus (CS) as the performance measure. CS measures the difference between what travellers are willing to pay for travel and what they actually pay. CS internalizes the effect of network congestion and the propensity of users to travel. For the same network and demand characteristics, a higher CS implies a better performing system. CS is expressed mathematically as:

$$CS_{\tau}^{rs} = n \left[\int_0^{N_{\tau}^{rs}} D_{\tau}^{rs-1}(v) dv - \pi_{\tau}^{rs} N_{\tau}^{rs} \right], \quad (6.22)$$

where CS_{τ}^{rs} is the CS of OD pair rs in year τ ; $D_{\tau}^{rs}(\cdot)$ is the travel demand function for OD pair rs in year τ . Inside the square brackets of (6.22) the first term represents what users would be willing to pay for travel while the second term represents the total travel cost (including the toll) that the travel demand actually pays. This study maximises the Total Discounted Consumer Surplus (TDCS) that is obtained by summing the consumer surpluses for all OD pairs, in present value terms, over the entire planning period. Mathematically the TDCS is expressed as:

$$TDCS = \sum_{\tau} \sum_{rs} \frac{CS_{\tau}^{rs}}{(1+i)^{\tau-1}}, \quad (6.23)$$

where the reciprocal of the denominator of (6.23) is the discount factor for year τ . The TDCS therefore represents the entire network over the whole planning period.

6.4.9 The DNDP-T Formulation

The time-dependent discrete network design problem (DNDP-T) can then be formulated as:

$$\max_{h,y,p} TDCS = \sum_{\tau} \sum_{rs} \frac{CS_{\tau}^{rs}}{(1+i)^{\tau-1}}$$

subject to: time-dependent traffic assignment constraints (6.1)-(6.4);
 definitional constraints (6.5)-(6.7);

travel demand functions (6.8), (6.10);
design and toll constraints (6.12)-(6.16);
cost and revenue functions (6.17)-(6.19), and
the OCR constraint (6.20) or the ACR constraint (6.21),

where $\mathbf{h}$, $\mathbf{y}$, and $\mathbf{p}$ are the vectors of route flows, capacity improvements and tolls. The existence of solutions depends on whether the non-linear constraints form a feasible solution set. Once the non-linear constraints form a feasible solution set, multiple solutions can exist although multiple solutions have not been encountered during the course of this study. Multiple solutions arise for the reason that the solution set is not convex and many network improvement and tolling plans can achieve the same TDCS in many situations. This highlights the non-uniqueness of the solutions to the DNDP-T and underscores the need to sort out the most appropriate strategy among the multiple solutions for the local situation and context. This brings up the need of defining secondary objectives to facilitate the selection process. However, it must be said that no model is designed to replace human design-making but rather the motivation of the model is to facilitate the decision maker in choosing the best strategy among many alternatives. Indeed multiple solutions to network design problems can be potentially advantageous since they provide the planner with many different alternative strategies to choose from, all of which give a similar TDCS.

6.5 SOLUTION METHOD

The proposed formulation is a single-level maximisation program, which can be solved by many existing optimisation algorithms. In the numerical section which follows, the DNDP-T formulation is solved using a combination of the Generalized Reduced Gradient (GRG) algorithm proposed by Abadie and Carpentier (1969), and the Branch and Bound (B & B) algorithm. The detailed algorithmic steps of the GRG are described in Lo and Szeto (2004) and in appendix A which follows this thesis. The B & B algorithm is invoked to deal with the discrete variables. The B & B algorithm starts by solving the relaxed problem (without the integer constraints) using the GRG algorithm, yielding an initial best bound for the problem including the integer constraints. The algorithm then begins branching and solving sub-problems with additional (or tighter) bounds on the integer variables. A sub-problem whose solution satisfies all of the integer constraints is a candidate for the solution of the overall problem; the candidate with the best objective value so far is saved as the

incumbent. The algorithm uses the best objective value of the remaining nodes to be fathomed to update the best bound. Each time the algorithm finds a new incumbent, it computes the relative difference between its objective value and the current best bound with respect to the current best bound, yielding an upper bound on the improvement in the objective value that might be obtained by continuing the solution process. If this relative difference is less than or equal to the tolerance specified, the algorithm stops.

It is noted here that the model contains non-convex constraints due to the time-dependent user-equilibrium conditions. Therefore, general-purpose non-linear optimisation codes will stop at local minima. This also means that the best bound used by the B & B algorithm will be based on local optima found by the GRG algorithm and this may not be the global optimum. Because of this, the B & B algorithm is not guaranteed to find the true integer optimum for non-linear problems, although it will often succeed in finding a 'good' integer solution. Therefore it is important to investigate how to take advantage of the special structure of this model to establish stable, efficient solution algorithms. One possibility is to develop effective sensitivity-based or penalty-based algorithms to deal with these constraints for larger-scale implementations. However, it is beyond the scope of this chapter/thesis and is left for future studies.

6.6 NUMERICAL STUDIES INVOLVING THE DNDP-T

In the following three examples, three aspects are illustrated: the importance of introducing discrete capacity constraints to NDP, the importance of introducing the time dimension to NDP, and cost recovery over time.

6.6.1 The Comparison of the DNDP-T and CNDP-T under the ACR Principle

This example aims at illustrating the potential problem of continuous approximation to NDP and bringing out the importance of modelling capacity as a discrete variable. The simple test network used to generate results is shown in figure 6.2 below and consists of one OD pair. The OD pair (1,2) connects node one (the origin) to node two (the destination) via only one possible route – route one. While this may be considered as a degenerate case since there is no route choice between origin and destination it is nonetheless sufficient for the purposes of this example.

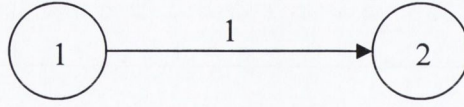


Figure 6.2: Test network 1

The following parameters which are chosen for illustrative purposes also apply:

Free-flow travel times: $t_1^0 = 10$ minutes

Initial link capacity: $c_1^0 = 10$ vehicles/hour

Demand parameters: $\tilde{N}_1^{12} = 150$ vph; $p^{12} = 0.04$; $\gamma^{12} = 0.15$

Maximum allowable link capacity: $u_1 = 200$ vph

Maximum allowable toll: $\rho_{1\max} = \text{€}10$

Value of time: $\psi = \text{€}60/\text{hour}$

Improvement cost parameters: $b_{1,0} = \text{€}43000/\text{veh}$; $b_{1,1} = \text{€}1/\text{veh}$

Interest and inflation rates: $i = 0.03$; $j = 0.01$

Factor to convert hourly flows to annual flows: $n = 8760$

Design period: $[0, T] = [0, 3]$ years

The above parameters can be broadly classified into three different types: network parameters; demand parameters; and financial parameters. Estimation of the network parameters such as the free-flow travel time and link capacity are relatively straightforward since they are related to the geometry of the network and the speed limit among other factors. For the networks used in these examples no data is available as the network parameters are chosen purely for illustrative purposes. The potential demand parameter can in general be estimated from traffic counts although again the values chosen in this study are for illustrative purposes. The demand growth rate parameter can be assigned a value based on factors such as population growth and other economic forecasts that would indicate future growth in car ownership. The demand growth may also increase if there are shifts in the local land-use patterns that may attract more trips to a particular destination. However, for simplicity, changes in land-use over time are not considered in this study. The value of time parameter is generally estimated from surveys of sections of the population in which respondents are generally asked to put a monetary value on how much

an hour of their time is worth. In general this value is related to the income of the person being interviewed. Therefore the final figure calculated for the ‘value of time’ is generally an average value since it will of course be higher for some sections of the population and lower for others. The rates of interest and inflation are index-linked and therefore can be estimated from economic forecasts. Other parameters such as the improvement cost parameter can be estimated from previous construction projects that have taken place in the particular country of interest.

This example considers two approaches: 1) the discrete approach (i.e. DNDP-T) and 2) the continuous approximation (i.e. CNDP-T). Both approaches adopt the ACR principle. The formulation of the DNDP-T under this principle is described in the formulation section of this chapter whereas the one for the CNDP-T can be obtained by removing the discrete capacity constraint (6.14). These two formulations can be solved by the solution method described earlier. Table 6.1 shows the optimal link improvements and tolls for the two approaches.

Table 6.1: Link improvements and tolls for DNDP-T and CNDP-T

Year	Link Improvement		Link Toll Charge	
	DNDP-T	CNDP-T	DNDP-T	CNDP-T
1	1	0.77	3.98	2.95
2	0	0.30	0.00	1.04
3	0	0.00	0.00	0.00

In terms of link improvements, the two approaches adopt quite different strategies. In particular, the project scale and the timing of improvements in the two approaches are quite different. The discrete approach only introduces capacity enhancement in the first year, whereas the continuous approach introduces improvements in both the first and second years. More importantly, the improvement in the second year under the continuous approach is quite small. This illustrates the importance of considering the discrete form of the NDP over time since the improvements to the network are in the form of indivisible facilities (or lanes in this case) unlike *in the continuous form of the problem where very small improvements can be obtained. These small improvements are often impractical to implement in reality.*

In terms of toll charges, the two approaches also yield different strategies. Tolls are collected only in the first year in the DNDP-T whereas tolls are collected in the first two years in CNDP-T. This is due to the ACR condition. A toll is charged in a year whenever

there is an improvement in that year so as to generate sufficient toll revenue for the improvement cost.

Table 6.2 shows the increase in hourly travel demand and TDCS in both approaches per year in the planning horizon. The discrete approach induces lower demand in the first year than the continuous approach as the toll is higher but the improvements in both approaches are similar according to table 6.1. In the second year, there is no link toll (and hence no improvement) in the discrete approach but there is in the continuous approach. Hence the demand is higher in the second year in the discrete approach. In the third year, both approaches charge no toll. Nevertheless, the total capacity enhancement within the three years in the continuous approach is slightly higher and therefore the demand is slightly higher in the third year.

Table 6.2: Hourly demand and TDCS for DNDP-T and CNDP-T

Year	Hourly Travel Demand		Annual Discounted Consumer Surplus		
	DNDP-T	CNDP-T	DNDP-T	CNDP-T	Difference
1	12.56	13.12	7.34E+05	7.66E+05	-4.18%
2	15.23	14.73	8.63E+05	8.35E+05	3.35%
3	15.37	15.45	8.46E+05	8.50E+05	-0.47%
			2.44E+06	2.45E+06	-0.41%

In terms of TDCS (table 6.2), the DNDP-T produces approximately 4.2% lower in annual TDCS than the CNDP-T in the first year. This is due to the higher toll charge (table 6.1) and corresponding lower travel demand (table 6.2) in the first year in the DNDP-T. Thereafter the situation is reversed quickly in the second year and the DNDP-T shows a much higher TDCS. This again is due to the toll charge and the demand in the second year. No toll is charged in the DNDP-T in the second year whereas a toll of €1.04 is collected in the CNDP-T. In the third year, the two consumer surpluses are very similar in both approaches, as the demands and tolls are very similar. The small difference is due to the slight difference in the overall improvement. The CNDP-T results in a slightly larger overall improvement and therefore a slightly larger TDCS.

To sum up, the timing and scale of project improvements of two design approaches can be quite different. More importantly, the DNDP-T is usually superior to the CNDP-T as the CNDP-T can introduce unrealistic small improvements.

6.6.2 The Comparison of the DNDP-T and DNDP under the OCR Principle

In this study the importance of explicitly considering the time dimension in the DNDP is demonstrated. This time, however, the study is performed under the principle of exact Overall Cost Recovery (OCR). For these purposes, this study considers two approaches under OCR, one with time consideration (DNDP-T) and one without (DNDP). Again, the first approach adopts the DNDP-T model whereas the second approach (i.e. DNDP) adopts the degenerated DNDP-T model – the DNDP-T model with only one time period. In the latter, the potential demands of the final year of the planning periods are considered. The second approach is then to maximise the TDCS once based on the potential demands in the final year under the OCR principle. The assumption of considering the final year potential demands is reasonable since presumably a planner would design the network with the future demands in mind and moreover it allows for a fair comparison between the two approaches. As you may notice, the main difference between the two approaches is that the DNDP-T under OCR can not alone capture the final year potential demands but also the gradually changing travel demands induced. By considering these two approaches, this study allows for a fair comparison between the first approach (the DNDP-T under OCR) and the second approach (the DNDP under OCR) while still considering the fundamental advantages of the DNDP-T (its ability to capture the gradually improved network and changing travel demands). The test network used in this comparison is shown in figure 6.3 below and consists of two OD pairs. The OD pair (1,2) connects node one (the origin) to node two (the destination) via two possible routes – route one which uses link one only and route two which uses links two and three. The other OD pair (3,2) connects the origin node three to the destination node two via one route only – route three which uses link three.

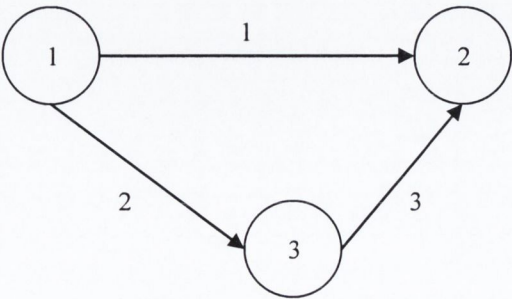


Figure 6.3: Test network 2

The following network parameters chosen for illustrative purposes also apply:

Free-flow travel times: $t_1^0 = 20$; $t_2^0 = 10$; $t_3^0 = 5$ minutes

Initial link capacities: $c_1^0 = 10$; $c_2^0 = 10$; $c_3^0 = 15$ vehicles/hour

Demand parameters: $\widetilde{N}_1^{12} = 40$ vph, $\widetilde{N}_1^{32} = 20$ vph; $p^{12} = 0.02$, $p^{32} = 0.04$; $\gamma^{12} = \gamma^{32} = 0.015$

Maximum allowable link capacities: $u_1 = u_2 = 20$ vph, $u_3 = 35$ vph

Maximum allowable toll: $\rho_{1,\max} = \rho_{2,\max} = \rho_{3,\max} = \text{€}10$

Value of time: $\psi = \text{€}60/\text{hour}$

Improvement cost parameters: $b_{1,0} = b_{2,0} = b_{3,0} = \text{€}43000/\text{veh}$; $b_{1,1} = b_{2,1} = b_{3,1} = \text{€}1/\text{veh}$

Interest and inflation rates: $i = 0.03$; $j = 0.01$

Factor to convert hourly flows to annual flows: $n = 8760$

Design period: $[0, T] = [0, 3]$ years

The DNDP-T and DNDP under the OCR principle are solved as described above to determine link improvements, link tolls, the travel demands induced, and the annual TDCS. Both the DNDP and DNDP-T are solved using the solution method described before. The comparison results between the two approaches are shown in tables 6.3-6.6.

Table 6.3: Link improvements for the DNDP-T and DNDP under OCR

Link	Link Improvements			
	DNDP-T			DNDP
	Year 1	Year 2	Year 3	Over Planning Horizon
1	0	0	0	0
2	5	0	0	0
3	18	0	0	10

Table 6.4: Link tolls for the DNDP-T and DNDP under OCR

Link	Link Tolls			
	DNDP-T			DNDP
	Year 1	Year 2	Year 3	Over Planning Horizon
1	0.00	0.00	0.00	0.50
2	3.92	3.89	4.04	4.2
3	0.01	0.10	0.03	0.38

Table 6.5: Link flows for the DNDP-T and DNDP under OCR

Hourly Link Flows				
Link	DNDP-T			DNDP
	Year 1	Year 2	Year 3	Over Planning Horizon
1	6.04	6.28	6.50	9.00
2	13.73	13.84	13.98	10.41
3	30.12	30.84	31.61	27.89

Table 6.6: Hourly travel demands and TDCS for the DNDP-T and DNDP under OCR

Hourly Travel Demands					Overall TDCS		
OD	DNDP-T			DNDP	DNDP-T	DNDP	Difference
	Year 1	Year 2	Year 3	Over Planning Horizon			
1-2	19.74	20.07	20.38	19.07	6.33E+07	6.06E+07	4.5%
3-2	16.53	17.11	17.81	17.44			

In terms of link improvements, both approaches concentrate on improving link 3 since this link carries the most flow as it is common to both OD pairs (see table 6.5 for the link flows). The DNDP-T approach also improves link 2 since a higher link flow is induced under this scheme. *Overall the DNDP-T improves the network by significantly more and improves links that were not improved at all under the DNDP approach.* This occurs since the DNDP-T is more sensitive to changes in travel demand patterns from year to year. However, although the DNDP chooses to improve link 3 in a somewhat similar way it fails to improve link 2 at all as it bases the design on the travel demands of the final year. The travel demands are indeed subject to change throughout the planning period. Only the DNDP-T has the ability to adjust the improvements accordingly and improves the network at the right locations.

As mentioned before, both approaches improve the network somewhat differently but it is worth noting that under the DNDP-T approach all improvements are carried out in the first year. This seems to be the best strategy since improvements carried out in the first year can benefit that year as well as subsequent years and the interest rate is relatively small. This result is possible due to the principle of overall cost recovery which allows the network to be built first and paid for later. A similar result would be possible if one considered a budget available to use as a lump sum at the beginning of the planning horizon.

In terms of tolling strategies (table 6.4), both models concentrate on tolling link 2. The DNDP-T model charges a lower toll on link 2 which is similar from year to year but can do so as the demand on link 2 is high enough and hence the toll revenue generated from link 2

over time is nearly sufficient to recover all the costs of construction of link 2 and link 3 in the first year. The DNDP approach on the other hand has to charge a small toll on links 1 and 3 in addition to collecting a toll on link 2 to recover the cost construction cost on link 3.

Under the principle of exact OCR both approaches must charge sufficient tolls to exactly cover the costs of improvements. Under the DNDP-T approach one can clearly see when these improvements should be paid for over time and how much should be paid for each year (i.e. the annual toll revenue generated in the network) whereas this is not clear from the DNDP results. Again this illustrates some of the advantages of explicitly considering the time dimension in the DNDP.

Table 6.6 shows the travel demands induced in the network. The demands under the DNDP-T approach increase steadily from year to year, and in comparison, the DNDP-T approach induces a slightly higher overall travel demand due to a larger overall network expansion. In addition, in terms of demand split, the DNDP-T approach attracts more users away from link 1 to link 2 since it expands the capacity of link 2 in the first year (table 6.5).

Table 6.6 also compares the overall TDCS between both the two approaches. As table 6.6 shows, the TDCS resulting from the DNDP-T model is approximately 4.5% higher than the DNDP model. This 4.5% represents a significant increase when one considers that the planning horizon used here in this study is relatively short (3 years). In practice you would expect that for longer planning horizons and for more complex networks, the difference in TDCS would be even greater in favour of the DNDP-T approach.

The last paragraph demonstrated the improved network performance (in terms of TDCS) when the DNDP-T model is adopted. However, it is also worthwhile to examine the effect of using the DNDP-T model from the perspective of individual network users and not just society as a whole. Users on each OD pair are concerned with their travel costs π_r^{rs} . These travel costs are directly related to CS_r^{rs} . The lower is the travel cost, the higher is the corresponding CS_r^{rs} . This study adopts a similar approach to that taken in Lo and Szeto (2003) to describe the change in travel cost 'before' and 'after' the improvements. The

percentage change (decrease) in OD travel cost of OD pair rs in period τ is defined here as follows:

$$PCC_{\tau}^{rs} = \frac{\pi_{\tau}^{rs,a} - \pi_{\tau}^{rs,b}}{\pi_{\tau}^{rs,b}} \times 100\% . \tag{6.24}$$

The superscripts b,a , respectively, refer to the cases of ‘before’ and ‘after’ the improvement. Table 6.7 compares the OD travel costs between the two approaches. As table 6.7 shows, the DNDP-T model consistently produces a larger percentage decrease in OD travel costs after the improvements have been made, in each year. By comparing the DNDP approach to the final year in the planning horizon of the DNDP-T model, it can be clearly seen that the DNDP-T model shows significant percentage decreases in OD travel costs for each OD pair whereas under the DNDP approach users are in fact worse off after the improvements on OD pair 1-2 since they face a percentage increase in travel costs of approximately 3%. On OD pair 3-2 the DNDP model also shows a significant percentage decrease in OD travel costs of approximately 40% but this is still much smaller than the corresponding decrease of 47.5% resulting from the DNDP-T model. Of course the reasons for the lower travel costs generated by the DNDP-T model are obvious since this is the model which also improves the network capacity by significantly more by the end of the planning horizon. This illustrates how the DNDP-T model not alone brings greater benefit to society but also benefits every OD pair through lower travel costs and therefore brings the greater benefit to individual travellers.

Table 6.7: OD travel costs for DNDP-T and DNDP

DNDP-T				DNDP
OD Travel Costs Before Improvements				
OD	Year 1	Year 2	Year 3	Over Planning Horizon
1-2	21.73	21.98	22.24	21.72
3-2	10.13	10.42	10.73	10.83
OD Travel Costs After Improvements				
1-2	20.45	20.55	20.67	22.38
3-2	5.52	5.66	5.64	6.49
Percentage Change in OD Travel Cost				
1-2	-5.91%	-6.52%	-7.07%	+3.07%
3-2	-45.48%	-45.74%	-47.50%	-40.04%

To sum up, this study illustrates the importance of considering the time dimension in the DNDP in terms of giving better network performance (i.e. TDCS), providing greater

benefits to individual users in terms of lower OD travel costs, and improving networks on the right links and to the right scale.

6.6.3 Comparison of OCR and ACR under the DNDP-T

This example adopts the framework of the DNDP-T to compare the two different principles of cost recovery (OCR and ACR) that were used in the previous two studies. Such a comparison is not possible without explicitly considering time in the DNDP. This study adopts the same network (figure 6.3) and parameters as the previous numerical study. Both models (the DNDP-T under OCR and the DNDP-T under ACR) are solved by the solution method described in section 6.5 of this chapter.

Figures 6.4a and 6.4b show the link capacities over time under the two principles of cost recovery.

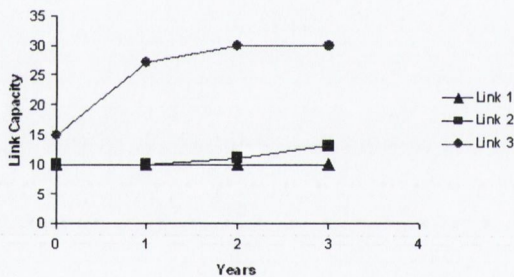


Figure 6.4a: Link capacities under ACR

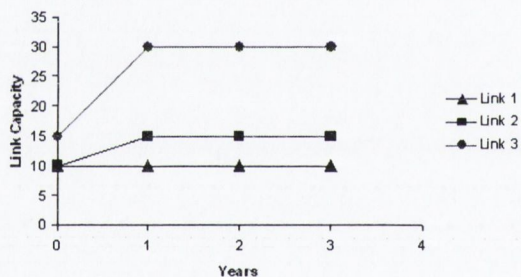


Figure 6.4b: Link capacities under OCR

The first observation is that for the DNDP-T model under the OCR approach all network improvements concentrate in the first year only and no other improvements take place thereafter (see figure 6.4b). However, the costs of expanding the network in year one are recovered over time by the tolls collected gradually over the planning period (table 6.8a). In contrast, as shown in figure 6.4a, the ACR approach determines that the network should be upgraded gradually over the planning period because the improvements within a given year must be funded by the toll revenues collected within that year. Another observation is that under the OCR approach it is possible to construct the network first and pay for it later whereas the ACR approach does not have this flexibility. The third observation is that by the end of the planning horizon both links 2 and 3 are improved under both approaches and link 1 is not improved by either in both approaches. This observation is case specific and related to potential demands. One more observation is that overall the OCR approach

introduces 10% more capacity improvements, due to the flexibility of build-first-pay-later. The toll revenue can be borrowed first to recover the construction cost at the first year.

Table 6.8 shows the tolling strategies of both approaches. Link 2 is tolled almost exclusively by both approaches but the manner in which the tolls are charged over time varies considerably. The OCR approach charges a similar toll on link 2 over the three years whereas the ACR approach charges a higher toll in year one and lowers it considerably in the subsequent years to reflect the decrease in capacity improvements. The OCR approach is more flexible and can build the network more quickly thereby allowing the benefits of the improvements to be felt in all years of the planning horizon. In contrast, the ACR approach charges a very high toll in the network in year one (€5.56 on link 2 and €1.53 on link 1) in order to begin expanding the network sufficiently that can benefit the users in this and subsequent years. The improvements in the second and third years are much smaller compared with the improvement in the first year and hence lower tolls are charged.

Table 6.8: Tolling strategies under OCR and ACR

a) OCR				b) ACR			
Link	Year 1	Year 2	Year 3	Link	Year 1	Year 2	Year 3
1	0.00	0.00	0.00	1	1.53	0.00	0.00
2	3.59	3.54	3.48	2	5.56	2.25	1.55
3	0.00	0.00	0.00	3	0.00	0.00	0.00

Overall both approaches produce similar values of TDCS with the OCR approach gaining approximately 1.5% more. However, it is worth noting that the OCR approach shows a 4.3% higher TDCS at the end of year one which underlines the benefits of constructing the network early in the planning horizon. The ACR approach ‘catches up’ by the end of the planning horizon shown here but the benefits of the improvement scheme under this approach cannot be fully felt until the end of the planning horizon. In larger networks over longer planning horizons this delay in experiencing the benefits of the improvements might be too long. Hence in practice the DNDP-T under the OCR principle might be more desirable assuming that the budget for the improvements are available as a ‘lump sum’ at the beginning of the planning horizon (for example, by borrowing from the bank), and can be repaid later from tolls charged.

6.6.4 Discussion

So far, this chapter extended the traditional/conventional DNDP by introducing the time dimension explicitly within the formulation. The resulting DNDP-T framework permits: modelling of time-dependent travel demands; capturing the gradually upgraded network over the planning period; the determination of optimal timing of project construction, project scale and phasing; determining optimal time-dependent tolls as well as the study of cost recovery over time. The DNDP-T was formulated as a single-level optimisation program, solved through the solution method outlined in section 6.5, and through numerical examples this study demonstrated the importance of introducing discrete capacity variables to NDP, and the superiority of the DNDP-T over the traditional DNDP. The framework developed herein was also used to provide some further insight on the effect of using different cost recovery principles.

Having demonstrated the advantages of the proposed DNDP-T model, the next part of this chapter formulates the proposed reliable DNDP-T (RDNDP-T) model.

6.7 THE RDNDP-T MODEL

In previous approaches to modelling the NDP under uncertainty and the reliable NDP (RNDP), the network improvements are considered to be continuous and furthermore to be one-off events. These approaches do not consider gradual network improvements or changing travel demands over a planning period. These problems are avoided here by proposing a time-dependent formulation of the DNDP. However, unlike the previous studies of both time-dependent NDP and the RNDP, the game theoretic approach to risk-averse traffic assignment is adopted to model the lower level problem. The resulting model allows network uncertainty to be explicitly in a time-dependent design framework with the capacity improvements modelled as discrete variables. This new RNDP model is known here as the Time Dependent Reliable Discrete Network Design Problem (RDNDP-T).

With the extension to the RDNDP-T one can not only consider time-dependent travel demands, changing scenario probabilities and the gradually upgraded network over the planning period but also more realistic capacity improvement schemes. The RDNDP-T is formulated as a single-level mathematical program which allows the use of existing non-

linear programming techniques for solutions. An example is given to: illustrate reliable network improvement strategies under a limited budget; demonstrate how only the most vulnerable links receive improvement; and show how the method can design the critical links over time under changing travel demands. Through the numerical example, important aspects of the RDNDP-T are highlighted such as the importance of introducing discrete variables to NDP over time and the importance of introducing the time dimension into the analysis.

6.8 FORMULATION OF THE RDNDP-T

A similar framework and set of assumptions as those assumed in section 6.4.1 are adopted here although in this case the link flow is assigned to the paths in the network under risk-averse traffic assignment. For ease of illustration, this chapter formulates the proposed RDNDP-T under the OD-specific demon formulation of risk-averse traffic assignment with fixed travel demands. However, the problem may also be formulated under the network-specific demon formulation of risk-averse assignment without conceptual difficulty.

6.8.1 Time-Dependent Risk-Averse Traffic Assignment Constraints

Similar to the conventional DNDP-T model proposed earlier in this chapter, the planning horizon of the RDNDP-T is composed of a number of different time ‘slices’. There are N_r^{rs} homogenous users between each OD pair rs in year τ , each of whom is assumed to seek the minimum cost route. There is also a demon intent on causing maximum delays (and hence increased costs) to the users by damaging one and only one link in each OD pair in the network. This assumption can be easily relaxed (as shown in chapters 3, 4 and 5) but this study considers one demon just for the ease of illustration. Moreover, it is assumed that no user knows which link a demon will choose to damage, and no demon knows in advance which link a user may decide to use. Furthermore, each link is assumed to have only two costs depending on whether it is damaged or not. With this setting, the risk-averse user equilibrium traffic assignment problem with OD-specific demons and with fixed demands can be formulated by applying the results in Nash (1951), Bell and Cassir (2002).

The proposed problem comprises two sub-problems: The user problem and the demon problem. The user problem can be viewed as a non-cooperative game, in which each user tries to select the route with minimum expected trip cost. This user problem can be approximated to deterministic traffic assignment when the number of homogenous players is large (Bell and Cassir, 2002). The first-order condition of deterministic traffic assignment for each time interval can be expressed as the following non-linear complementarity conditions:

$$h_{j\tau}^{rs} \left[\sum_k g_{jk\tau}^{rs}(\mathbf{h}) q_{k\tau} - \min_{d_\tau} \left(\sum_k g_{dk\tau}^{rs}(\mathbf{h}) q_{k\tau} \right) \right] = 0, \quad (6.25)$$

$$\sum_k g_{jk\tau}^{rs}(\mathbf{h}) q_{k\tau} - \min_{d_\tau} \left(\sum_k g_{dk\tau}^{rs}(\mathbf{h}) q_{k\tau} \right) \geq 0, \quad (6.26)$$

$$h_{j\tau}^{rs} \geq 0, \quad (6.27)$$

where $h_{j\tau}^{rs}$ is the flow on route j between OD pair rs in year τ ; $\mathbf{h}$ is the route flow vector; $g_{jk\tau}^{rs}(\mathbf{h})$ represents the cost of route j between OD pair rs in scenario k in year τ based on the flow vector $\mathbf{h}$; d_τ is the minimum expected travel cost route between OD pair rs in year τ ; $q_{k\tau}$ is the probability of scenario k occurring in year τ .

According to (6.25), if route j connecting OD pair rs in year τ carries flow ($h_{j\tau}^{rs} > 0$), the corresponding expected route travel cost between OD pair rs in year τ , $\sum_k g_{jk\tau}^{rs}(\mathbf{h}) q_{k\tau}$, must be equal to the minimum expected travel cost $\min_{d_\tau} \left(\sum_k g_{dk\tau}^{rs}(\mathbf{h}) q_{k\tau} \right)$ between OD pair rs in year τ , as the term in the square bracket in (6.25) must equal zero. If route j connecting OD pair rs in year τ carries no flow ($h_{j\tau}^{rs} = 0$), the corresponding expected route travel cost between OD pair rs in year τ must be greater than or equal to the minimum expected travel cost between OD pair rs in year τ based on (6.26). Condition (6.27) is the non-negativity condition of route flows.

The route flow in year τ , $h_{j\tau}^{rs}$, in (6.25)-(6.27) must satisfy route flow conservation, expressed as:

$$N_{\tau}^{rs} = \sum_j \delta_j^{rs} h_{j\tau}, \quad (6.28)$$

where N_{τ}^{rs} denotes the total number of network users between OD pair rs in year τ ; δ_j^{rs} is the route-OD incidence indicator. $\delta_j^{rs} = 1$ if j connects OD pair rs , and $\delta_j^{rs} = 0$ otherwise.

The demon problem describes the objective of the demon. The demon seeks a mixed strategy to maximise its own individual expected pay-offs or the total expected network cost to the users. The problem can be written for each time interval as the following non-linear complementarity conditions:

$$q_{k\tau} \left(\max_{w_{\tau}} \sum_j g_{jw\tau}^{rs}(\mathbf{h}) h_{j\tau}^{rs} - \sum_j g_{jk\tau}^{rs}(\mathbf{h}) h_{j\tau}^{rs} \right) = 0, \quad (6.29)$$

$$\max_{w_{\tau}} \sum_j g_{jw\tau}^{rs}(\mathbf{h}) h_{j\tau}^{rs} - \sum_j g_{jk\tau}^{rs}(\mathbf{h}) h_{j\tau}^{rs} \geq 0, \quad (6.30)$$

$$q_{k\tau} \geq 0, \quad (6.31)$$

where the probability $q_{k\tau}$ must satisfy:

$$\sum_k q_{k\tau} - 1 = 0. \quad (6.32)$$

and be non-negative as shown in (6.31) by definition, and $g_{jk\tau}^{rs}(\mathbf{h})$ follows the definition in (6.33).

Conditions (6.29)-(6.30) represent the necessary and sufficient conditions of the Nash equilibrium of a non-cooperative, mixed-strategy game for the demon (see Nash 1951). According to (6.29), if scenario k in year τ is selected ($q_{k\tau} > 0$), the maximum total expected network cost in year τ , $\max_{w_{\tau}} \sum_j g_{jw\tau}^{rs}(\mathbf{h}) h_{j\tau}^{rs}$, must be equal to the total expected network cost in scenario k in year τ , $\sum_j g_{jk\tau}^{rs}(\mathbf{h}) h_{j\tau}^{rs}$, through the condition

$$\left(\max_{w_{\tau}} \sum_j g_{jw\tau}^{rs}(\mathbf{h}) h_{j\tau}^{rs} - \sum_j g_{jk\tau}^{rs}(\mathbf{h}) h_{j\tau}^{rs} \right) = 0. \text{ If scenario } k \text{ in year } \tau, \text{ is not selected}$$

($q_{k\tau} = 0$), the maximum total expected network cost in year τ must be greater than or equal to the total expected network cost in scenario k in year τ according to (6.30).

6.8.2 The Definitional Constraints

The cost of route j between OD pair rs in scenario k in year τ , $g_{jk\tau}^{rs}(\mathbf{h})$, appeared in (6.25) and (6.26) and is the total cost on every link that is on this route in year τ :

$$g_{jk\tau}^{rs}(\mathbf{h}) = \sum_a \delta_{ja} t_{ak\tau}(v_{a\tau}), \quad (6.33)$$

where δ_{ja} is 1 if link a is on route j , and 0 otherwise; $t_{ak\tau}(v_{a\tau})$ denotes the flow-dependent cost on link a in scenario k in year τ and is defined by the Bureau of Public Roads (BPR) function in (6.34) below:

$$t_{ak\tau}(v_{a\tau}) = t_a^0 \left(1 + 0.15 \left(\frac{v_{a\tau}}{\overline{c_{ak\tau}}} \right)^4 \right), \quad (6.34)$$

where t_a^0 is the free-flow travel time of link a ; $\overline{c_{ak\tau}}$ is the capacity of link a in scenario k in year τ and $v_{a\tau}$ is the link flow in year τ . Both $\overline{c_{ak\tau}}$ and $v_{a\tau}$ are defined in (6.35) and (6.36) respectively:

$$\overline{c_{ak\tau}} = \lambda_{ak\tau} \left(c_a^0 + \sum_{i=1}^{\tau} y_{ai} \right), \quad (6.35)$$

$$v_{a\tau} = \sum_{rs} \sum_j \delta_{ja} h_{jr}^{rs}. \quad (6.36)$$

In (6.35) c_a^0 is the initial undamaged capacity of link a in year; the summation term in (6.35) represents the total capacity enhancements of link a up to the period τ and when added to the initial undamaged capacity of link a denotes the link capacity in period τ after implementing the capacity enhancements and inclusive of period τ ; while $\lambda_{ak\tau}$ is a link-damage indicator which, in general, equals one if there is no damage to link a in scenario k in year τ and equals 0.5 if there is damage to link a in scenario k in year τ . Equation (6.36) states that the link flow in year τ is simply the sum of the route flows

using that link in year τ .

6.8.3 The Potential Travel Demand Function

For simplicity, the RDNDP-T model proposed in this chapter employs the linear potential demand function defined in equation (6.11).

6.8.4 Design Constraints

The rationale behind the design constraints adopted here is similar to that described in section 6.4.5 for the conventional DNDP-T. Mathematically, these constraints may be expressed as:

$$\overline{c_{ak\tau}} \leq u_a, \forall a, k, \tau, \quad (6.37)$$

$$y_{a\tau} \geq 0, \forall a, \tau, \quad (6.38)$$

$$y_{a\tau} = 0, l, 2l, \dots, \forall a, \tau, \quad (6.39)$$

where l is the capacity of a standard lane. Equation (6.37) constrains the design to be less than its maximum allowable capacity by limiting the total capacity to be less than its maximum allowable capacity u_a . Condition (6.38) is the non-negativity constraint of capacity enhancements while equation (6.39) constrains capacity enhancements to be the multiple of the capacity of a standard lane. For simplicity, l is taken as 1 without loss of generality.

6.8.5 Improvement Cost Functions and Budgetary Constraints

The annual improvement cost, K_τ , and the time-dependent improvement cost function, $m_{a,\tau}(\cdot)$, adopted here are described in equations (6.18) and (6.19) respectively.

The budgetary constraints can be expressed as:

$$K_1 + R_1 = B_1, \quad (6.40)$$

$$K_\tau + R_\tau = R_{\tau-1} + B_\tau, \forall \tau > 1, \quad (6.41)$$

where B_τ is the network improvement budget allocated by the government at the beginning of year τ . The variable R_τ represents in this case the cumulative allocated funds that have not yet been spent in year τ and are available for use at the beginning of year $\tau + 1$. Equation (6.41) states that the sum of the total cost of the improvements for year τ and the funds unspent or saved for the future years after τ is equal to the budget allocated by the government for year τ plus the funds saved from year 1 to $\tau - 1$. Equation (6.40) describes the beginning of the planning horizon, assuming that no funds were saved before the planning horizon. This set of recursive equations permits the modelling of various budget and expenditure scenarios. For example, if the government provides a lump sum only at the beginning of the planning horizon, then $B_1 > 0$ and $B_\tau = 0$, $\tau > 1$. Then (6.40)-(6.41) can be reduced to $\sum_{\tau} K_{\tau} \leq B_1$.

6.8.6 The Objective Function

Similar to the RCNDP models proposed in chapter 5, this study adopts an objective function with a straightforward interpretation. This study considers the minimisation of expected total system travel time (ETSTT) over the entire planning horizon. Mathematically, it can be expressed as:

$$ETSTT = \sum_{\tau} \sum_a \sum_k n(v_{a\tau}(t_{ak\tau}(v_{a\tau}))q_{k\tau}), \quad (6.42)$$

where $v_{a\tau}$ is the link flow in year τ as given in equation (6.36); $t_{ak\tau}(v_{a\tau})$ is the link travel time in scenario k in year τ given in equation (6.34); $q_{k\tau}$ is the probability of scenario k occurring in year τ , n is a parameter to convert hourly flows to annual flows. The product $n(v_{a\tau}(t_{ak\tau}(v_{a\tau}))q_{k\tau})$ is evaluated for every link in each scenario in the network over the entire planning horizon.

6.8.7 The RDNDP-T Formulation

The time-dependent reliable discrete network design problem (RDNDP-T) can then be formulated as a single single-level minimisation program by setting the lower level problem as a set of constraints:

$$\min_{\mathbf{h}, \mathbf{y}, \mathbf{q}} ETSTT = \sum_{\tau} \sum_a \sum_k n \left(v_{a\tau} \left(t_{ak\tau} \left(v_{a\tau} \right) \right) q_{k\tau} \right),$$

subject to: time-dependent risk-averse traffic assignment constraints (6.25)-(6.32),
 definitional constraints (6.33)-(6.36),
 potential travel demand function (6.11),
 design constraints (6.37)-(6.39),
 improvement cost functions (6.18)-(6.19),
 budgetary constraints (6.40)-(6.41),

where $\mathbf{h}$, $\mathbf{y}$, and $\mathbf{q}$ are the vectors of route flows, capacity improvements and scenario probabilities. This formulation is a single-level minimisation program which can be solved by many existing optimisation algorithms. In the numerical section which follows, the RDNDP-T formulation is solved using the solution method described in section 6.5.

6.9 NUMERICAL EXAMPLE INVOLVING THE RDNDP-T

In section 6.6, the results from a number of numerical studies described various aspects of the DNDP-T model and the advantages it holds over the traditional static or DNDP approach and the corresponding continuous time-dependent approach (CNDP-T). This section focuses mainly on the design approach of the RDNDP-T. To avoid repetition, only the results from the RDNDP-T model are considered and no comparison is made to either the DNDP or CNDP-T models. Instead, this numerical study focuses on the selection and design strategies for the vulnerable network links in a time-dependent setting with discrete capacity improvement variables. The advantages of the time-dependent discrete model have already been demonstrated earlier in this chapter. However, it is worth mentioning that in the case of RDNDP-T, the inclusion of the time element means that the changing link selection probabilities can be captured over the planning horizon which can provide further useful information to a planning authority.

In the following example, the following aspects are illustrated: the design approach of the RDNDP-T model under a budget and the selection and design of the vulnerable links over time under different assumptions on the type of demons assumed and the corresponding form of link capacity reduction. Overall (though without direct comparison to static RCNDP models), this example shows the importance of considering the time dimension in

the modelling approach and the importance of using discrete capacity improvement variables over time in reliable network design.

6.9.1 The Design Approach of the RDNDP-T Model

In this example, a scenario is given which aims at illustrating the design approach of the RDNDP-T with a budget under two different assumptions of link capacity reduction. In the first scenario, the demons are assumed to be OD-specific (as formulated in section 6.8) and the link capacity reduction is assumed to be 50%. In the second scenario, the demons are assumed to be network-specific and for comparison purposes non-linear capacity reduction is assumed. The network-specific demon approach to RDNDP-T is similar to the OD-specific approach and may be formulated by replacing the OD-specific risk-averse traffic assignment constraints (6.25)-(6.32) with a time-dependent network-specific demon model of risk-averse traffic assignment. A time-dependent formulation of risk-averse traffic assignment with network-specific demons can be developed by introducing the time dimension into the formulation proposed in chapter 3 of this thesis (similar to the manner in which the time dimension was introduced into the OD-specific demon approach).

It is also assumed that the budget is provided by the government or planning authority as a lump sum at the beginning of the planning horizon. The simple test network used to generate results is shown in figure 6.5 below and consists of two OD pairs sharing a common origin (node one). The OD pair (1,2) connects node one (the origin) to node two (the destination) via one link - link one. The OD pair (1,3) connects node one (the origin) to node three (the destination) via one link – link 2. While this may be considered as a degenerate case since there is no route choice between the origin and destination nodes, it is nonetheless sufficient for the purpose of illustrating the design approach of the RDNDP-T model. This allows the results for the improvement strategies to be more easily explained without being complicated by route choice factors and complicated link selection scenarios. The following parameters which are chosen for illustrative purposes also apply:

Free-flow travel times: $t_1^0 = t_2^0 = 10$ minutes

Initial undamaged link capacity: $c_1^0 = c_2^0 = 10$ vph

Demand parameters: $N_1^{12} = 16$ vph, $N_1^{13} = 18$ vph; $p^{12} = 0.35$, $p^{13} = 0.05$

Maximum allowable link capacity: $u_1 = u_2 = 40$ vph

Factor to convert hourly flows to annual flows: $n=8760$

Design period: $[0, T] = [0, 3]$ years

Budget: $B = \text{€}2.6 \times 10^6$

Damaged link capacities (with improvements):

- $\lambda_{ak} = 1$ if link a is not selected by any demon(s) in scenario k ,
- $\lambda_{ak} = 0.5$ if link a is selected by any number of OD-demon(s) in scenario k ,
- $\lambda_{ak} = 0.5$ if link a is selected by one network-demon in scenario k ,
- $\lambda_{ak} = 0.25$ if link a is selected by two network-demons simultaneously in scenario k ,

where λ_{ak} is a link-damage incidence indicator.

The same financial parameters as those used in previous numerical studies in this chapter are also adopted here.

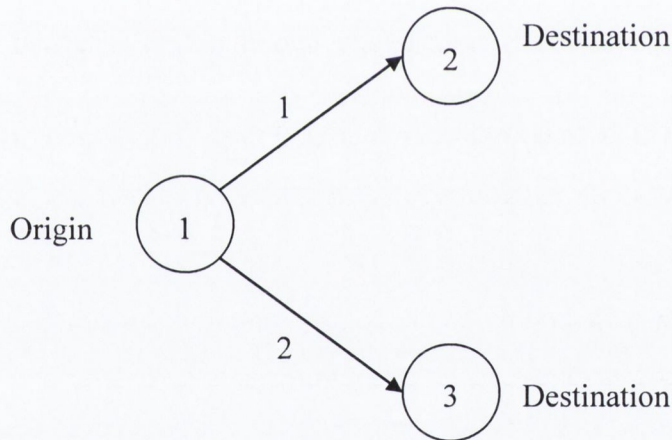


Figure 6.5: Test network 3

In this example, two OD pairs are considered which have the same initial undamaged link capacity and broadly similar base year travel demands. The scenario set up here is that one OD pair is expected to have low potential demand growth while the other has quite high potential demand growth. In practice, this type of scenario may arise due to a shift in land-use patterns at the various destinations over time and while land-use elements are not explicitly considered here, it is nonetheless possible to incorporate them into the framework.

Table 6.9 describes the link selection probabilities of the network-specific and OD-specific demons over the planning horizon when no budget is provided and consequently when no link capacity improvements occur. The link selection probabilities of the network-specific and OD-specific demons can be explained by the link flows over the design period and the assumption on capacity reduction used. The corresponding link flows due to demand growth over the same period are shown in table 6.10. At the beginning of the planning period, OD pair 1-3 (or link 2) has a higher demand than OD pair 1-2 (or link 1). Therefore, both network-specific demons select link 2 to damage in order to maximise their expected pay-off under the assumption of non-linear capacity reduction. In subsequent years of the planning horizon (i.e., years 2 and 3), the network demons combine to select link 1 since the expected pay-off is higher on this link due to the rapid growth in demand on OD pair 1-2. The OD-specific demons do not have the same flexibility and cannot combine to select the same link to damage and as they are forced to damage each OD pair separately. It can be observed that the network demon link selection probabilities may change over the planning period which emphasises the importance of introducing the time-dimension to the problem. The OD-specific demon link selection probabilities do not change over the planning period in this particular example but they may do so where there are more paths connecting a particular OD pair or where a number of links are common to a number of OD pairs. Nevertheless, the network demon approach allows for greater flexibility since the demons can combine to inflict greater damage in terms of capacity reduction. Furthermore, the network demon approach can also capture the OD demon approach, if the link damage indicator λ_{ak} is equal to 0.5 for any number of network demons selecting a link to damage simultaneously.

Table 6.9: Link selection probabilities for the network and OD-Specific demons over time

Link Selection Probabilities of the Demons				
	Network Demon		OD Demon	
	A	B	OD 1-2	OD 1-3
	Year 1			
Link 1	0	0	1	0
Link 2	1	1	0	1
Year 2				
Link 1	1	1	1	0
Link 2	0	0	0	1
Year 3				
Link 1	1	1	1	0
Link 2	0	0	0	1

Table 6.10: Link flows over time

Link Flows over the Planning Horizon (vph)				
OD	Link	Year 1	Year 2	Year 3
1-2	1	16	21.6	29.16
1-3	2	18	18.9	19.845

As detailed in the description of parameters above, a small budget is provided by the government or planning authority as a lump sum at the beginning of the planning horizon. The link capacity improvements for the RDNDP-T model under network-specific and OD-specific demons are shown in table 6.11. Under both the network-specific and OD-specific demon approach, all capacity improvements for the RDNDP-T are completed within the first year. In fact, this ought to be the best strategy as improvements completed within the first year can benefit both the first year and subsequent years. This result, of course, is subject to the assumption that the budget is available as a lump sum at the beginning of the planning horizon.

Table 6.11: Link capacity improvements for the RDNDP-T model

Link Capacity Improvements				
Demon	Link	Year 1	Year 2	Year 3
Network-Specific	1	6	0	0
	2	0	0	0
OD-Specific	1	5	0	0
	2	1	0	0

Under the network-specific approach, all capacity improvements are on link 1 since this eventually attracts the largest hourly demand and in the long term is the link selected for damage by the network-specific demons. The OD-specific approach also concentrates on link 1 due to the high projected demand but also introduces some small improvements to link 2 reflecting the fact that this link is selected for damage under this approach. Nonetheless, it can be observed that under a tight budget, both approaches concentrate improvements on the links which are vulnerable to failure. Unlike the static approach to reliable network design, the time dimension of the RDNDP-T allows for an appreciation of the changing link selection probabilities over time. In fact, this would be even more important in scenarios where the network improvements to certain links would induce more users to these links (i.e., where elastic demand is assumed) and therefore possibly have an effect on the demon link selection probabilities. Furthermore, unlike the continuous versions of the network design problem (i.e., CNCP, CNDP-T, RCNDP), the

improvements introduced here are in terms of additional lanes of traffic. This makes the capacity improvements potentially possible to implement in a reality.

Overall, the effect of using the time-dependent model for reliable network design is clear since this makes it possible to observe which links are the vulnerable ones in each year of the planning horizon. This gives the model the flexibility of adapting to these changes and improving links gradually to the right scale at the right time. In this example, all improvements take place in the first year since the budget is provided as a lump sum at the beginning of the planning horizon. However, if the budget was provided gradually over the planning period then the link improvements in a particular year may vary depending on the demand level in the various network links in that year and the link selection strategy of the demons in that year. Finally, since network improvements may induce new users into the network, it may be worthwhile in the future to incorporate elastic demands into the formulation of RDNDP-T. This extension will, however, call up the need for more appropriate objective functions to adopt in any future RDNDP-T model where elastic demands are incorporated. Minimising ETSTT may be inappropriate when the demand is elastic because this could be achieved through minimizing travel demand and thus result in undesirable solutions involving less investment. Yang and Bell (1998) discuss this issue for network design problems in greater detail. Adopting more appropriate objective functions for this approach, represents a worthy future research direction.

6.10 CONCLUDING REMARKS

The first part of this chapter extended the traditional or conventional DNDP by introducing the time dimension explicitly within the formulation. The resulting DNDP-T framework permits: modelling of time-dependent travel demands, capturing the gradually upgraded network over the planning period, the determination of optimal timing of project construction, project scale and phasing, determining optimal time-dependent tolls as well as the study of cost recovery over time. This chapter formulated the DNDP-T as a single-level optimisation program, solved it through the solution scheme described in section 6.5, and through numerical examples demonstrated the importance of introducing discrete capacity variables to NDP, and the superiority of the DNDP-T over the traditional DNDP. The framework developed herein was also used to provide some further insight on the effect of using different cost recovery principles.

This chapter also proposed the RDNDP-T by modelling the lower level problem under risk-averse traffic assignment. The resulting model allows one to explicitly consider uncertainty in the network design problem and does not rely on data on user behaviour in a damaged network. Similar to the proposed conventional DNDP-T, the time-dependent framework of the RDNDP-T permits modelling of: time-dependent travel demands; capturing the gradually upgraded network over the planning period; the determination of optimal timing of project construction; and project scale and phasing. In addition, the RDNDP-T can determine the vulnerable network links and allow these links to receive priority in the capacity improvement scheme, thereby resulting in a reliable network design. By basing the network design on the expected travel costs of the users when confronted by supply uncertainty, this approach introduces improvements which cater for the worst case scenarios of network reliability. This chapter formulated the RDNDP-T as a single-level optimisation program, solved it through the solution scheme described in section 6.5, and through a numerical example demonstrated how the model determines and designs the critical links under a budget when either OD-specific or network-specific demands are assumed.

CHAPTER 7

Conclusions

7.1 RESEARCH SUMMARY

Interest in transport network reliability has significantly increased in the last number of years and reliability issues are likely to become even more important in the future. This is due to the increased expectations of the users of transport systems and the growing awareness among planners of the important role of transportation networks in acting as crucial lifelines following major incidents which may occur in the network. Furthermore, there is increasing evidence that environmental pollution is increasing the threat of natural disasters (for example, flooding and hurricanes). Thus, interest in the area of network reliability and reliable network design is likely to only intensify in the future. Finally, transport systems are increasingly the target of malevolent acts (or attacks) by those who seek to cause severe hardship to society by disrupting transport systems and in particular targeting links which, when damaged, are most likely to impact on the greatest number of network users. Assessing the network reliability, considering these scenarios where the event probabilities are largely unknown, is likely therefore to attract even greater attention from researchers for many years to come.

On the less severe end of this scale of impacts, there are a number of lower impact but higher frequency events such as road traffic accidents, recurrent congestion, adverse weather and road maintenance, all of which may disrupt or damage the network links. Allowing for the possibility that events such as these may occur anywhere in the network and potentially at the same time or in the same location is addressed in this thesis. Catering for simultaneous disruptive events to occur on the same links in the network within the same day may also result in a reduction in link capacity which is proportional to the number of disruptive events occurring. This thesis addresses these issues through formulating the multiple network demon (game theoretic) approach to network reliability. The effect of assuming multiple network demons and relaxing the capacity assumption to expected network cost and network reliability measurement was demonstrated through

numerical studies. The results show just how important it is to accurately capture the link capacity degradations particularly when the presence of more than one demon is assumed. Relaxing this assumption also allows for more operational states to be modelled. Allowing for users' response to these events is also important including re-routing behaviour and the decision not to travel. This thesis addresses these issues by formulating the risk-averse traffic assignment with elastic demands. Obtaining globally optimal solutions to this formulation is also addressed through the proposed modified α BB solution method with guaranteed ε convergence. In applying the techniques developed in this thesis to assess network reliability, one is aware that the number of variables increases for application to larger real-life networks with many scenarios and a larger number of network demons. This thesis attempted to address this issue by proposing link-based formulations of the proposed multiple network demon formulation of risk-averse traffic assignment. This model can avoid path enumeration which is important particularly when dealing with multiple network demons which already increases computation times due to the large number of scenarios of link damage to be considered. This becomes even more significant when applying the model to larger real-life networks. Under this framework it was proven that there is at least one solution to this link-based problem and this problem is equivalent to the path-based one. This formulation can be reformulated into a VIP and an unconstrained minimisation program for solutions.

Through numerical studies, this thesis has shown that the expected network cost arising from the game between network demons and users is at least the same as the one played against OD demons and in general is higher. Following on from Bell's definition of reliability (see chapter 1) this is considered as a broad measure of the network reliability and since it is generally higher for the network demon formulations, is therefore also considered to capture the worst case scenarios of network reliability from the viewpoint of pessimistic network users. Based on the expected network costs, link selection probabilities and route/link flows, a number of specific measures or indicators of network reliability are defined in this thesis, including encountered and travel time reliability. The former gives the probability of not encountering link degradations on the path with least perceived costs and is defined on the path and OD levels. The latter is the probability that a trip can be made successfully within a given acceptable time interval and is defined on the path, OD and network levels. While the path travel time reliability measures the performance of the path and gives useful information to travellers, the OD and network

travel time reliabilities measure the performances of OD pairs and networks respectively and can be used by network planners for network planning purposes.

Furthermore, to date no study had considered the existence of analogues of Braess' paradox in risk-averse user equilibrium assignment or the resulting implications for travel time and encountered reliability measurements. This thesis demonstrated the existence of stochastic Braess' paradox and reliability paradox (analogues of Braess' paradox for expected network travel cost and reliability) in risk-averse assignment, and provided some further insights into the implications for network design if these paradoxes are ignored.

To cater for future disruptive events, it is important that reliability studies are used to obtain an optimal modification of the network. This is particularly important since the vulnerable network links should receive priority in the improvement scheme. This thesis addresses this issue by proposing a reliable continuous network design problem (RCNDP) where the lower level problem of the RCNDP is formulated under the multiple network demon formulation of risk-averse traffic assignment. However, since network improvements are not one-off events and to capture changing travel demands and gradual network upgrades, the NDP is also formulated with the time dimension explicitly considered in this thesis. Moreover, it is important to consider the addition of link capacity which is practical to implement with long term planning in mind. To this end, this thesis formulates the discrete network design problem over time (DNDP-T) and shows through numerical studies how it is superior to the static DNDP and the traditional time-dependent CNDP. Following this, the thesis proposes the reliable DNDP-T (RDNDP-T) and shows how this can capture changing travel demands, gradual network upgrades and changing link damage probabilities over time.

The rest of this chapter is organised as follows: the specific contributions of each chapter of this thesis are now described in greater detail in section 7.2. Following this, a critical assessment of the work in this thesis is given in section 7.3 and arising from this discussion a number of possible future extensions are proposed in section 7.4.

7.2 RESEARCH FINDINGS AND CONTRIBUTIONS

This section presents the important research findings and contributions of this thesis, which are summarised from chapters 2 to 6 as follows:

Chapter 2 develops an NCP formulation to model the risk-averse traffic assignment with elastic demands, and proposes a finite ε convergence method to solve for global solutions. A numerical study is also performed to illustrate the effects of bound improving strategies, sample size, and user defined α on the solution speed. The major contributions of this chapter are three-fold: first, it captures the elastic demand in the risk-averse assignment. Second, a solution method with guaranteed ε convergence is developed. Third, a sampling method is developed for determining α .

Chapter 3 develops a formulation of the risk-averse traffic assignment that allows for the presence of multiple network-specific demons and therefore allows for the option to consider the effect of greater network uncertainty in terms of capacity degradation. Numerical studies are performed to demonstrate just how important the assumption on how many demons to use is. A numerical study is also carried out to consider the effect of relaxing the assumption on capacity reduction when links are selected for damage by one or more demons. The results show just how important it is to accurately capture the link capacity degradations. Relaxing this assumption also allows for more operational states to be modelled. Apart from these two contributions, this chapter also provides some further insights into the demon behaviour in selecting links to damage and what their optimal damage strategies are, demonstrates the existence of multiple solutions to this route choice game, and compares the network demon approach proposed here and the traditional OD-specific demon approach.

Chapter 4 develops a link-based formulation of the risk-averse traffic assignment via the NCP approach that allows for the presence of multiple network-specific demons and therefore allows for consideration of the effect of greater network uncertainty in terms of capacity degradation. This model can avoid path enumeration which is important particularly when dealing with multiple network demons which already increases computation times due to the large number of scenarios of link damage to be considered. This becomes even more significant when applying the model to larger real-life networks. Under this framework it was proven that there is at least one solution to this link-based problem and this problem is equivalent to the path-based one. This formulation can be reformulated into a VIP and an unconstrained minimisation program for solutions. This chapter also demonstrated the effect of assuming multiple demons and relaxing the capacity assumption to network reliability measurement through numerical studies. The results show just how important it is to accurately capture the link capacity degradations

particularly when the presence of more than one demon is assumed. Relaxing this assumption also allows for more operational states to be modelled. Apart from these contributions, this chapter also demonstrated the existence of stochastic Braess' paradox and reliability paradox (analogues of Braess' paradox for expected network travel cost and reliability) in risk-averse assignment, and provided some further insights into the implications for network design if these paradoxes are ignored.

Chapter 5 proposes a formulation of the reliable continuous network design problem by modelling the lower level problem of the NDP via the multiple network demon approach to risk-averse traffic assignment. The resulting model allows a planner to explicitly consider uncertainty in the transport network design problem and does not rely on data on user behaviour in a damaged network. Furthermore, the proposed network demon approach to modelling the users route choice behaviour in reliable network design allows for the presence of multiple network-specific demons and therefore allows for the option to consider the effect of greater network uncertainty in terms of capacity degradation. The framework of the RCNDP permits modelling of many more operational states than could be previously considered and determining the links most vulnerable to failure and allowing these links to receive priority in the improvement scheme thereby resulting in a reliable network design. By basing the network design on the expected travel costs of the users when confronted by supply uncertainty, this approach introduces improvements which cater for the worst case scenarios of network reliability. This chapter formulated the RCNDP as a single-level optimisation program, solved it through the GRG solution algorithm, and through numerical examples showed the importance of the type of demon to use and the amount of capacity reduction to assume. Overall, the results show just how important it is to accurately capture the link capacity degradations in reliable transport network design.

Chapter 6 extended the traditional DNDP by introducing the time dimension explicitly within the formulation. The resulting DNDP-T framework permits: modelling of time-dependent travel demands; capturing the gradually upgraded network over the planning period; the determination of optimal timing of project construction, project scale and phasing; determining optimal time-dependent tolls as well as the study of cost recovery over time. This chapter formulated the DNDP-T as a single-level optimisation program, solved it through a combination of the GRG and branch-and-bound solution algorithm, and through numerical examples, demonstrated the importance of introducing discrete capacity

variables to NDP, and the superiority of the DNDP-T over the traditional DNDP. The framework developed herein was also used to give some further insight on the effect of using different cost recovery principles.

Following this, *chapter 6* also proposed a time-dependent formulation of the reliable discrete network design problem (RDNDP-T) by modelling the lower level problem under risk-averse traffic assignment. The resulting model allows a planner to explicitly consider uncertainty in the network design problem and does not rely on data on user behaviour in a damaged network. Similar to the DNDP-T model, the time-dependent framework of the RDNDP-T permits modelling of: time-dependent travel demands; capturing of the gradually upgraded network over the planning period; the determination of optimal timing of project construction; project scale and phasing; determining the critical links and allowing these links to receive priority in the improvement scheme, thereby resulting in a reliable network design. By basing the network design on the expected travel costs of the users when confronted by supply uncertainty, this approach introduces improvements which cater for the worst case scenarios of network reliability. This chapter formulated the RDNDP-T as a single-level optimisation program, solved it through the solution method described in chapter 6, and through a numerical example demonstrated how the model determines and designs the critical links under a budget when either OD-specific or network-specific demands are assumed.

7.3 CRITICAL ASSESSMENT

In the wider field of network reliability and reliable network design, a number of useful indicators have been proposed to measure the transport network reliability and many techniques have been developed to calculate these indicators. The usefulness and limitations of these indicators was discussed in chapter 1 of this thesis and while there is no single 'perfect' indicator or technique, the choice of which to use is largely determined by the viewpoint of the analyst and the amount of information or data available on capacity reducing incidents. In addition, the type and range of interventions being considered is also an important factor. In the absence of available data on previous incidents in a network it is clear that it is much more difficult and laborious to estimate event probabilities reasonably accurately than it is to estimate the impact of network degradation if it occurs. The game theoretic approach tries to circumvent this problem by looking at worst-case scenarios.

One common criticism of the game theoretic approach is that it is overly pessimistic since the demons are forced to damage at least one link per day. However, if one wishes to assess the network reliability with the motivation of finding critical links and determining worst-case scenarios then the game theoretic approach is one such useful way of doing so. The multiple network demon game theoretic approach, developed in this thesis, raises the burning question of – how to determine exactly how many demons to be chosen to be active in the network (or in an OD pair)? This is a difficult question to answer since it is also linked to the other question of how much one should assume that each demon contributes to the link capacity reduction when the demon selects a link to damage. Bell's formulations limit the number of demons to be equal to the number of OD pairs. However, under the multiple network-specific demon approach proposed in this thesis any number of demons may be assumed and this number can be greater than or less than the number of OD pairs, unlike the Bell and Cassir (2002) approach. One possible method may be to choose the number of demons that are suitable based on some available data on incidents that have led to past link failures in the network. Very often, however, this kind of data may be unavailable. Furthermore, the question of what constitutes the 'worst-case scenario' needs to be considered in order to decide on the number of demons to assume. In chapter 5, the possibility of absurd results arising from applying the OD demon approach to the Chicago regional network of 3,202,310 OD pairs and 39,018 links was discussed. Clearly this very large number of demons to damage considerably fewer links than OD pairs would be inappropriate. It was concluded that the network demon approach is more appropriate for application to realistic networks since it allows the decision maker the discretion to specify the number of demons to assume. On the other hand, in some networks there may exist many more links than OD pairs in which case the proposed network demon approach can allow for more demons if necessary thereby resulting in higher expected network costs and hence a better 'worst-case scenario' design. Thus, it can be concluded that one major advantage of the network demon approach is that it allows one the flexibility to choose as many demons as is appropriate.

Nicholson et al. (2003) suggest that there is a need to develop techniques for estimating event probabilities, even if only approximately. One could then obtain an approximate estimate of the expected cost of degradation and ascertain the sensitivity of that estimate to variations in the approximate probabilities. One could then decide whether to proceed with a more detailed and thorough evaluation of the impacts. Nonetheless, while the game theory approach does not predict which links will fail (with a certain probability) it remains

attractive since it gives information on which links will impact most on the network users *when selected* for failure. The broad measure of expected network cost can give a sense of the overall network reliability while specific measures such as the travel time and encountered reliabilities can provide more information on the probabilities of arriving on time and not encountering link failure during a journey. Moreover, the approach developed here can be used to design reliable road networks as demonstrated in chapters 5-6 of this thesis.

In terms of solution methods, chapter 2 of this thesis proposed the modified α BB algorithm which compares favourably with the heuristic method proposed to solve Bell's models. On the other hand the results given here are based on a very small network. Different, larger networks should be adopted to validate whether the proposed strategy is always efficient. The models proposed in chapters 3-6 of this thesis were solved by the generalized reduced gradient (GRG) algorithm. This can be a good solution method to solve problems in small networks. However, when considering large networks with many feasible paths, it would be difficult and time-consuming to find a global solution. In chapter 4, a link-based formulation was proposed which obviated the need for path enumeration and makes the formulation easier to solve for larger networks. Notwithstanding this, there would still be a large number of variables to determine and the presence of many demons would still generate a large number of scenarios and multiple solutions. Therefore, it would be more desirable to find a global optimum by applying or designing more efficient solution algorithms and computational methods.

Nevertheless, these issues offer opportunities for further research. A detailed discussion of potential and worthy future research directions is provided in the next section.

7.4 RECOMMENDATIONS FOR FUTURE RESEARCH

This thesis opens up many possible future research directions. A number of these extensions are common to every chapter and are outlined in sections 7.4.1-7.4.4. Following this, a number of possible extensions specific to each chapter are given.

7.4.1 Formulating the Problem as a Stackelberg Game

The formulations of risk-averse traffic assignment proposed in chapters 2-4 and used to formulate the reliable network design formulations in chapters 5-6 are defined as a Nash game where the demon(s) cannot predict user response to link damage. An alternative to these models would be to formulate the problems as Stackelberg games in which the demon is placed at the upper level. This may allow for a better approximation of the worst-case expected travel cost. However, to determine whether this approach is a good model of risk-averse behaviour requires further investigation. Bell and Cassir (2002) argue that in some contexts, such as network testing, this may be an appropriate approach. However, even if pessimistic, network users retain uncertainty about which link will be damaged and so route choice is unlikely to be controlled as directly by the demon(s) as the Stackelberg game would imply. Nonetheless, the nature of the route choice game is worthy of further investigation.

7.4.2 Application to Larger Networks

The studies throughout the thesis on the behaviour of multiple demons are carried out on small test networks. Different, larger networks are required to further investigate the demons behaviour and to determine if they always select certain links to damage over others in the network with a higher probability, in particular those critical links near the origin and destination nodes.

7.4.3 The Game of Increasing Adversity

The game between the demons and the network users can also be extended to the case where the degree of adversity towards network users shown in the choice of scenarios is varied. Bell (1999) suggests that the demon could choose a scenario according to a logit model with a sensitivity parameter α . When α is zero, each scenario has an equal probability of being selected since the demon is insensitive to travel costs. As α increases, the scenario choice probabilities change (along with the route choice probabilities) so that expected network cost is progressively reduced. The sensitivity parameter could therefore be said to measure the degree of pessimism of network users. This extension would allow the planner the discretion to set the degree of pessimism of network users, perhaps based on surveys of commuters or other information collected. Furthermore, the degree of risk-

averseness displayed by network users is likely also to depend on the purpose of the trip with those trips of a higher importance (for example, a student travelling to University to sit an exam) attracting a greater level of risk-averseness. Formulating the games in this thesis in this manner represents a worthy future extension.

7.4.4 Pessimistic versus Realistic Demon Link Selection Probabilities

The formulations proposed in this thesis and the ones proposed in Bell (2000) and Bell and Cassir (2002) do not require any information about link incident or link selection/scenario probabilities. In many respects this can be advantageous particularly when this information is absent or unreliable. However, in its present format, this approach cannot take into account *any* previous information about link incident probabilities. This implies that route flows are obtained based on ‘extremely pessimistic’ incident probabilities. Throughout this thesis the advantages of this pessimistic approach have been highlighted (for example, it localises the vulnerable links and gives the worst-case scenarios). However, modifying the approach to a framework in which the information on the likelihood of an incident can be exploited to choose a ‘reasonable’ link incident probability represents a worthy future research direction. Recently, in a working paper, Nagae and Akamatsu (2009) introduced the concept of *relative entropy* into the hazardous material (Hazmat) risk-averse routing model. The relative entropy of the subjective (pessimistic) probability with respect to the objective (nominal) probability, measures the discrepancy between these two link incident probabilities. This allows the link incident probabilities to be selected pessimistically but not to the extent so as to be extremely unrealistic. A similar concept could be adopted for the models developed in this thesis.

7.4.5 Other Future Research Directions

There are also a number of worthy research directions arising from specific issues in each chapter of this thesis and these issues are discussed in the following paragraphs.

Chapter 2 opens up many future research directions. First, the results given here are based on a very small network. Different, larger networks should be adopted to validate whether the proposed strategy is always efficient. Second, based on computation experience (not shown here), the sampling approach to obtain the α value is better than the approach based on the interval Hessian derived from the natural interval extension. Other approaches

to derive the interval Hessian may result in shorter computation time. Also, the use of other smooth gap functions may improve the speed. These issues are worthy of future research.

Chapter 3 also opens up many future research directions. First, the proposed multiple network demon formulation and the ones in Bell and Cassir (2002) and chapter 2/Szeto et al. (2006) could be formulated into one model of risk-averse traffic assignment. This would have the benefit of providing one clearly defined framework for the problem. Second, the very important question of how many demons are appropriate to assume remains unanswered. Extensive studies on real networks involving data on previous incidents are required to rectify this. Third, the studies here on the behaviour of two demons are carried out on small test networks. Different, larger networks are required to further investigate the demons behaviour and to determine if they always select certain links to damage over others in the network with a higher probability, in particular those critical links near the origin and destination nodes.

Chapter 4 raises some interesting questions and poses some issues requiring future research. First, it is not clear whether the occurrence of stochastic Braess' paradox implies the occurrence of reliability paradox and vice versa. This is left to future studies. Second, the number of users here is fixed. In reality, the demand is elastic and uncertain. Incorporating demand elasticity (as in chapter 2/Szeto et al., 2006) and demand stochasticity (as in Ukkusuri et al., 2007) into the present modelling framework is therefore another important research extension.

Chapter 5 opens up many future research directions. First, the results given here are based on small networks. Different, larger networks should be adopted to validate whether the RCNDP model is always superior to other approaches. Second, the RCNDP models presented here adopt ETSTT as the objective function. One important future extension would be to adopt other objective functions which incorporate generalised costs, particularly when tolling strategies are being considered. This may bring up some interesting and important future research directions.

Chapter 6 proposes the DNDP-T and RDNDP-T models. Both of these models and associated numerical studies offer some interesting future research directions. First, the proposed formulation of the DNDP-T does not take into account changing patterns in land use over time. One interesting study would be to incorporate land use elements into this

model. Second, only the costs associated with capacity improvements are considered in the DNDP-T model proposed in this chapter. Throughout the world it is becoming increasingly common for private companies to build and operate new highways for at least a generation (or approximately twenty years) after construction has been completed. Therefore, one possible future study is to extend the overall cost recovery considerations to include operations and maintenance costs and to consider the effect on the optimal capacity improvements over a much longer planning horizon. Third, the algorithm used here to generate solutions is a general purpose algorithm and may not be best suited to handle the highly non-convex DNDP-T. Further research is required to develop a suitable solution scheme to handle the discrete capacity constraints, particularly when one considers larger networks, and the non-convexity. Fourth, the proposed model is only applicable to the situation when the stated assumptions in the text are nearly satisfied including that the demand structure and financial environment do not substantially differ from the ones predicted. The uncertainties in the demand structure and financial environment are not considered in this chapter. Capturing these uncertainties in the formulation so as to allow alternative designs represents a worthy future research direction. Furthermore, it is also acknowledged here that most of the major infrastructure investment decisions are made through political processes but despite this the use of models such as the one proposed in this chapter are helpful to the planner when involved in the decision making process.

Lastly, most of the extensions mentioned above can also be incorporated into the RDNDP-T model. In addition, the RDNDP-T model proposed in this thesis restricts the investment money to be used solely for link widening. However, it may also be interesting in the future to consider the effect of using the budget to improve the physical robustness of the links, i.e., improving the structural toughness of the link against natural disaster.

Finally, using the risk-averse frameworks proposed in this thesis, one can assess the reliability of transport networks, develop a reliable network design, and analyse the effect of assumptions on traveller behaviour in each approach on network design and reliability. This may bring up some interesting and important future research directions.

REFERENCES

- Abadie, J. and Carpentier, J. (1969) Generalization of the Wolfe Reduced Gradient Method to the Case of Nonlinear Constraints. In R. Fletcher (ed.), *Optimization*, pp. 37-47. New York: Academic Press.
- Abdel-Aty, M.A., Kitamura, R. and Jovanis, P.P. (1997) Using Stated Preference Data for Studying the Effect of Advanced Traffic Information on Drivers' Route Choice. *Transportation Research*, Vol. 5C, pp. 39-50.
- Adjiman, C.S., Androulakis, I.P. and Floudas, C.A. (1998b) A Global Optimisation Method, α BB, for General Twice Differentiable Constrained NLPs – II. Implementation and Computational Results. *Computers and Chemical Engineering*, Vol. 22, No. 9, pp. 1159-1179.
- Adjiman, C.S., Dallwig, S., Floudas, C.A. and Neumaier, A. (1998a) A Global Optimisation Method, α BB, for General Twice Differentiable Constrained NLPs – I. Theoretical Advances. *Computers and Chemical Engineering*, Vol. 22, No. 9, pp. 1137-1158.
- Arnott, R. and Small, K. (1994) The Economics of Traffic Congestion. *American Scientist*, Vol. 82, pp. 446-455.
- Asakura, Y. (1996) Reliability Measures of an Origin and Destination Pair in a Deteriorated Road Network with Variable Flows. In M.G.H. Bell (ed.), *Transportation Networks: Recent Methodological Advances*, pp. 273-288. Pergamon Press, Oxford.
- Asakura, Y., Hato, E. and Kashiwadani, M. (2003) Stochastic Network Design Problem: An Optimal Link Investment Model for Reliable Network. In M.G.H. Bell and Y. Iida (eds.), *The Network Reliability of Transport: Proceedings of the 1st International Symposium on Transportation Network Reliability*, pp 245-259. Pergamon, Oxford.

- Asakura, Y. and Kashiwadani, M. (1991) Road Network Reliability Caused by Daily Fluctuation of Traffic Flow. *Proceedings of the 19th PTRC Summer Annual Meeting*, 1991, pp. 73-84. Brighton.
- Bass, T. (1992) Road to Ruin. *Discover*, May, pp. 56-61.
- Bazzan, A.L.C. and Klugl, F. (2005) Case Studies on the Braess Paradox: Simulating Route Recommendation and Learning in Abstract and Microscopic Models. *Transportation Research*, Vol. 13C, pp. 299-319.
- Bell, M.G.H. (1999) Measuring Network Reliability: A Game Theoretic Approach. *Journal of Advanced Transportation*, Vol. 32, pp. 135-146.
- Bell, M.G.H. (2000) A Game Theory Approach to Measuring the Performance Reliability of Transport Networks. *Transportation Research*, Vol. 34B, pp. 533-546.
- Bell, M.G.H. (2003) The Use of Game Theory to Measure the Vulnerability of Stochastic Networks. *IEEE Transactions on Reliability*, Vol. 52, No. 1, pp. 63-68.
- Bell, M.G.H. (2004) Games, Heuristics, and Risk Averseness in Vehicle Routing Problems. *Journal of Urban Planning and Development*, Vol. 130, No. 1, pp. 97-41.
- Bell, M.G.H. (2007) Mixed Routing Strategies for Hazardous Materials: Decision-Making under Complete Uncertainty. *International Journal of Sustainable Transportation*, Vol. 1, No. 2, pp. 133-142.
- Bell, M.G.H. and Cassir, C. (2001) *Reliability of Transport Networks*. Research Studies Press, UK.
- Bell, M.G.H. and Cassir, C. (2002) Risk-Averse User Equilibrium Traffic Assignment: An Application of Game Theory. *Transportation Research*, Vol. 36B, pp. 671-681.
- Bell, M.G.H., Cassir, C., Iida, Y. and Lam, W.H.K. (1999) A Sensitivity-Based Approach to Network Reliability Assessment. *Proceedings of the 14th International Symposium on Transportation and Traffic Theory*, Jerusalem, pp. 283-300.

- Bell, M.G.H., Kanturska, U., Schmöcker, J.-D. and Fonzone, A. (2008) Attacker-Defender Models and Road Network Vulnerability. *Philosophical Transactions of the Royal Society A*. Vol. 366, pp. 1893-1906.
- Bell, M.G.H. and Iida, Y. (1997) *Transportation Network Analysis*. John Wiley and Sons, Chichester.
- Bell, M.G.H. and Iida, Y. (2003) *The Network Reliability of Transport: Proceedings of the 1st International Symposium on Transportation Network Reliability (INSTR)*. Pergamon, Oxford.
- Bell, M.G.H. and Schmöcker, J.D. (2002) Public Transport Network Reliability: Topological Effects. *Proceedings of the 3rd International Conference on Traffic and Transportation Studies (ICTTS)*, 2002, pp. 453 – 460. Guilin.
- Berdica, K. (2000) Analyzing Vulnerability in the Road Transportation System: Putting Theory into Practice Using Sweden as an Example. Working paper TRITA-IP FR 00-76, 2000, Department of Infrastructure and Planning, Royal Institute of Technology, Stockholm, Sweden.
- Berdica, K. (2001) Vulnerability – A Model-Based Case Study of the Road Network in the City of Stockholm. Paper Presented at: *1st International Symposium on Transportation Network Reliability*. Kyoto, Japan.
- Berdica, K. (2002) An Introduction to Road Vulnerability: What Has Been Done, Is Done and Should Be Done. *Transport Policy*, Vol. 9, No. 2, pp. 117-127.
- Boyce, D.E. (1984) Urban Transportation Network Equilibrium and Design Models: Recent Achievements and Future Prospects. *Environment and Planning*, Vol. 16A, pp. 1445-74.
- Boyce, D.E., Ran, B. and Li, Y. (1998) Considering Travellers' Risk-Taking Behavior in Dynamic Traffic Assignment. In M.G.H. Bell (ed.), *Transportation Networks: Recent Methodological Advances*, pp. 67-81. Oxford: Elsevier.

- Braess, D. and Uber ein Paradoxon der Verkehrsplanung. (1968) *Unternehmensforschung*, Vol. 12, pp. 258–268.
- Cassir, C. and Bell, M.G.H. (2001) The N+M Person Game Approach to Network Reliability. In M.G.H. Bell and C. Cassir (eds.), *Reliability of Transport Networks*. Research Studies Press, UK.
- Cassir, C., Bell, M.G.H., and Iida, Y. (2001) Reliability of Transport Networks: Introduction. In M.G.H. Bell and C. Cassir (eds.), *Reliability of Transport Networks*, pp. vii-xi. Research Studies Press Ltd., UK.
- Cassir, C., Bell, M.G.H. and Schmöcker, J.-D. (2003) A Normative Assessment of Transport Network Reliability Based on Game Theory. In M.G.H. Bell and Y. Iida (eds.), *The Network Reliability of Transport*, pp. 225-244. Pergamon, Oxford.
- Chan, K.S. and Lam, W.H.K. (2005) Impact of Road Pricing on the Network Reliability. *Journal of the Eastern Asia Society for Transportation Studies*, Vol. 6, pp. 2060 – 2075.
- Chen, M.Y. and Alfa, A.S. (1991) A Network Design Algorithm Using a Stochastic Incremental Traffic Assignment Approach. *Transportation Science*, Vol. 25, No. 3, pp. 215-224.
- Chen, A., Ji, Z.W. and Recker, W. (2002) Travel Time Reliability with Risk-Sensitive Travellers. In *Transportation Research Record: Journal of Transportation Research Board*, No. 1783, Transportation Research Board of the National Academies, Washington, D.C., 2002, pp. 27-33.
- Chen, A., Kim, J., Zhou, Z. and Chootinan, P. (2007) Alpha Reliable Network Design Problem. In *Transportation Research Record: Journal of Transportation Research Board*, No. 2029, pp. 49-57.
- Chen, A. and Recker, W. (2000) Considering Risk-Taking Behaviour in Travel Time Reliability. Paper Presented in: *Matsuyama Workshop on Transport Network Analysis*, Matsuyama, Kyoto, Japan.

- Chen, A., Subprasom, K. and Ji, Z. (2003) Mean-Variance Model for the Build-Operate-Transfer Model under Demand Uncertainty. In *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 1857, pp. 93-101.
- Chen, A., Tatneni, M., Lee, D.-H. and Yang, H. (2000) Effect of Route Choice Models on Estimating Network Capacity Reliability. *Transportation Research Record Journal of the Transportation Research Board*, Vol.1733, pp. 63-70.
- Chen, A. and Yang, C. (2004) Stochastic Transportation Network Design Problem with Spatial Equity Constraint. In *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 1882, pp. 97-104.
- Chen, A., Yang, H., Lo, H. and Tang, W.H. (2002) Capacity Reliability of a Road Network: An Assessment Methodology and Numerical Results. *Transportation Research*, Vol. 36B, pp. 225 – 252.
- Chiou, S.W. (2005) Bilevel Programming for the Continuous Transport Network Design Problem. *Transportation Research*, Vol. 39B, No. 4, pp. 361-383.
- Chootinan, P., Wong, S.C. and Chen, A. (2005) A Reliability-Based Network Design Problem. *Journal of Advanced Transportation*, Vol. 39, No. 3, pp. 247-270.
- Connors, R., Sumalee, A. and Watling, D. (2005) Equitable Network Design. *Journal of Eastern Asia Society for Transportation Studies*, Vol. 6, No. 1, pp. 1382-1397.
- Dafermos, S. and Nagurney, A. (1984) On Some Traffic Equilibrium Theory Paradoxes. *Transportation Research*, Vol. 18B, pp. 101–110.
- Dalziell, E.P. and Nicholson, A.J. (2001) Risk and Impact of Natural Hazards on a Road Network. *ASCE Journal of Transportation Engineering*, Vol. 127, No. 2, pp. 159-166.
- Davis, G.A. (1994) Exact Local Solution of the Continuous Network Design Problem via Stochastic User Equilibrium Assignment. *Transportation Research*, Vol. 28B, No. 1, pp. 61-75.

- Davis, G.A. and Sanderson, K. (2001) *Building Our Way out of congestion? Highway Capacity for Twin cities*, Report No. MN/RC – 2002-01, Minnesota Department of Transportation, St. Paul, Minnesota, USA.
- D'Este, G.M. and Taylor, M.A.P. (2003) Network Vulnerability: An Issue for Regional, National and International Strategic Networks. In M.G.H. Bell and Y. Iida (eds.), *The Network Reliability of Transport*, pp. 23-44. Pergamon, Oxford.
- Du, Z.P. and Nicholson, A.J. (1997) Degradable Transportation Systems: Sensitivity and Reliability Analysis. *Transportation Research*, Vol. 31B, No. 3, pp. 225-237.
- Fisk, C. (1979) More Paradoxes in the Equilibrium Assignment Problem. *Transportation Research*, Vol. 1B, pp. 305-309.
- Friesz, T.L. (1985) Transportation Network Equilibrium, Design and Aggregation: Key Developments and Research Opportunities. *Transportation Research*, Vol. 19A, 5-6, pp. 413-427.
- Friesz, T.L. and Shah, S. (2001) An Overview of Non-Traditional Formulations of Static and Dynamic Network Design. *Transportation Research*, Vol. 35B, No. 1, pp. 5-22.
- Gordon, A., Van Vuren, T., Watling, D.P., Polak, J., Noland, R.B., Porter, S. and Taylor, N. (2001) Incorporating Variable Travel Time Effects into Route Choice Models. *Proceedings of the PTRC European Transport Conference* (Methodological Innovations), PTRC Education and Research Services Ltd., London.
- Heydecker, B.G. (2002) Dynamic Equilibrium Network Design. In M.A.P. Taylor (ed.), *Transportation and Traffic Theory*, pp. 349-370. Elsevier Science.
- Hollander, Y.Z. and Prashker, J.N. (2006) The Applicability of Non-Cooperative Game Theory in Transport Analysis. *Transportation*, Vol. 33, No. 5, pp. 481-496.

- Huang, H. J. and Bell, M. G. H. (1998) Continuous Equilibrium Network Design Problem with Elastic Demand: Derivative-Free Solution Methods. In M.G.H. Bell (ed.), *Transportation Networks: Recent Methodological Advances*, pp. 175-193. Pergamon, Oxford.
- Jackson, W. B. and Jucker, J.V. (1981) An Empirical Study of Travel Time Variability and Travel Choice Behaviour. *Transportation Science*, Vol. 16, No. 4, pp. 460-475.
- Kanturska, U., Schmöcker, J.-D., Fonzone, A. and Bell, M.G.H. (2009) Improving Reliability Through Multi-Path Routing and Link Defence. In V.M. Bier and M.N. Azaiez (eds.), *Game Theoretic Risk Analysis of Security Threats*, pp. 199-228. Springer, New York.
- Kanzow, C. and Fukushima, M. (1996) Equivalence of the Generalized Complementarity Problem to Differentiable Unconstrained Minimization. *Journal of Optimization Theory and Applications*, Vol. 90, No. 3, pp. 581-603.
- Lam, W.H.K., Tam, M.L. and Bell, M.G.H. (2002) Optimal Road Tolls and Parking Charges for Balancing Demand and Supply of Road Transport Facilities. In M.A.P. Taylor (ed.), *Transportation and Traffic Theory in the 21st Century: Proceedings of the 15th International Symposium on Transportation and Traffic Theory*, pp. 561-582. New York: Elsevier Science.
- Lam, W.H.K. and Zhang, N. (2000) A New Concept of Travel Demand Satisfaction Reliability for Assessing Road Network Performance. Presented in *Matsuyama Workshop on Transport Network Analysis*, Matsuyama, Japan.
- LeBlanc, L.J. (1975) An Algorithm for Discrete Network Design Problem. *Transportation Science*, Vol. 9, No. 3, pp. 183-199.
- Lim, Y.T., Heydecker, B. and Lee, S. (2004) A Continuous Network Design Model in Stochastic User Equilibrium Based on Sensitivity Analysis. *Proceedings of the International Workshop on Behaviour in Networks*, pp. 223-233.

- Liu, H.X., Ban, X., Ran, B. and Mirchandani, P. (2002) Analytical Dynamic Traffic Assignment Model with Probabilistic Network and Travellers' Perceptions. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1783, pp. 125-133.
- Liu, H.X., Recker, W. and Chen, A. (2004) Uncovering the Contribution of Travel Time Reliability to Dynamic Route Choice Using Real-Time Loop Data. *Transportation Research*, Vol. 38A, No.6, pp. 435-453.
- Lo, H. (2002) Trip Travel Time Reliability in Degradable Transport Networks. In M.A.P. Taylor (ed.), *Transportation and Traffic Theory in the 21st Century: Proceedings of the 15th International Symposium on Transportation and Traffic Theory*, pp. 541-560. New York: Elsevier Science.
- Lo, H. and Chen, A. (2000) Traffic Equilibrium Problem with Route-Specific Costs: Formulation and Algorithms. *Transportation Research*, Vol. 34B, pp. 493-513.
- Lo, H.K. and Szeto, W.Y. (2009) Time-Dependent Transport Network Design under Cost-Recovery, *Transportation Research B*, Vol. 43B, No. 1, pp. 142-158.
- Lo, H.K. and Szeto, W.Y. (2004) Planning Transport Network Improvement over Time. In D.H. Lee (ed.), *Urban and Regional Transportation Modelling: Essays in Honour of David Boyce*, Chapter 9, pp. 157-176. Edward Elgar.
- Lo, H.K. and Szeto, W.Y. (2003) Time-Dependent Transport Network Design: A Study on Budget Sensitivity. *Journal of Eastern Asia Society for Transportation Studies*, Vol. 5, pp. 1124-1139.
- Lo, H. and Tung, Y.K. (2000) A Chance Constrained Network Capacity Model. In M.G.H. Bell and C. Cassir (eds.), *Reliability of Transport Networks*, pp. 159-172. Research Studies Press Ltd., UK.
- Lo, H. and Tung, Y.K. (2003) Network with Degradable Links: Capacity Analysis and Design. *Transportation Research*, Vol. 37B, No. 4, pp. 345-363.

- Magnanti, T.L. and Wong, R.T. (1984) Network Design and Transportation Planning: Models and Algorithms. *Transportation Science*, Vol.18, No.1, pp. 1-55.
- Maher, M.J., Zhang, X. and Van Vliet, D. (2001) A Bi-Level Programming Approach for Trip Matrix Estimation and Traffic Control Problems with Stochastic User Equilibrium Link Flows. *Transportation Research*, Vol. 35B, No. 1, pp. 23-40.
- Maranas, C.D. and Floudas, C.A. (1994a) A Deterministic Global Optimisation Approach for Molecular Structure Determination. *Journal of Chemical Physics*, Vol. 100, No. 2, pp. 1247-1261.
- Maranas, C.D. and Floudas, C.A. (1994b) Global Minimum Potential Energy Conformations for Small Molecules. *Journal of Global Optimisation*, Vol. 4, No. 3, pp. 135-140.
- Maranas, C.D. and Floudas, C.A. (1993) Global Optimization for Molecular Conformation Problems. *Annals of Operations Research*, Vol. 42, pp. 85-117.
- Maranas, C.D. and Floudas, C.A. (1992) A Global Optimization Approach for Lennard-Jones Micro-Clusters. *Journal of Chemical Physics*, Vol. 97, No. 10, pp. 7667-7677.
- Meng, Q., Lee, D. H., Yang, H. and Huang, H. J. (2004) Transportation Network Optimization Problems with Stochastic User Equilibrium Constraints. In Transportation Research Record: *Journal of Transportation Research Board*, No. 1882, pp. 113-119.
- Meng, Q., Yang, H. and Bell, M.G.H. (2001) An Equivalent Continuously Differentiable Model and Locally Convergent Algorithm for the Continuous Network Design Problem. *Transportation Research*, Vol. 35B, No. 1, pp. 83-105.
- Mirchandani, P. and Soroush, H. (1987) Generalized Traffic Equilibrium with Probabilistic Travel Times and Perceptions. *Transportation Science*, Vol. 21, pp. 133-152.
- Murchland, J. D. (1970) Braess' Paradox of Traffic Flow. *Transportation Research*, Vol. 4, pp. 391-394.

- Nagurney, A. (1993) *Network Economics: A Variational Inequality Approach*. Kluwer Academic Publishers, Norwell, Massachusetts.
- Nash, J. (1951) Non-Cooperative Games. *Annals of Mathematics*, Vol. 54, pp. 286–295.
- Negae, T. and Akamatsu, T. (2009) Robust Mixed Strategies for Routing of Hazardous Materials. *Operations Research*, (submitted).
- Nicholson, A., Schmöcker, J.D., Bell, M.G.H. and Iida, Y. (2003) Assessing Transport Network Reliability: Malevolence and User Knowledge. In M.G.H. Bell and Y. Iida (eds.), *The Network Reliability of Transport*, pp. 1-22. Pergamon, Oxford.
- Nie, Y. and Fan, Y.Y. (2006) Arriving-on-Time Problem – Discrete Algorithm that Ensures Convergence. In *Transportation Research Record: Journal of Transportation Research Board*, No.1964, pp. 193-200.
- Noland, R.B. and Polak, J.W. (2002) Travel Time Variability: A Review of Theoretical and Empirical Issues. *Transport Reviews* Vol. 22, No. 1, pp. 39-54.
- Noland, R.B., Small, K.A., Koskenoja, P.M. and Chu, X. (1998) Simulating Travel Reliability. *Regional Science and Urban Economics* Vol. 28, pp. 535-564.
- O'Brien, L. and O'Mahony, M. (2010) Time-Dependent Reliable Transport Network Design. *The Annals of Regional Science*, (submitted).
- O'Brien, L. and Szeto, W.Y. (2007) The Discrete Network Design Problem Over Time. *HKIE Transactions*, Vol. 14, No. 4, pp. 47-55.
- O'Brien, L. and Szeto, W.Y. (2008) Designing Reliable Transport Networks: The Multiple Network Spoiler Approach. *HKIE Transactions*, Vol. 15, No. 4, pp. 1-9.

- O'Brien, L., Szeto, W.Y. and O'Mahony, M. (2006) Transport Network Design under Risk-Averse Traffic Assignment: A Game Theory Approach. *Proceedings of the EWGT/EURO Conference Handling Uncertainty in Transportation: Analyses, New Paradigms, Applications and Comparisons*, pp. 494-502. Bari, Italy, 2006.
- O'Brien, L., Szeto, W.Y. and O'Mahony, M. (2007a) Time-Dependent Discrete Network Design Problem. *Journal of Eastern Asia Society for Transportation Studies*, Vol. 7, pp. 2230-2245.
- O'Brien, L., Szeto, W.Y. and O'Mahony, M. (2007b) Time-Dependent Discrete Network Design Problem. *Proceedings of the 7th Conference of the Eastern Asia Society for Transportation Studies*, Dalian, China, September 2007.
- Pas, E. and Principio, S.I. (1997) Braess' Paradox: Some New Insights. *Transportation Research*, Vol. 31B, pp. 265-276.
- Satayapiwat, P., Suksomboon, K. and Aswakul, C. (2008) Vulnerability Analysis in Multi-Commodity Stochastic Networks by Game Theory. *5th International Conference on Electrical Engineering/Electronics, Computers, Telecommunications and Information Technology, ECTI-CON*, pp. 357-360.
- Schmöcker, J.-D., Bell, M.G.H. and Lee, C. (2002) An Application of Congested Transit Network Loading with the Markov Chain Approach. *Proceedings of the 9th Meeting of the Euro Working Group on Transportation*, pp. 443-447. Bari, Italy, 2002.
- Shao, H., Lam, W.H.K., Tam, M.L. and Yuan X.M. (2008) Modelling Rain Effects on Risk-Taking Behaviours of Multi-User Classes in Road Networks with Uncertainty. *Journal of Advanced Transportation*, Vol. 42, No.3, pp. 265-290.
- Sheffi, Y. (1985) *Urban Transportation Networks – Equilibrium Analysis with Mathematical Programming Methods*, Prentice-Hall, Englewood Cliffs, NJ.
- Sumalee, A., Connors, R. and Batley R. (2005) The Applicability of Prospect Theory to the Analysis of Transport Networks. *Proceedings of the European Transport Conference*. Strasbourg, France.

- Sumalee, A. and Kurauchi, F. (2006) Guest editorial: Reliability and Emergency Issues in Transportation Network Analysis. *Networks and Spatial Economics*, Vol. 6, No. 3-4, pp. 169-172.
- Sumalee, A., Watling, D.P. and Nakayama, S. (2006) Reliable Network Design Problem - Case with Uncertain Demand and Total Travel Time Reliability. In *Transportation Research Record: Journal of Transportation Research Board*, No. 1964, pp. 81-90.
- Szeto, W.Y. and Lo, H.K. (2005) Strategies for Road Network Design over Time: Robustness Under Uncertainty. *Transportmetrica*, Vol. 1, No. 1, pp. 47-63.
- Szeto, W.Y. and Lo, H.K. (2006) Transportation Network Improvement and Tolling Strategies: The Issue of Intergeneration Equity. *Transportation Research*, Vol. 40A, No. 3, pp. 227-243.
- Szeto, W.Y., O'Brien, L. and O'Mahony, M. (2006) Risk-Averse Traffic Assignment with Elastic Demands: NCP Formulation and Solution Method for Assessing Performance Reliability. *Networks and Spatial Economics*, Vol. 6, pp. 313-332.
- Szeto, W.Y., O'Brien, L. and O'Mahony, M. (2007) Generalisation of Risk-Averse Traffic Assignment. In R. Allsop, M.G.H. Bell and B.G. Heydecker (eds.), *Transportation and Traffic Theory*, pp. 127-153. Elsevier, Amsterdam.
- Szeto, W.Y., O'Brien, L. and O'Mahony, M. (2009) Measuring Network Reliability Considering Paradoxes: The Multiple Network Demon Approach. In *Transportation Research Record: Journal of the Transportation Research Board*, (In Press).
- Tzeng, G.H. and Tsaur, S.H. (1997) Application of Multiple Criteria Decision Making for Network Improvement. *Journal of Advanced Transportation*, Vol. 31, No. 1, pp. 49-74.
- Uchida, T. and Iida, Y. (1993) Risk Assignment: A New Traffic Assignment Model Considering Risk of Travel Time Variation. In C.F. Daganzo (ed.), *Proceedings of the 12th International Symposium on Transportation and Traffic Theory*, pp. 89-105. Elsevier, Amsterdam.

- Ukkusuri, S.V., Mathew, T.V. and Travis Waller, S. (2007) Robust Transportation Network Design under Demand Uncertainty. *Computer-Aided Civil and Infrastructure Engineering*, Vol.22, No.1, pp. 6-18.
- Ukkusuri, S.V. and Patil, G. (2009) Multi-Period Transportation Network Design under Demand Uncertainty. *Transportation Research*, Vol. 43B, pp. 625-642.
- Wakabayahsi, H. and Iida, Y. (1992) Upper and Lower Bounds of Terminal Reliability of Road Networks: An Efficient Method with Boolean Algebra. *Journal of Natural Disaster Science*, Vol. 14, pp. 29-44.
- Waller, S.T. and Ziliaskopoulos, A.K. (2001) Stochastic Dynamic Network Design Problem. In *Transportation Research Record: Journal of Transportation Research Record Board*, No. 1771, pp. 106-113.
- Wan, K.H., Lo., H. and Yip, C.W. (2002) Optimal Integrated Transit Network Design. In C.P. Wang et al. (ed.), *Proceedings of the Seventh International Conference of the ASCE Applications of Advanced Technologies in Transportation, USA: American Society of Civil Engineers*, pp. 736-743.
- Wardrop, J. (1952) Some Theoretical Aspects of Road Traffic Research. *Proceedings of the Institute of Civil Engineers, Part II*, pp. 325-378.
- Wong, S.C. and Yang, H. (1997) Reserve Capacity of a Signal-Controlled Road Network. *Transportation Research*, Vol. 31B, pp. 397-402.
- Wu, Z.X., Lam, W.H.K. and Chan, K.S. (2005) Multi-Modal Network Design: Selection of Pedestrianisation Location. *Journal of Eastern Asia Society for Transportation Studies*, Vol. 6, No. 1, pp. 2275-2290.
- Yang, H. and Bell, M.G.H. (1998) Models and Algorithms for Road Network Design: A Review and some New Developments. *Transport Reviews*, Vol. 18, No. 3, pp. 257-278.

- Yang, H., Lo, H.K. and Tang, W.H. (2000) Travel Time versus Capacity Reliability of a Road Network. In M.G.H. Bell and C. Cassir (eds.), *Reliability of Transport Networks*. Research Studies Press Ltd., UK.
- Yang, H. and Meng, Q. (2000) Highway Pricing and Capacity Choice in a Road Network under a Build-Operate-Transfer Scheme. *Transportation Research*, Vol. 34A, No. 3, pp. 207-222.
- Yin, Y. and Ieda, H. (2001) Assessing Performance Reliability of Road Networks under Non-Recurrent Congestion. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1771, pp. 148–155.
- Yin, Y. and Ieda, H. (2002) Optimal Improvement Scheme for Network Reliability. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1783, pp. 1-6.
- Yin, Y., Lam, W.H.K. and Ieda, H. (2002) Modeling Risk-Taking Behavior in Queuing Networks with Advanced Traveler Information Systems. In M.A.P. Taylor (ed.), *Transportation and Traffic Theory in the 21st Century*, pp. 309-328. Oxford: Elsevier.
- Yin, Y., Lam, W.H.K. and Ieda, H. (2004) New Technology and the Modelling of Risk-Taking Behaviour in Congested Road Networks. *Transportation Research* Vol. 12C, pp. 171-192.

APPENDIX A: OUTLINE OF THE GENERALIZED REDUCED GRADIENT (GRG) ALGORITHM

The following are the detailed algorithmic steps of the Generalized Reduced Gradient Algorithm:

Step 0: **Initialisation**

Choose stopping tolerance $\varepsilon > 0$ and any starting feasible solution $\mathbf{x}^0$.

Partition vector $\mathbf{x}$ into basic variables $\mathbf{x}_b$ and non-basic variables $\mathbf{x}_n$.

Therefore, the initial point $\mathbf{x}^0$ can be represented by $\begin{bmatrix} \mathbf{x}_b^0 \\ \mathbf{x}_n^0 \end{bmatrix}$. Construct the

basis matrix $\mathbf{B} = \frac{\partial \mathbf{h}}{\partial \mathbf{x}_b}(\mathbf{x}^0)$ and the non-basis matrix $\mathbf{N} = \frac{\partial \mathbf{h}}{\partial \mathbf{x}_n}(\mathbf{x}^0)$. Set

$k = 0$.

Step 1: **Direction finding**

Compute the generalised reduced gradients $\mathbf{r}_n^k$ at $\mathbf{x}^k$ by:

$$\mathbf{r}_n^k = \frac{\partial CS}{\partial \mathbf{x}_n}(\mathbf{x}^k) - \frac{\partial CS}{\partial \mathbf{x}_b}(\mathbf{x}^k) \mathbf{B}^{-1} \mathbf{N}.$$

Determine the displacement direction $\mathbf{d}^k = \begin{bmatrix} \mathbf{d}_b^k \\ \mathbf{d}_n^k \end{bmatrix}$ by:

$$\mathbf{d}_b^k = -\mathbf{B}^{-1} \mathbf{N} \mathbf{d}_n^k,$$

$$\mathbf{d}_n^k = \begin{bmatrix} d_n^k \end{bmatrix},$$

where $d_n^k = \begin{cases} r_n^k & \text{if } r_n^k > 0 \text{ or } x_n^k > 0 \\ 0 & \text{otherwise} \end{cases}$, $r_n^k(x_n^k)$ is the element of $\mathbf{r}_n^k(\mathbf{x}^k)$.

Step 2: **Convergence test**

If $\|\mathbf{d}^k\| \leq \varepsilon$, stop.

Step 3: **Feasibility limit**

Compute feasibility limiting step $\lambda_{\max}^k$ by:

$$\lambda_{\max}^k = \begin{cases} \infty & \text{if } \mathbf{d} > \mathbf{0} \\ \min \left\{ \frac{x_j^k}{-d_j^k}, \forall j \in \{n, b\} \mid d_j^k < 0 \right\} & \text{otherwise} \end{cases}$$

in which d_b^k is the element of $\mathbf{d}_b^k$.

Step 4: Line search

Perform a one-dimensional search to determine the step size

λ^k solving:

$$\max CS(\mathbf{x}^k + \lambda^k \mathbf{d}^k)$$

subject to $0 \leq \lambda^k \leq \lambda_{\max}^k$.

Step 5: Move

Set $\mathbf{x}'^k = \mathbf{x}^k + \lambda^k \mathbf{d}^k$.

Step 6: Correction step:

Step 6.1 Set $l = 0$ and $\mathbf{z}^0 = \mathbf{x}_b'^k$.

Step 6.2 Compute $\mathbf{z}^{l+1}$ by:

$$\mathbf{z}^{l+1} = \mathbf{z}^l - \mathbf{B}^{-1} \mathbf{h}(\mathbf{z}^l, \mathbf{x}_n'^k).$$

Step 6.3 If $\mathbf{z}^{l+1} \geq \mathbf{0}$ and $\|\mathbf{h}(\mathbf{z}^{l+1}, \mathbf{x}_n'^k)\| \leq \varepsilon$, Set $\mathbf{x}_b^{k+1} = \mathbf{z}^{l+1}$, $\mathbf{x}_n^{k+1} = \mathbf{x}_n'^k$, and $k = k + 1$. Go to Step 1.

Step 6.4 If $\mathbf{z}^{l+1} \geq \mathbf{0}$, set $l = l + 1$ and go to Step 6.2.

Step 6.5 Substitute the basic variable that becomes negative by a non-basic variable. Form a new basis $\mathbf{B} = \frac{\partial \mathbf{h}}{\partial \mathbf{x}_b}(\mathbf{x}^k)$ and $\mathbf{N} = \frac{\partial \mathbf{h}}{\partial \mathbf{x}_n}(\mathbf{x}^k)$. Go to Step 1.

APPENDIX B: LIST OF AUTHOR'S PUBLICATIONS AND AWARDS

B.1 REFEREED JOURNAL PAPERS

1. Szeto, W.Y., **O'Brien, L.**, and O'Mahony, M. (2009) Measuring Network Reliability Considering Paradoxes: The Multiple Network Demon Approach, *Transportation Research Record: Journal of the Transportation Research Board*, 2090, 42-50.
2. **O'Brien, L.** and Szeto, W.Y. (2008) Designing Reliable Transport Networks: The Multiple Network Spoiler Approach, *HKIE Transactions*, 15, 1-9. (This is an awarded paper for the Hong Kong Institution of Engineers Outstanding Paper Award for Young Engineers/Researchers 2008).
3. **O'Brien, L.** and Szeto, W.Y. (2007) The Discrete Network Design Problem Over Time, *HKIE Transactions*, 14, 47-55. (This is a shortlisted paper for the Hong Kong Institution of Engineers Outstanding Paper Award for Young Engineers/Researchers 2007).
4. **O'Brien, L.**, Szeto W.Y., and O'Mahony, M. (2007) Time-Dependent Discrete Network Design Problem, *Journal of Eastern Asia Society for Transportation Studies*, 7, 2230-2245. (A conference version of this paper was awarded first runner-up prize of the Graduates & Students Division of the Institution of Civil Engineers, Republic of Ireland Branch Paper Competition 2007).
5. Szeto, W.Y., **O'Brien, L.**, and O'Mahony, M. (2006) Risk-averse Traffic Assignment with Elastic Demands: NCP Formulation and Solution Method for Assessing Performance Reliability, *Networks and Spatial Economics*, 6, 313-332.

B.2 REFEREED BOOK CHAPTER

1. Szeto, W.Y., **O'Brien, L.**, and O'Mahony, M. (2007) Generalisation of the Risk-averse Traffic Assignment, *Transportation and Traffic Theory* (eds. R. Allsop, M.G.H. Bell & B.G. Heydecker), 127-153, Elsevier, Amsterdam.

B.3 PAPERS IN REFEREED CONFERENCE PROCEEDINGS

1. Szeto, W.Y., **O'Brien, L.**, and O'Mahony, M. (2009) Measuring Network Reliability Considering Paradoxes: The Multiple Network Demon Approach., *Proceedings of the 88th Annual Meeting of the Transportation Research Board*, Washington D.C., January 2009.
2. **O'Brien, L.**, Szeto W.Y., and O'Mahony, M. (2007) Time-Dependent Discrete Network Design Problem, *Proceedings of the 7th Conference of the Eastern Asia Society for Transportation Studies*, Dalian, China, 24-27 September 2007.

3. **O'Brien, L.**, Szeto, W.Y., and O'Mahony, M. (2006) Transport Network Design under Risk-Averse Traffic Assignment: A Game Theory Approach. *Proceedings of the EWGT/EURO Conference Handling Uncertainty in Transportation: Analyses, New Paradigms, Applications and Comparisons*, Bari, Italy, 27-29 September 2006, 494-502.
4. **O'Brien, L.** (2007) Discrete Network Design: The Time-Dependent Approach, Institution of Civil Engineers Graduates and Students Paper Competition Seminar (Republic of Ireland Branch), Dublin, 6 February 2007.

B.4 JOURNAL PAPERS SUBMITTED FOR PUBLICATION

1. **O'Brien, L.** and O'Mahony, M. (2010) Time-Dependent Reliable Transport Network Design, *The Annals of Regional Science*, (submitted).

B.5 RESEARCH AND ACADEMIC AWARDS

1. Awarded 'The Hong Kong Institution of Engineers International Outstanding Paper Award for Young Engineers/Researchers (2008)'.
2. Short-listed for 'The Hong Kong Institution of Engineers International Outstanding Paper Award for Young Engineers/Researchers (2007)'.
3. Awarded the 1st Runner-up Prize of 'The Graduates & Students Division of the Institution of Civil Engineers, Republic of Ireland Branch Paper Competition (2007)'.
4. Trinity College Dublin Postgraduate Studentship (2005-2008).

