SWINBURNE UNIVERSITY OF TECHNOLOGY,

MELBOURNE, AUSTRALIA

DOCTORAL THESIS

The dynamic traffic assignment models with real-time information

---

Author:

Nam Hong HOANG

Supervisors:

Prof. Hai L. VU

Dr. Manoj PANDA

Dr. Antonio CRICENTI

A thesis submitted in fulfilment of the requirements

for the degree of Doctor of Philosophy

in the

Faculty of Science, Engineering and Technology

February 15, 2018

Declaration of Authorship

I, Nam Hong HOANG, declare that this thesis titled, "The dynamic traffic assignment models with real-time information" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at Swinburne University of Technology.
- This thesis has been submitted for the degree of *Doctor of Philosophy* at Swinburne University of Technology, and not to any other institution or for any other qualification.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed: 

Date: 15/02/2018

"If you cannot do great things, do small things in a great way."

Napoleon Hill (1883-1970)

Abstract

Nam Hong HOANG

The dynamic traffic assignment models with real-time information

Dynamic traffic assignment (DTA) is an important problem in the transportation planning to prepare for the development of a competitive economy and sustainable lifestyle. It has a wide range of applications in reality, from the evaluation of planning schemes, congestion management, to the forecast of road traffic in the network. Additionally, the development of communication technologies, the evolution of the automotive industry, and the movement of smart city bring a new behaviour of travellers in daily life, e.g., they make routing decisions based on the up-to-date traffic information collected on-line during their travel. Although DTA problem is extensively studied in many applications, only a few works analyse this new behaviour in the context of traffic assignment (or optimal routing). This thesis aims to fill this gap by developing a series of novel analytical frameworks to study the adaptive routing in the transportation network for a given evolution of updated information.

First, the thesis' contributions enhance the traditional DTA models (implicitly assuming full information about travel cost) via the development of a linear mathematical programming framework for both system optimal and user equilibrium route choices in a single origin-destination dynamic network. Relying on kinematic-wave theory, the proposed models are capable of capturing the spillover phenomenon, shock-wave propagation and especially predicting the congested queue at any link. The proposed DTA formulations improve the computational performance because of its linearity without the large enumeration of paths in both sets of variables and constraints. Last but not least, the analytical study shows the relation between two typical principles of route choice, i.e., user equilibrium (UE) and system optimal

(SO) principles, and the role of First-In-First-Out (FIFO) conditions in obtaining UE solutions.

These previous studies of the SO and UE DTA problems are extended to a more general problem where users only know the most recent update of network states. In this thesis, a proposed information model represents the evolution of probabilistic traffic demand and network capacity. The integration of this model in the traditional DTA framework forms linear formulations in the study of information-based stochastic system optimal (ISSO) solutions that internally consider users' preferences of route choices and the en-routing behaviour given an updated information. In the ISSO DTA model, users are assumed to comply with the route advices from system operators as long as they align with the users' preference.

For the selfish routing aided by real-time information, this thesis develops a novel linear mathematical framework to formulate the information-based stochastic user equilibrium (ISUE) DTA problem for a single origin-destination network. The linkage between UE and SO solutions enables the development of incremental loading method to obtain the ISUE solutions efficiently by solving a sequence of linear programs. Moreover, due to the use of link-based model, this method is more scalable than path-based methods in the literature by avoiding a large enumeration of paths especially in large-scale networks. The numerical examples show the impact of information to both route choices and network performance, and demonstrate the significant improvements in the obtained ISUE solution both in terms of accuracy and computational complexity.

Acknowledgements

I would like to acknowledge with particular gratitude the assistance of my supervisors, Prof. Hai L. Vu, Dr. Manoj Panda and Dr. Tony Cricenti. I am also indebted to a number of other people, in particular, Prof. Dong Ngoduy (University of Canterbury, New Zealand), Prof. Dave Watling (University of Leeds, UK), Prof. Hong K. Lo (Hong Kong University of Science and Technology), Prof. Mike Smith (University of York, UK) who collaborated with me on various research papers and provided much needed support. Thanks also to the CAIA group for friendship and supports during my PhD journey.

I am grateful to my current employer, Swinburne University of Technology, for providing the resources and support via the SUPRA scholarship to pursue a research higher degree.

Finally, I would like to thank my family for their loving forbearance during the long period it has taken me to conduct the research and write up this thesis.

Contents

List of Figures

List of Tables

List of Abbreviations

ATIS	**A**dvanced **T**raveller **I**nformation **S**ystems
BPR	**B**ureau of **P**ublic **R**oads
CTM	**C**ell **T**ransmission **M**odel
CUE	**C**apacitated **U**ser **E**quilibrium
DQM	**D**ouble **Q**ueue **M**odel
DSO	**D**ynamic **S**ystem **O**ptimal
DTA	**D**ynamic **T**raffic **A**ssignment
DUE	**D**ynamic **U**ser **E**quilibrium
DVI	**D**ifferential **V**ariational **I**nequality
FIFO	**F**irst **I**n **F**irst **O**ut
ISM	**I**ncremental **S**olution **M**ethod
ISSO	**I**nformation-based **S**tochastic **S**ystem **O**ptimal
ISUE	**I**nformation-based **S**tochastic **U**ser **E**quilibrium
KKT	**K**arush-**K**uhn-**T**ucker
KWM	**K**inematic-**W**ave **M**ethod
LP	**L**inear **P**rogramming
LTM	**L**ink **T**ransmission **M**odel
MCL	**M**inimum total **C**ongested **L**ength
MDCL	**M**inimum **D**ifferentiation of **C**ongested **L**ength
MILP	**M**ixed-**I**nteger **L**inear **P**rogramming
MIP	**M**ixed-**I**nteger **P**rogramming
MPC	**M**odel **P**redictive **C**ontrol
MSA	**M**ethod of **S**uccessive **A**verages
NCP	**N**on-linear **C**omplementarity **P**roblem
NHB	**N**on **H**olding **B**ack
NLP	**N**on **L**inear **P**rogramming
O-D	**O**rigin-**D**estination
PQM	**P**oint **Q**ueue **M**odel
SO	**S**ystem **O**ptimal
SQM	**S**patial **Q**ueue **M**odel
TTM	**T**wo-regime **T**ransmission **M**odel
UE	**U**ser **E**quilibrium
VI	**V**ariational **I**nequality
VMS	**V**ariable **M**essage **S**igns

List of Symbols

	General cell network
$\mathbb{T}$	Finite discrete time domain (starting at time 1 and ending at time T).
$\mathbb{C}$	Set of nodes.
$\mathbb{A}$	Set of directed arcs connecting nodes.
$\mathbb{P}$	Set of all possible paths in the network.
$\mathbb{P}^{(ij)}$	Set of all routes from node i to node j.
Γ_i^+	Set of downstream neighbours of node i, e.g. $\Gamma_i^+ = \{j : j \in \mathbb{C}, (i,j) \in \mathbb{A}\}$.
Γ_i^-	Set of upstream neighbours of node i, e.g. $\Gamma_i^- = \{j : j \in \mathbb{C}, (j,i) \in \mathbb{A}\}$.
D_t	Set of all demands entering the network at time t.
$D_t^{(ij)}$	Demand from node i to node j at time t.
$D_{t\|x}^{(ij)}$	Demand from node i to node j at time t for traffic scenario x.
S_t	Supply at time t, characterising the road capacity of any links.
	The information model
$\mathcal{X}$	Set of scenarios, describing all realisations of traffic demand and supply.
$\tilde{x}$	Random variable represents a traffic scenario $x \in \mathcal{X}$, $\Pr\{\tilde{x} = x\} = q_x$.
$x(t)$	The t-traced pattern of scenario x at time t.
$\tilde{x}(t)$	Random variable represents t-traced pattern.
$\mathcal{X}_{x(t)}$	t-realisable scenario set for a given t-traced pattern at time t, defined as a hyperscene.
$\tilde{\mathcal{X}}_{x(t)}$	Random set represents t-realisable scenario set at time t.
$\mathbb{X}_t$	The hyperscene set, describing all possible t-realisable scenario sets, at time t.
$\mathcal{Y}$	Set of all realisations of the traffic state in terms of link traffic flows in the network.
$\mathcal{K}_t$	The global real-time information at time t.
	The choice model
$H^{(ij)}$	A non-empty set of some paths from node i to node j, called a hyperpath from i to j.
$\mathbb{H}^{(ij)}$	Defines it as the power set of $H^{(ij)}$, all hyperpath from i to j.
$v_t^{(is)}$	The preferred paths (or hyperpath) from node i to destination s for each hyperscene at time t.
$L_{x(t)}^{(is)}$	Real-time choice set after applying user preference $v_t^{(is)}$ under the current real-time information $x(t)$ at time t.
μ	Policy choice function.
π	Path choice function.
	The cell transmission model with information
$\mathbb{C}_{\{R,S,I\}}$	Set of sources, sinks and intermediate cells.
$Q_{i,t\|x}$	Maximum flow in or out of cell i during time interval t in scenario x.
$N_{i,t\|x}$	Maximum number of vehicles that can be present in cell i at time t in scenario x.
$\omega_{i\|x}$	Backward propagation ratio at cell i in scenario x.

$n_{i,t\|x}$	Number of vehicles on cell i at time interval t in scenario x.
$n_{i,p,t\|x}$	Number of vehicles on cell i following path p at time t in scenario x.
$n_{i,t\|x}^{(-s)}$	Number of vehicles on cell i at time t, targeting the destination s, in scenario x.
$y_{ij,t\|x}$	Number of vehicles moving from cell i to cell j at time interval t in scenario x.
$y_{ij,p,t\|x}$	Number of vehicles moving from cell i to cell j, following path p at time t in scenario x.
$y_{ij,t\|x}^{(-s)}$	Number of vehicles moving from cell i to cell j at time t, targeting the destination s, in scenario x.
$f_{p,t\|x}$	Number of vehicles choosing path p for the demand at time t in scenario x.

For link-based networks

$\mathcal{G}$ $(\mathbb{V}, \mathbb{A})$	=	Set of nodes ($\mathbb{V}$) and directed arcs ($\mathbb{A}$).
$\mathbb{V}_{R,S,I}$		Set of sources, sinks and intermediate nodes respectively.
$\mathbb{A}_{R,S,I}$		Set of sources, sinks and intermediate links respectively.
Y_a^-		Set of inflow links to link a.
Y_a^+		Set of outflow links from link a.
$\mathbb{P}$		Set of all paths in $\mathcal{G}$.
$\mathbb{P}^{(rs)}$		Set of all paths from r to s in $\mathcal{G}$.
$\mathcal{T}$		Maximum continuous time horizon.
T		Maximum discrete time horizon.
$\delta = \frac{\mathcal{T}}{T}$		Unit time interval.
$\mathbb{T}$		Set of discrete times, i.e., $\mathbb{T} = \{1, 2, \ldots, T\}$.
r, s, a		Index for source link r, sink link s and any link a.
p		Index for a path $p \in \mathbb{P}$.
τ		The continuous time $\tau \in [0, \mathcal{T}]$.
t		The discrete interval time $t \in \mathbb{T}$.
$d^{(rs)}(\tau)$		Continuous-time traffic demand function from r to s at time τ.
$D_t^{(rs)}$		Amount of traffic demand from r to s at the discrete interval time t.
D		Total amount of traffic demands.
$q_a(\tau)$		Continuous-time flow capacity of link a at time τ.
$Q_{a,t}$		Maximum flow capacity to or out of link a at the discrete time t.
K_a		Vehicle jam density at link a.
V_a		Unsaturated speed in link a.
W_a		Backward speed in link a.
L_a		Length of link a.

Link-based variables

Φ	Set of all trips for a traveler moving from a source to a destination.
φ	A trip in Φ.
f_φ	Amount of traffic following trip φ.
Φ_p^t	Set of trips starting at time t on path p.
$\Phi_{a,t}$	Set of trips entering link a at time t.
$\boldsymbol{f}$	Set of all trip flows.
$f_{p,t}$	Amount of traffic at source r choosing path p at time t.
$f_{p,t,h}$	A partial flow of $f_{p,t}$ arriving the destination at time h ($h > t$).
$T_{p,t}$	The travel time averaged over all flows departing at time t on path p.
$\tilde{T}_{p,t}$	The marginal travel time for the path flow $f_{p,t}$.

$T_{p,t}(u)$	The travel time function averaged over an amount u of flow in $f_{p,t}$.
$\tilde{T}_{p,t}(u)$	The marginal travel time function for an amount u of flow in $f_{p,t}$.
$u_{a,t}$	Amount of traffic to link a at time t.
$v_{a,t}$	Amount of traffic out of link a at time t.
$f_{ab,t}$	Amount of traffic from a to b at time t.
$f_{a,p,t}$	Amount of traffic on path p entering link a at time t.
$f_{a,p,t,h}$	Amount of traffic on path p entering link a at time t and exiting link a at time h.
$\mathcal{D}_{a,t}$	Demand capacity at link a at time t.
$\mathcal{S}_{a,t}$	Supply capacity at link a at time t.
	For signal control
$\mathcal{C}$	The fixed signal cycle time.
$\mathbb{J}$	The intersection set and $\mathbb{J} \subset \mathbb{V}$.
Π_j	The set of signalling phases at the intersection $j \in \mathbb{J}$.
$\mathbb{T}_c$	The time domain corresponding to the c^{th} cycle.
$g_{\pi,c}$	Green split for phase $\pi \in \Pi$ at the c^{th} cycle.
G_0	The minimum green split for each phase at intersections.
$f_{\pi_n,t}$	Amount of signalling time assigned to phase π_n within time $t \in \mathbb{T}_c$.

To my family, Son & Dung

And my parent, Mr. & Mrs. Hong

Chapter 1

Introduction

1.1 Motivation

Improving the efficiency of transport could enable a more competitive economy and sustainable lifestyles (Mackie et al., 2014). However, there are a number of challenges for transport development, including the dynamic use of different transport modes (e.g., private cars, buses, trains), the management of congestion dynamically, driving safety, etc. The necessity of careful and optimised planning is required in any transport appraisals because the congestion cost is significant and increasing as shown in BITRE (2014).

To tackle the above challenges, the development of information services, such as Advanced Traveller Information Systems (ATIS), brings more valuable data for monitoring and operating traffic in real time. By utilising intelligent transport systems (ITS), a recent study in Australia (BITRE, 2017) has estimated a great improvement of the annual social cost associated with the road transport, i.e., avoidable $27 billion for road accidents and $30 billion for congestion cost predicted in the year 2030. Additionally, these advanced technologies can lead to a new travelling behaviour e.g., travellers potentially rely on a decision support system to make routing decisions based on the real-time updates of the network traffic conditions or traffic states. The insight understanding of this behaviour (e.g., its impact on reducing congestion) is the key to unlock the power of ITS, however, only limited work in the literature investigates such new behaviour in the context of optimal routing and traffic planning. As reviewed in Balakrishna et al. (2013), there are still a number of open issues spanning three aspects: information processing (e.g., big-data in the

large-scale networks), the modelling of drivers' responses to real-time information, and the optimisation of network performance.

There are different methods to tackle these challenges, herein we focus on the analytical approach based on the fundamental dynamic traffic assignment (DTA) problem which helps to find an optimal routing or traffic split for a given traffic demand in a transport network. It is worth noting that this DTA problem is closely related to other problems in transport, including network design, strategy control optimisation (signal timing, detour, speed limit etc.), and evacuation planning etc. In the literature, the extensive study of this problem has shown that it is capable of predicting the traffic evolution under different user behaviour models and that it incorporates a variety of criteria for cost-benefit analysis (CBA) of the optimal solutions (Szeto et al., 2012). Therefore, the wide range of DTA applications makes it suitable for addressing the above challenges in a transportation network underpinned by our development of informed routing decision models within the DTA framework presented in this thesis.

1.2 Thesis statement and challenges

The objective of this research is to analyse the impact of real-time information on traffic routing and planning under a DTA framework. In this section, we will initially present the thesis statement and then the major challenges in the study of this problem.

1.2.1 The problem statement

The following research questions help to identify the problem, its outcomes and methodology.

How to find optimal routing for a given real-time information? The thesis assumes the availability of real-time information that enables the behaviour of rerouteing traffic, for example, the update of incidents on any link, the real-time information of the peak-hour traffic demand or the current locations of traffic congestion in the

network. There are two typical classes of routing, either decided by the system operator to achieve system-wide performance or by individual travellers to minimise their own cost. To answer this question for each type of routing, we will first study an information model to represent the stochastic network states, particularly the evolution of probabilistic traffic demand and network capacity. We will then integrate this model with the two types of routing above in the optimisation framework that assigns optimal route flows on any links in the network.

How to analyse the impact of real-time information on the optimal routing and network performance? Proposing a closed-form mathematical formulation to compute the optimal routing aided by real-time information is the key to analyse its impact on both routing and network performance. We aim to apply the methodology of dynamic optimisation (Friesz, 2010) and analyse the optimal solutions via the well-known Karush-Kuhn-Tucker (KKT) conditions. Furthermore, numerical computation is another way to evaluate these impacts in different scenarios.

How to develop an efficient method for large-scale networks? Since the number of real-time DTA applications is growing rapidly, we need efficient methods to obtain optimal solutions, especially for large-scale networks, in a reasonable time. The literature presents work (Ziliaskopoulos, 2000; Ukkusuri et al., 2008) that provides efficient linear formulations for the system optimal DTA problems, unfortunately, there was no such formulation for user optimal route choice. To answer this question, we firstly utilise the advanced linearisable dynamic traffic models and then extend this to include the proposed information model. We apply the analytical methodologies that give insight understanding of optimal solutions to develop an efficient linear framework for the computation of both system and user optimal routing.

1.2.2 Major challenges

Given the scope of problem studied in this thesis, the major challenges are follows: the dynamic urban traffic network, the modelling of user behaviour, and the computational complexity. First, the literature has shown the difficulties in understanding

transportation networks (Ran et al., 2012), i.e., it took about 30 years (1950s-1980s) to establish essential theories for static models, including the fundamental relation between traffic flow and vehicle density, the dependence between travel time and flow (Bureau of Public Roads, 1964), Beckmann's formulation for the static traffic assignment (Beckmann et al., 1956) etc. From the 1980s-2000s, the transportation models focused on the dynamic networks with the development of kinematic-wave theory and the support of early intelligent transport systems. Different features of an urban network were investigated, including spill-back, hysteresis phenomena, backward kinematic-wave propagation, actuated signal controls, etc. After the 2000s, the rapid development of wireless communication technologies has changed not only the user behaviour (from users' travel experience to decision support systems) but also the action of system operators (from static controls to actively response according to the current network states). The interaction between operators and travellers in real-time yields a complex interaction in a dynamic urban traffic network, that still needs further research investigation.

Furthermore, the modelling of user behaviour is a crucial component in any study, therefore, many research activities have been dedicated to this problem in the past. The variety of these models was reviewed in Szeto et al. (2012), including the role of economical and psychological factors. Additionally, the involvement of real-time information initiates the dynamic assessment of travel cost made by users, that is not only more complex but also more difficult to be modelled accurately.

Finally, the general DTA problem is usually difficult due to its non-convexity and non-linearity (Carey, 1992). The development of efficient methods to solve this problem spreads across many applications, e.g., traffic operating systems, information provision system etc. Although computing infrastructure and technologies have advanced dramatically in recent years, more studies are still needed to efficiently utilise them for large-scale DTA problems. Particularly, an in-depth analytical study of this DTA in light of the available real-time information is necessary to discover useful properties or structure of optimal solutions and enable a practical deployment of such a system. In summary, the three mentioned challenges are the main obstacles in transport modelling and analysing of real-time information on traffic routing and planning under the DTA framework.

1.3 Contributions

This thesis develops a novel analytical framework to study adaptive routing, both system optimal (SO) and user equilibrium (UE) choices, in a transportation network for a given evolution of updated information. Herein the information is assumed to be complete and perfect (i.e., no erroneous data). The original contributions of this thesis are summarised as follows.

SO DTA (Chapter 3) The proposed SO DTA model, published in Ngoduy et al. (2016), formulates the kinematic-wave theory as a set of linear constraints by applying a recently proposed Two-regime Transmission Model (TTM) in Balijepalli et al. (2013). The TTM principle is more desirable for the optimal distribution of the queues over links than using others in literature because, on the one hand, it utilises the entry and exit flows to describe the link state similarly to the LTM (Yperman, 2007) or DQM (Osorio et al., 2011), and, on the other hand, it provides the time and space evolution of the queue lengths. It shows that this evaluation of queue lengths can be determined by the inflow at the upstream node and outflow at the downstream node in TTM. While LTM or DQM cannot describe such evolution of the queue lengths, CTM (Daganzo, 1994) can do it with high computational cost, especially for the SO-DTA problem. The major contribution in this research, therefore, is to establish a new optimisation framework to find SO solutions which can both minimise the total system travel time and optimally distribute the queue lengths over links so that potential spillbacks can be minimised or free-flow traffic inside the network can be attained.

This SO-DTA problem is suitable for the study of optimal traffic signal control. This chapter also proposes a continuous signal control model that takes into account two important signalling features, i.e., green time split and phase ordering. This model is integrated in the above SO-DTA framework for single-destination networks. Via numerical examples, it shows the impact of phase ordering on both route choices and green splits which cannot be seen by other continuous signal control models in the literature.

UE DTA (Chapter 4) The chapter proposes a new mathematical framework for the DTA problem with general capacitated constraints. The novelty of the proposed model stems from a cost function based on a new concept of trip separability as opposed to that of link costs used in the static models. This novel cost definition enables us to formulate a single and explicit UE objective rather than embedding the UE conditions in the constraints of the optimisation problem as usually done in the literature. It provides insights on the relationship between the SO and UE solutions based on the different perspectives of the resulting travel cost functions.

Furthermore, the mathematical proof presented in this chapter shows that in the limit, the SO objective could be used to obtain the UE solution as the system time step decreases to zero, given the satisfaction of the FIFO constraint. The difference between UE and SO objectives represents a measure to evaluate the accuracy of the proposed solution method.

As the result, we develop an incremental solution method to effectively solve the UE-DTA problem where the convergence is guaranteed with a predetermined number of iterations. The numerical results illustrate that UE-DTA problems can be effectively solved using the proposed linear programming framework, which is in stark contrast to the existing methods in the literature.

SO DTA with evolving information (Chapter 5) This chapter proposes a novel information model describing the evolution of knowledge (in the form of the real-time information) that accounts for the (stochastic) uncertainty in demand and network capacity. Our model departs from existing approaches in that it is based on the demand and supply rather than the link travel time distributions as in Gao et al. (2006). The proposed information and choice model has several advantages such as it enables us to study separately the impact of the key influencing factors on the route choice strategies, namely, demand, supply and user preference. Moreover, the preference of route choice in our model could also be a function of the information revealed by the operators (see Example 1 in Section 5.5.1).

Given the information, the drivers either decide on the route to their destination only at the beginning of their journey (referred to as pre-trip *path choice*) (Lindsey

et al., 2014), or gradually find the route toward the destination (i.e., en-route) by re-assessing their decision at intermediate points in a so-called *policy choice* (Gao, 2012; Dong et al., 2013). The new information model enables us to study the impact of information on route choices and network performance. This study is complementary to the large body of existing literature in that it provides the SO solution with real-time information access which can serve as a benchmark for network performance achievable with real-time information. Our approach provides a solution that is guaranteed to be SO under an evolving information model. In this SO framework, users comply with the route advice decided by the system operator as long as they align with the users' preferences.

We show that the proposed information model can be seamlessly integrated into a linear programming based optimisation framework. In particular, we develop two new linear programming formulations for the SO-DTA problem encompassing *path* and *policy* based adaptive route choice behaviours.

UE DTA with evolving information (Chapter 6) The thesis proposes a novel mathematical programming model for the information-based stochastic user equilibrium (ISUE) DTA problem under the evolution of knowledge, in the form of real-time information, that accounts for the (stochastic) uncertainty in demand and network capacity. In particular, travellers adapt their route choices at the source according to the latest information.

It shows mathematically the relationship between ISUE and the information-based stochastic system optimal (ISSO) objective and proves that the ISSO objective is approximately the ISUE objective as the time step goes to zero, given the satisfaction of the FIFO constraint.

Additionally, it proves the impact of information in terms of total system travel time in the increasing order of following solutions: ISSO, without-information SO and without-information UE solutions. For the ISUE solution, we discuss its performance in both analytical study and numerical results that it could be as good as an ISSO, or worse than a without-information UE solution.

Furthermore, it develops a link-based incremental solution method to effectively

solve the ISUE-DTA problem where the convergence is guaranteed with a predetermined number of iterations. Via numerical examples, this chapter shows that the ISUE-DTA problem can be efficiently solved using the proposed linear programming framework, which is very different from the existing methods in the literature.

1.4 Organisation of the thesis

The remaining thesis has six chapters which are organised as follows:

Chapter 2 presents the background knowledge of the general DTA framework which has two main components, i.e., traffic model and route choice model. It then provides a review of DTA models with two types of travel choice, i.e., system optimal and user equilibrium, with or without real-time information.

The thesis' contributions are presented in Chapters 3–6. In Chapter 3, we first make an overview of the two-regime transmission model (TTM), proposed by Balijepalli et al. (2013). We then develop optimisation framework to optimise queue placement in the dynamic network based on the linear-constraint formulation of the TTM-based SO-DTA model.

Chapter 4 proposes a novel mathematical framework for studying DTA problems in a dynamic capacitated network that applies to both the SO and UE formulations in a unified manner. Based on the analytical study, we develop a novel linear programming (LP) based method to solve the UE-DTA problem efficiently by exploiting the relationship between the SO and UE objectives.

Chapter 5 first proposes an information model to describe the evolving probabilistic traffic demand and network capacity. It is associated with a choice framework that utilises the real-time update of information to make routing decisions. In this framework, the system operator considers users' preference of route choice in its routing decision, and therefore, users comply with the route advice from the system operator. This chapter proposes two SO DTA models with information, each for policy choice and path choice. It also discusses the complexity of these models and makes comparison between solutions of policy and path choices.

Similarly, Chapter 6 utilises the information model proposed in Chapter 5, but

for the computation of user equilibrium path choice. It proposes an information-based stochastic user equilibrium (ISUE) DTA model, taking into account the evolving real-time information. It then develops an incremental solution method for the proposed model and the linear approximation of ISUE objective based on the study of relationship between ISUE and information-based stochastic SO (ISSO) solutions.

Finally, Chapter 7 concludes the thesis by providing a summary of all contributions (Chapters 3–6), and a discussion of the potential extensions of this work.

1.5 Publications

All contributions in this thesis have been published or are under-review. The list of publications is as follows.

1. Dong Ngoduy, N. H. Hoang, Hai Le Vu, and D Watling (2016). "Optimal queue placement in dynamic system optimum solutions for single origin-destination traffic networks". In: *Transportation Research Part B: Methodological*. DOI: 10.1016/j.trb.2015.11.011.
   In this work, I contributed significantly in Sections 4, 5 and 6 of this publication which are included in Chapter 3 of this thesis.

2. N. H. Hoang, H. L. Vu, M. Smith, and D. Ngoduy (2016). "Convex signal control model in a single-destination dynamic traffic assignment". In: *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, pp. 790–794. DOI: 10.1109/ITSC.2016.7795645.

3. N. H. Hoang, Hai Le Vu, Panda Manoj, and K. Hong Lo (2017a). "A linear framework for dynamic user equilibrium traffic assignment in a single origin-destination capacitated network". In: *Transportation Research Part B: Methodological*. DOI: 10.1016/j.trb.2017.11.013.

4. N. H. Hoang, Hai Le Vu, Panda Manoj, and Dong Ngoduy (2017b). "An information model for real-time adaptive routing in the dynamic system optimal assignment problem". *Under-review to Networks and Spatial Economics*.

5. N. H. Hoang and Hai Le Vu (2017c). "An informed user equilibrium dynamic traffic assignment problem in a single origin-destination stochastic network". *Under-review to Transportation Research Part B: Methodological*.

Chapter 2

Literature review

Dynamic traffic assignment (DTA) is the last step in the four-stage process of transportation planning (Willumsen, 2011), in which each individual trip between an origin and a destination is allocated to a route in the corresponding network. The literature reviews in Peeta et al. (2001), Szeto et al. (2012), and Ran et al. (2012) showed a considerable amount of works in the study of DTA problem especially since the 2000s, with a wide range of applications in reality, e.g., congestion management, a forecast of future road states, evaluation of planning schemes, etc. This part summarises the modelling and solution methods to the DTA problem in the literature.

2.1 A general framework of DTA problem

The general description of DTA problem is that for a given transport network (supply side) and a set of travel demands (demand side), find the optimal routes for each demand. The supply side provides a cost model to evaluate the travel cost between any origin-destination (O-D) pairs in the network while the demand side makes routing decisions according to this travel cost. From the economical point of view, DTA represents the interaction between demand (the transportation network) and supply (the amount of traffic seeking for routes) via the dependence between travel cost and route choice. This section reviews different approaches to study the supply and demand sides in the literature.

2.1.1 The supply-side dynamic network loading (DNL) models

In DTA, the typical travel cost is the residual time a vehicle staying in a network, however, there are other definitions of cost depended on the objective of route choices, e.g., in terms of fuel consumption (Ma et al., 2017), battery usage (He et al., 2013) or monetary cost (Yang et al., 2004b). For more accurate estimation of travel cost, the cost models also capture the properties of traffic propagation, including First-In-First-Out (FIFO) or spill-over phenomenon. In the literature, there are three different approaches, i.e., microscopic, macroscopic and mesoscopic, to model the dynamic propagation of traffic over the network.

First, the microscopic approach describes each vehicle individually, i.e., its length, speed and location etc., and constrains vehicles' behaviour to obey traffic signs, its reaction to the ahead traffic (car-following model), lane changing, etc. Due to the detail of traffic at the level of vehicle, this approach enables the studies in a diversity of user behaviour (car following or lane changing) and vehicle types (e.g., trucks, cars or buses). As a result of this complexity, this method is computationally expensive and difficult for calibration (due to a large number of attributes). For this approach, we refer to the following works in Brackstone et al. (1999), Punzo et al. (2005), and Rahman et al. (2013).

In contrast, the macroscopic approaches regulate the aggregated traffic flow to follow a fundamental relationship, e.g., between flow and density in Fig. 2.1. Therefore, this method uses less number of variables to characterise the network state. Major analytical works in DTA rely on this macroscopic approach to efficiently analyse the optimal solutions and provide the insight understanding that is unable (or difficult) to observe using the other approaches. However, as a drawback, it is unable to describe the route for each vehicle, and difficult to deal with a heterogeneous traffic environment.

The hybrid approaches (or mesoscopic models) combine the advantages of the above two approaches by considering route choice individually but aggregating traffic propagation on links (meaning that vehicles do not have individual trajectory within link but follow a macroscopic traffic model). CONTRAM (Taylor, 2003) grouped vehicles into packets where their speed is decided by the density for a given

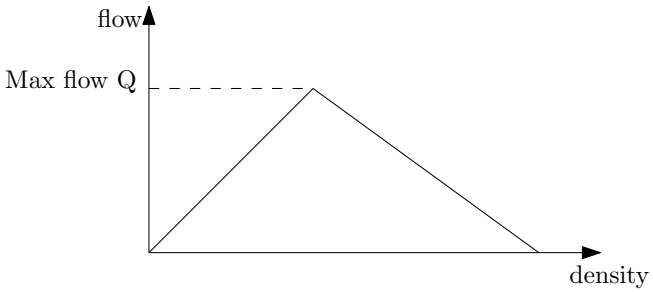


FIGURE 2.1: The triangular fundamental diagram between flow (veh/s) and density (veh/m).

speed-density relation function. DYNASMART (Jayakrishnan et al., 1994), FASTLANE (Gawron, 1999), DTASQ (Mahut et al., 2003) proposed queue-server models where a road is divided into two parts, i.e., running (from upstream) and queueing (at downstream). Nagel et al. (1992) applied the method of cellular automata that vehicles follow a minimal set of behaviour rules to determine their traverse over cells in the network. The disadvantage of mesoscopic models is the integration of micro- and macroscopic models which requires a complex interaction between the study of route choice (at the micro level) and traffic propagation (at the macro level). We refer to Joueiai et al. (2015), Nuzzolo et al. (2016), and Toledo et al. (2010) for further detail discussions of this method.

This thesis focuses on the analytical study of DTA problem, therefore, applies the macroscopic traffic network loading models in the proposed DTA models. The following part summaries their state-of-the-art formulation in the literature.

The macroscopic approach The general behaviour of traffic propagation on a road is similar to queueing operations, i.e., traffic, entering a road at its upstream node, is stored and dispatched at the downstream node. Therefore, a number of queue-based models are proposed, including the point queue (PQ) (Merchant et al., 1978; Vickrey, 1969), spatial queue (SQ) (Zhang et al., 2013), double queue (DQ) (Osorio et al., 2011) or kinematic-based queue (KQ or LWR as the abbreviation of authors' names) (Lighthill et al., 1955; Richards, 1956).

The mathematical differences among PQ, SQ and KQ were studied in Zhang et al. (2013) and Shen et al. (2014). The PQ models consider a queue with infinite storage, that enable us to treat roads or links independently. However, they

are unable to represent the effect of backward flow spillover in which the reduction of downstream flow is propagated back to the upstream flow. To enable this spillover phenomenon, SQ models respect to the physical space of each vehicle and link, i.e., equivalently queue has finite storage. However, as they do not base on the kinematic-wave theory, the SQ models still not able to describe the formation of shock-wave traffic which normally prevents free-flow movement, therefore vehicles take a longer time to be discharged at the downstream point. The DQ models construct the one-way road as two queues: the upstream queue (also called running queue) and downstream queue (or bottleneck queue). The challenge in this approach is that they allow random variable of queue lengths, but restricted to a fixed total length. Via simulation, this model coincides with the simplified theory of kinematic waves proposed previously by Newell (1993).

Based on the kinematic-wave theory, the cell transmission model (CTM), proposed by Daganzo (1994), is the well-known LWR-based traffic model with the capability of capturing shock-waves and spillover phenomenon. The relationship between vehicle flow and density follows the triangular fundamental diagram in Fig. 2.1. The properties of CTM were well-studied in Daganzo (1994), Daganzo (1995), and Mohan et al. (2013) not only its advantages but also the lack of describing anticipation and user-driven inertial effect. For the comparison with previous models, Nie et al. (2005) showed that CTM gave better insight traffic on links, but paid for a large computational cost. Similarly, the link transmission model (LTM), first presented in Yperman (2007), based on the simplified kinematic-wave theory (Newell, 1993). However, it is more efficient and scalable than CTM because the model is discrete only in time instead of in both time and space, shown in Balijepalli et al. (2013), Jin (2015), and Gentile (2010).

For causal characteristics (car-following or hysteresis effects), they are often modelled in the relationship between speed and density in the higher-order LWR-based models which are reviewed in Mohan et al. (2013). Due to their burden of complexity, these models have not been used widely in the DTA formulation.

2.1.2 The demand-side choice models

The demand model, also called the travel choice model, represents a selection of routes according to the travel cost that is provided, observed or perceived from the network loading models. The pioneer work of Wardrop (1952) introduced two fundamental principles of travel choices which are widely used in many DTA models, described below:

System optimal (SO) principle The total cost spent by all users is minimised.

User-equilibrium (UE) principle No users could unilaterally change their choices to get a better cost.

In SO principle, drivers behave cooperatively to ensure the minimisation of total travel costs. It is suitable for the studies of traffic planning in which different control strategies such as signal timing, lane allocations, and road pricing are employed to influence traffic so that the total cost is minimised. This principle is also useful for any centrally controlled systems, for example to depict SO flows in logistics (Amaral et al., 2015), public transports (Ibarra-Rojas et al., 2015), or in evacuation systems (Ng et al., 2010).

Differently, UE principle represents the selfish behaviour of drivers that rationally select the lowest-cost routes. It is widely used in many DTA models to represent a reasonable user behaviour in practice. Given this principle, the decision of UE choice requires two components: the choice dimension and the definition of travel cost. In the context of a dynamic network, the choice dimension could be pure route choices (originally studied in static models), pure departure time choices, or simultaneous route and departure time choices. For travel cost, it is depended on the domain of applications and is derived from the supply-side models introduced in Section 2.1.1.

According to the review of travel choice models in Szeto et al. (2012), these principles are extended in different ways: definitions of travel cost (e.g., travel time or budget), psychological factors (e.g., perception error, risk aversion, or inertial behaviour), uncertain travel time caused by uncertainties of demand or network capacity, bounded rationality (in which travellers only changes the current routes to another if the difference of costs exceeds a threshold), or elastic demand (in which

users have right to cancel their trips). In stochastic models, some papers (Sun et al., 2017; Szeto et al., 2011b) also considered the stability or reliability of stochastic solutions via evaluating both mean and deviation of some random quantities (e.g., stochastic travel time). In multiple modal traffic which includes different types of vehicles, e.g., buses, cars or trucks, the interaction caused by a mixture of route choices was studied in Bliemer et al. (2003) and Lo et al. (1996).

Due to the variety of travel choice models, there are different methods to formulate them. The SO principle can also be formulated either as an objective function of minimal total cost (Ziliaskopoulos, 2000) or as a complementarity condition using marginal route travel time (Wie et al., 1998). It is worth to note that UE concept is seen equivalently in other theories, i.e., economic concept of utility maximisation, game theory. Therefore, UE choices could be formulated in various methods, including the Nonlinear Complementarity Problem (NCP) (Wie et al., 2002), Variational Inequality Problem (VIP) (Lo et al., 2002; Friesz et al., 1993), Mathematical Programming Problem (MPP) (Wie et al., 1990), Fixed-Point Problem (FPP) (Huang et al., 2002), Optimal Control Problem (OCP) (Lam et al., 1995; Friesz et al., 1989), or Continuum Modeling Problem (CMP) (Wang et al., 2016; Jiang et al., 2011).

The following sections review the DTA models in the combination of route choice and network loading models. In this thesis, they are focused on:

- Macroscopic network loading model: kinematic-wave method.
- Route choice model: inelastic demand, homogeneous vehicles, the decision of route choices relied on a decision support system (DSS) that is not influenced by human factors, e.g., perception error, inertial behaviour etc.

2.2 System optimal DTA model

The SO-DTA problem predicts the optimal time-dependent routing pattern of travellers in a network such that the given time-dependent origin-destination demands are satisfied and the total system travel time spent by all travellers is minimised, assuming some models for dynamic network loading. There has been much effort in literature undertaken to formulate the kinematic wave model (KWM) of Lighthill

et al. (1955) and Richards (1956), based on an analogy between traffic flow and certain types of wave motion in fluids, for the SO-DTA problem. Especially, a discrete version of the KWM, called the cell transmission model (CTM) (Daganzo, 1994; Daganzo, 1995) has been formulated as side constraints in the SO-DTA problem (Carey et al., 2012a; Carey et al., 2012b; Gentile et al., 2005; Gentile et al., 2007; Lo et al., 2002; Nie et al., 2005; Shen et al., 2014; Szeto et al., 2011a; Ukkusuri et al., 2012; Ukkusuri et al., 2008; Ziliaskopoulos, 2000). It is because knowing the time-space dynamics of traffic flow within the link will facilitate a better understanding of the resulting SO solution, for example, the level of congestion and spillback location in the network, etc. An important property of the CTM is the possibility to reformulate it as a relaxed set of linear constraints so that a linear programming model for the SO-DTA problem for a network can be solved efficiently (Beard et al., 2006; Li et al., 2003; Ukkusuri et al., 2008; Ziliaskopoulos, 2000).

However, the choice to adopt CTM is not without its computational overheads. As noted in Nie et al. (2005), if we just consider the issue of DNL (of given route in-flow profiles) then the computational time for the CTM is directly proportional to the number of cells, and hence the computational efficiency is inversely proportional to the accuracy (in recovering the LWR). Thus the choice of discretisation level effectively means a choice/compromise between computational efficiency and the level of agreement with the (continuum) KWM model. If we then consider the wider issue of how the model is integrated within a SO-DTA framework, then we are faced with further computational issues: if the CTM is specified as a set of side constraints as suggested by Peeta et al. (2001) then the number of constraints grows with the fineness of the discretisation, whereas if we represent it using a route-based mapping as in Lo et al. (2002), then we are led down a path of route enumeration with all the computational difficulties that it is known to bring. As noted in Bar-Gera (2005), the 'computational requirements reduce the attractiveness of this model for large-scale long duration applications'. Such computational issues of CTM based SO-DTA problems are further explored in Shen et al. (2008) for different network sizes, where the number of constraints is increased polynomially with the network size.

Recently, a well known Lax-Hopf (LH) formula for the Hamilton-Jacobi (HJ) type partial differential equations has been used to solve the KWM, which can avoid the

discretisation, hence enhance the computational efficiency. The LH formula can be used to provide a variational formulation of the HJ equation solutions describing the evolution of a cumulative number of vehicles at the two ends of the link. In principle, the LH formula has been derived in various ways including the traffic flow theory by Daganzo (2005) which actually generalises the theory of Newell (1993), the viability theory by Aubin et al. (2011) for given boundary conditions at two ends of the link, which is then extended to include the internal condition (i.e. information of probe vehicles) by Claudel et al. (2010a) and Claudel et al. (2010b), and the technique of calculus of variations by Evans et al. (2012). The Link Transmission Model (LTM) in either discrete form (Yperman et al., 2006) or continuous form (Han et al., 2016; Jin, 2015) has been developed using the Newell's theory which is a special formulation of LH formula above, where the state of the whole link (i.e., either free-flow or congested) will be determined by the entry and exit flow. More specifically, the flow propagation in LTM is based on Newell's cumulative flow curves applied at the entry/exit of each link, with node models used to calculate the transition flows, which are based on conservation of flow between the incoming and outgoing flows. Sending/receiving flows, together with transition flows and other flow constraints, form the basis for updating the cumulative flows at the link boundaries.

Osorio et al. (2011) and Osorio et al. (2014) have developed another version of LTM, which is a so-called Double Queue (DQ) Model. In the DQ model, the link is treated as a set of two queues, referred to as the upstream queue and the downstream queue. Both LTM and DQ model can properly capture the free-flow travel time delay (when the link is in a free-flow state) and the backward shockwave time delay (when the link is in a congested state), which makes them possible to capture queue spillbacks. The DQ model was used in Ma et al. (2014) to find a free-flow DSO solution where spillback is tracked by the traffic state at the link entrance (in free-flowing) or at the link exit (in congested) accounting for some time shift. Nevertheless, either LTM or DQ model does not determine explicitly the propagation of the front shocks within a link and thus is unsuitable for providing the detailed traffic state within the link.

Signal control model in DTA The existing interaction between signal controls and traffic flow has been found in the literature. Allsop (1974) first pointed out that network control parameters (such as traffic signals) could influence user route choices. The benefits of adaptive route choice have been studied in Yang et al. (2004a) and Sims et al. (1980) where a choice was made according to a known signal timings and online traffic states, respectively or in Wallace et al. (1991) with coordinate controls among relevant intersections.

As the traffic network modelling advances, the concept of signal control has been extended to the DTA problem, first introduced by Gartner et al. (1998), through a framework that includes both signal control and traffic assignment models. Han et al. (2015a), Sun et al. (2006), and Ukkusuri et al. (2013) proposed bi-level approaches where the optimal controls subject to the optimal route choices for a given control in the lower-level sub-problem. Similarly, Li et al. (2015) studied the combination of route guidance and optimal signal control models. Although these approaches are sufficient to describe the problem in general, all of them face the issue of computational complexity. In fact, it is not necessary to always form a bi-level model, especially when travellers follow suggestions from the network operator as long as it aligns with their preference of route choices (Jahn et al., 2005; Paz et al., 2009b; Le et al., 2017; Le et al., 2013). Therefore, this thesis will base on a single-level approach and study signal control models embedded inside the DTA framework with an open option for describing user constraints as in Jahn et al. (2005).

In particular, this thesis focuses on the methodology of modelling traffic signal inside the DTA framework. Not long after the appearance of cell transmission model (CTM) (Daganzo, 1994), Lo (1999) and then Lin et al. (2004) developed CTM-based mixed-integer programming (MIP) models to include signal controls. Related to the DTA problem, several works relied on Model Predictive Control (MPC) framework to develop mixed-integer linear programming (MILP) models (Kamal et al., 2015; Haddad et al., 2013) which were more efficient than MIP in general. Although MIP or MILP approaches are straightforward to describe the phase transitions (red to green and vice versa), they are difficult to be solved. Therefore, a number of algorithms were proposed for these problems, e.g., decomposing methods (Rinaldi et al.,

2015; Timotheou et al., 2015), genetic approach (Sun et al., 2006) or Iterative Assignment Optimization procedure (Ukkusuri et al., 2013).

Besides the above methods for the exact description of traffic signalling characteristics, other works approximated their effects to obtain a continuous model of signal control. For example, focusing only on the green split for each phase in a predefined cycle time, Chen et al. (1998) and Ukkusuri et al. (2010) proposed efficient linear models instead of MIP models. Later, Han et al. (2014) studied this signal control continuum approximation in the link transmission model (LTM) (Yperman, 2007). Note that, the property of continuity is essential in many works, especially for the user equilibrium traffic assignment problems. However, by approximation, studying the effect of phase ordering to routing is infeasible in these works.

In conclusion, Table 2.1 summarises LWR-based SO-DTA models in the literature, and compares them with the SO-DTA model proposed in Chapter 3. The following section reviews the LWR-based models for the UE-DTA problem.

2.3 User equilibrium DTA model

Dynamic user equilibrium traffic assignment (UE-DTA) formulation has been used in many applications, such as transport planning or network design, and studied intensively in the literature. However, finding the user equilibrium solutions remains a big challenge due to the complexity of current non-linear models that are not scalable for large networks.

In the literature, most of the existing works enforce the UE principle (also known as Wardrop's first principle), where an individual user does not have an incentive to unilaterally change its decision about route choice, as a constraint in their mathematical models. These works include variational inequality (VI) models (Lo et al., 2002; Han et al., 2015b; Zhong et al., 2011; Perakis et al., 2006; Gentile, 2016), differential variational inequality (DVI) models (Friesz et al., 2011), or non-linear complementarity problem (NCP) (Ukkusuri et al., 2012). Due to the difficulty of solving these models, there have been a number of approximated solution methods proposed, e.g., fixed-point algorithms, the method of successive averages (MSA), alternating direction method, projection method, etc. The performance of these iterative methods,

TABLE 2.1: Comparison among selected LWR-based SO-DTA models in the literature. Abbreviations: **Y**: Yes, **N**: No, **C**: Continuous, **D**: Discrete, **MI**: Mixed-Integer.

Papers	**Traffic model**	**Holding-free**	**Congested queue length**	**Signal control formulation**	**Phase ordering**	**Time setting**	**Space setting**
Ziliaskopoulos (2000) and Shen et al. (2014)	CTM	Y	N	N	N	D	D
Ukkusuri et al. (2008)	CTM	N	N	N	N	D	D
Ma et al. (2014)	DQM	Y	N	N	N	D + C	C
Beard et al. (2006) and Ukkusuri et al. (2013)	CTM	N	N	MI	N	D	D
Lo (1999), Lin et al. (2004), and Timotheou et al. (2015)	CTM	Y	N	MI	N	D	D
Han et al. (2015a)	LWR	Y	N	C	N	C	C
Ukkusuri et al. (2010)	CTM	N	N	C	N	D	D
This thesis (Chapter 3)	TTM	Y	Y	C	Y	D	C

e.g., their speed of convergence, partially depends on the chosen initial solutions and other parameter settings. A performance comparison among existing methods can be found in Carey et al. (2012a).

In contrast to the above, the works of Beckmann et al. (1956) and Larsson et al. (1999) and Correa et al. (2004) which apply to static networks with infinite and finite capacity links, respectively, incorporate the UE principle through suitably defined objective functions and not as constraints. The framework developed in this thesis provides an extension of Beckmann's formulation to dynamic networks. So far there has not been a successful attempt for such a dynamic network problem as Lin et al. (2000) showed that an incorrect extension of Beckmann's formulation would not guarantee the UE condition in general. It is because the above mentioned static UE model relies on the separability of link cost functions without any interaction among them, which is not valid in the dynamic scenarios (i.e., DTA problem) due to the complexity of traffic propagation over both time and space.

In the proposed framework, like Beckman's model but for dynamic networks, we show that an appropriately chosen objective function will drive the solution to satisfy either the first Wardrop's principle (i.e., achieving UE solution) or the second Wardrop's principle (i.e., achieving SO solution). However, rather than using link costs, this thesis will introduce a concept of trips that captures the vehicle trajectory in both time and space, and thus enables the separation of path travel costs.

Furthermore, several works exploit the relationship between the UE and SO solution to solve the UE-DTA problem. Carey (2009) utilised the SO problem to find a UE solution, however, via a bi-level framework that separated the route flow assignment (in the upper-level SO model) and link flow propagation (in the lower-level model). In particular, the equilibrium flow was solved by iteratively updating the link travel time. In contrast to Carey's approach, this thesis will formulate the UE-DTA problem in a single-level optimization model with an explicit objective function. Moreover, its solving approach is incremental without the need of updating the assigned flows. As a result, the non-linear First-In-First-Out (FIFO) constraints can be transformed to linear formulas yielding a linear model in each and every solution step in our incremental method.

Note that, the incremental method was first developed by Van Vliet (1976) but

TABLE 2.2: Summary of selected LWR-based UE-DTA models.
Abbreviations: **ISM**: Incremental Solution Method, **N**: No, **Y**: Yes.

Paper(s)	**Traffic model**	**UE formulation**	**Quantity of SO-UE difference**	**Solution method**
Lo et al. (2002)	Path-based CTM	VI	N	Alternating direction
Han et al. (2015b)	Path-based LWR	VI	N	Fixed point, self-adaptive projection, proximal point
Perakis et al. (2006)	Path-based LWR	VI	N	Dynamic Frank-Wolfe
Gentile (2016)	General LTM	VI	N	Gradient projection
Ukkusuri et al. (2012)	Path-based CTM	NCP	N	Projection method
This thesis (Chapter 4)	Path-based LTM	Non-linear objective function	Y	Linear link-based ISM

without any theoretical analysis of the SO-UE relationship. It simply solved the UE problem by loading some fraction of demands, choosing the shortest paths then updating the travel cost before repeating these steps for the remaining demands. Herein, the work in Chapter 4 will utilize the incremental approach and develop a unified mathematical formulation to theoretically obtain the UE solution by leveraging the relationship between the SO and UE objectives while preserving the FIFO principle.

In summary, Table 2.2 summarises the selected LWR-based UE-DTA models and compares with the UE-DTA model proposed in Chapter 4. The following section reviews the DTA models with information-based adaptive routing.

2.4 DTA models with adaptive routing based on real-time information

With the advancement of technologies, information becomes more accessible with improved quality and diversity (e.g. via sensors, phones and smart devices, or social networks, etc.). Increasingly, information is being used to enhance travellers' experience and to improve traffic flows in the transportation networks (Ahmed et al., 2016). For example, in advanced traveller information systems (ATIS) (Chorus et al., 2006), information such as traffic condition can now be provided for road users

to make informed routing decision in real-time. Besides, more and more information is being used by the road network operators to reduce congestion and better manage traffic flows in the network. For this reason, the value of information and its impact on users' behaviour and network performance is a very important topic and has been researched intensively in the literature. Most of the existing works (with the exception of Waller et al. (2013)) focus on the user's point of view where questions regarding the improvement in travel time of the individual, or the equilibria of that of many individuals (i.e., network perspective) have been studied. In the following, this section reviews the relevant literature on the impact of information on users' behaviour and motivates the development of a new information model to be integrated into the well-researched DTA framework.

2.4.1 Information impact on routing

The benefit of information for road users (i.e., from the individual point of view) was first studied in Hall (1983) and subsequently in Yang (1998), Arnott et al. (1999), Gao (2012), and Balakrishna et al. (2013). It is found that the total travel time of users who adaptively select route using information is reduced, but the benefit decreases as the market penetration increases. This relation suggests the trade-off between the amount of information measured by the degree of user participation and its benefit to the travellers. Furthermore, better information coverage both in terms of space and time seems to have a positive impact on users' ability to adapt in real-time (Gao et al., 2012).

Given the information, users' adaptation (i.e. behaviour) can also come in different forms such as varying the departure times, transport modes or selecting the optimal route (Balakrishna et al., 2013). The latter has been studied intensively in the literature where drivers, based on the information available to them, either decide on the route to their destination only at the beginning of their journey (referred to as pre-trip *path choice*) (Lindsey et al., 2014), or gradually find the route toward the destination by reassessing their decision at intermediate points in a so-called *policy choice* (Gao, 2012; Dong et al., 2013).

In path choice, due to the lack of perfect, real-time information, travellers might perform an imprecise evaluation of travel cost (Han, 2003; Connors et al., 2009).

Further study of the personal perception was conducted with an additional penalty of late arrival (Watling, 2006), or with the representation of inertial phenomenon (Xie et al., 2014) in route choice behaviour. Besides that, the effect of information coverage on path choice is investigated in the study of inefficiency (also referred as the price of anarchy) between user and system optimal solution (Liu et al., 2009), or the benefit of pre-trip information to routing decision was discussed in Lindsey et al. (2014).

In policy choice, Polychronopoulos et al. (1996), then followed by those in Gao et al. (2006), Gao et al. (2012), and Gao (2012), introduced the concept of learning where users could eliminate infeasible realisations of link cost via the knowledge gained as they traversed the network. With stochastic and flow dependent arc cost, Unnikrishnan et al. (2009) assumed that users could know the link travel costs only when they arrive at the upstream node of these links, which in turns enable them to choose the best next link before continuing their journey. Other studies by Kim et al. (2005), Xiao et al. (2014), and Dong et al. (2013) exploited the knowledge of road state transition (or link costs correlation) to reduce the uncertainty of travel time estimation in the policy routing. Instead of path or policy choice, some studies, e.g., Spiess et al. (1989), Marcotte et al. (1998), Hamdouch et al. (2004), and Ukkusuri et al. (2007), considered hyperpath (or multi-path) choice, also referred to as strategy choice, in which a group of paths between source and destination is predefined, but travellers make a routing decision depending on the current traffic information.

2.4.2 The adaptive SO routing with information

Although important, there are surprisingly not many works existing in the literature that examine the benefit of information usage from the network's perspective. According to the emerging trends of the role of information technology in transportation, e.g., autonomous vehicles (Diakaki et al., 2015), Internet of Things (Wang et al., 2015), etc., it is expected that network operators will have more control and options to influence users' behaviour to deal with global issues (e.g., serious disasters, massive congestion, etc.). To the best of our knowledge, research by Waller et al. (2013) is the only work so far that studied the system optimality (SO) of the DTA problem with the strategic routing (i.e., policy choice with a predefined group of paths). Their

strategic routing turns out to be a kind of adaptive path choice based on historical (a priori) information. Consequently, to-date, the use of real-time information and the resulting users' adaptive route choice behaviour on network performance have been little known and is still an open problem in the literature.

Instead of studying the benefit of information use to individual user route choice, Waller et al. (2013) modelled the system optimal (SO) path choice based on the historical (a priori) information from the network's perspective. In terms of route guidance, the works of Paz et al. (2009a), Balakrishna et al. (2005), Peeta et al. (2002), and Mahmassani (2001) steer the network towards SO solutions via information provision. In their work, they approached this problem by developing the rolling-horizon framework for updating real-time information and using fuzzy-logic approach for route choice.

2.4.3 The adaptive UE routing with information

Studying the impact of information services, such as the advanced traveller information systems (ATIS), on the system performance is usually based on the observation that the available information helps to reduce the uncertainty of the perceived travel cost for users equipped with ATIS. Hall (1996) studied this impact on different types of users and shown a simple two-route example that information cannot degrade network performance. He also suggested the use of ATIS to achieve the user equilibria instead of shaping the traffic toward system optimal point because users "certainly will not rely on sources that run counter to their own observations." The benefit in path choice from a mixture of guided and unguided travellers has been studied in Yang (1998). Informed drivers are assumed to know exactly the realisation of link cost via messages from ATIS, while others choose their routes with random perception errors in their evaluation of the travel cost. Using a more advanced traffic model based on the cell transmission model, Lo et al. (2004) developed a complementary formulation for the stochastic UE route choice and compared the results between static and dynamic paradigms. Bifulco et al. (2016) focused on the stability of UE under ATIS in the day-to-day dynamics of the traffic network. Works in this category include mathematical programming models (Spiess et al., 1989; Marcotte et al., 1998; Yang, 1998; Unnikrishnan et al., 2009), variational inequality models

(Han, 2003; Hamdouch et al., 2004; Connors et al., 2009; Liu et al., 2009; Xie et al., 2014), fixed-point models (Gao, 2012), dynamic programming models (Kim et al., 2005; Xiao et al., 2014), and heuristic approaches (Gao et al., 2006; Ukkusuri et al., 2007; Gao et al., 2012). The literature on this direction of research is extensive, for more details, see Chorus et al. (2006), De Palma et al. (2012), Huang et al. (2008), and Lam et al. (2008).

Given an up-to-date information, travellers adapt their decision of route choices according to the current network state. Particularly, Polychronopoulos et al. (1996), then followed by those in Gao et al. (2006), Gao et al. (2012), and Gao (2012), introduced the concept of learning where users could eliminate infeasible realisations of link cost via the knowledge gained as they traversed the network. To investigate the network-level impact of information using DTA, Gao (2012) developed a fixed-point framework in which the information is the link travel time distribution inputting into their route choice model.

Similarly, with stochastic and flow dependent arc cost, Unnikrishnan et al. (2009) assumed that users know the link travel costs but only when they arrive at the upstream node of these links, which enables them to choose the best next link before continuing their journey. Other studies by Kim et al. (2005), Xiao et al. (2014), and Dong et al. (2013) relied on the knowledge of road state transition (or link costs correlation) to reduce the uncertainty of travel time estimation in routing. In transit networks, Hamdouch et al. (2004) and Hamdouch et al. (2014) studied the strategic decisions aided by the link access probability due to incidents or supply capacity drop. Their problem was formulated in the form of variational inequalities to achieve the user equilibrium solutions.

In summary, Table 2.3 compares the above LWR-based DTA models with the proposed models in Chapters 5 and 6 in the context of adaptive routing with real-time information. The following chapters detail the thesis' contributions.

TABLE 2.3: Summary of selected LWR-based information-based DTA models.
Abbreviations: **ISM**: Incremental Solution Method, **N**: No, **Y**: Yes.

Paper(s)	**SO or UE**	**Traffic model**	**Route choice**	**Quantity of SO-UE difference**	**Uncertainty**	**Evolving information**	**Solution method**
Waller et al. (2013)	Path-based SO	CTM	Path	-	Demand	N	Linear method
Lo et al. (2004)	Path-based UE NCP	CTM	Path	N	Travel time	N	Projection method
Gao (2012)	UE Fixed-point	Simulation-based	Policy	N	Travel time	Y	Fixed-point method
This thesis (Chapter 5)	SO	CTM	Path + Policy	-	Demand + network capacity	Y	Linear method
This thesis (Chapter 6)	non-linear objective UE	path-based LTM	Path	Y	Demand + network capacity	Y	Link-based linear ISM

Chapter 3

A SO-DTA model for single O-D traffic networks

The system optimal dynamic traffic assignment (SO-DTA) problem aims to determine a time-dependent routing pattern of travellers in a network such that the given time-dependent origin-destination demands are satisfied and the total travel time is at a minimum, assuming some model for dynamic network loading. The network kinematic wave model is now widely accepted as such a model, given its realism in reproducing phenomena such as transient queues and spillback to upstream links. An attractive solution strategy for SO-DTA based on such a model is to reformulate as a set of side constraints apply a standard solver, and to this end two methods have been previously proposed, one based on the discretisation scheme known as the Cell Transmission Model (CTM), and the other based on the Link Transmission Model (LTM) derived from variational theory. This chapter aims to combine the advantages of CTM (in tracking time-dependent congestion formation within a link) with those of LTM (avoiding cell discretisation, providing a more computationally attractive with much fewer constraints). The motivation for this work is the previously-reported possibility for DSO to have multiple solutions, which differ in where queues are formed and dissipated in the network. This work aims to find DSO solutions that optimally distribute the congestion over links inside the network which essentially eliminate avoidable queue spillbacks. In order to do so, it requires more information than the LTM can offer, but wish to avoid the computational burden of CTM for DSO. It thus adopts an extension of the LTM called the Two-regime Transmission Model (TTM), which is consistent with LTM at link entries and exits

but which is additionally able to accurately track the spatial and temporal formation of the congestion boundary within a link (which we later show to be a critical element, relative to LTM). It sets out the theoretical background necessary for the formulation of the network-level TTM as a set of linear side constraints. Numerical experiments are used to illustrate the application of the method to determine DSO solutions avoiding spillbacks, reduce/eliminate the congestion and to show the distinctive elements of adopting TTM over LTM. Furthermore, in comparison to a fine-level CTM-based DSO method, our formulation is seen to significantly reduce the number of linear constraints while maintaining a reasonable accuracy.

The organisation of content in this chapter follows. Section 3.2 briefly reviews the development of the TTM for dynamic network loading problems which has been published previously in Balijepalli et al. (2013). Section 3.3 formulates the TTM as a set of linear constraints for a DSO traffic assignment problem in which the queue lengths are calculated explicitly. Section 3.4 presents a new optimisation framework which finds DSO solutions optimally distributing the queue lengths over links in the network using the linear constraints formulated in Section 3.3. Section 3.6 illustrates our numerical studies for a small and a reasonably large network in order to support the advantages of this proposed approach. Finally, we conclude this chapter in Section 3.7.

3.1 Notation

A traffic network is considered a directed graph, which consists of a set of links connected via a set of nodes. The notation below will define a traffic network being considered in this chapter.

- $\mathbb{T}$: a set of discrete time slots, $\mathbb{T} \subset \mathbb{N}$. For continuous time domain, we use notation $\mathbb{T}_{\mathbb{R}}$ ($\mathbb{T}_{\mathbb{R}} \subset \mathbb{R}$). We also define $T = |\mathbb{T}|$ as the number of time slots.
- $\mathbb{V}$: Set of nodes. There are two subsets: set of source nodes $\mathbb{V}_R$ and set of sink nodes $\mathbb{V}_S$, such that: $(\mathbb{V}_R \cup \mathbb{V}_S) \subseteq \mathbb{V}$ and, $\mathbb{V}_R \cap \mathbb{V}_S = \varnothing$.
- $\mathbb{A}$: Set of directed links, combined by any two nodes in $\mathbb{V}$, e.g. if $(n, m) \in \mathbb{A}$ then we call n the upstream node, and m the downstream node of this link.

There are three subsets of links: normal, source ($\mathbb{A}_R$) and sink ($\mathbb{A}_S$) links. The last two types are the specialised and virtual links to provide features of source nodes and sink nodes. For each link $a \in \mathbb{A}$:

- l_a: physical length.
- Υ_a^-: set of inflow links to link a.
- Υ_a^+: set of outflow links from link a.
- $S_a(t)$: upstream capacity (supply).
- $D_a(t)$: downstream capacity (demand).
- $u_a(t)$: real-time upstream traffic.
- $v_a(t)$: real-time downstream traffic.
  - Note that: $u_a(t)$, $v_a(t)$, $S_a(t)$ and $D_a(t)$ are continuous in time domain $\mathbb{T}_{\mathbb{R}}$.

Also for each node $n \in \mathbb{V}$:

- Γ_n^-: set of inflow links to node n.
- Γ_n^+: set of outflow links from node n.

$$(\Gamma_n^- \cup \Gamma_n^+) \subseteq \mathbb{A}$$
$$\Gamma_n^- \cap \Gamma_n^+ = \varnothing$$
$$\Upsilon_a^- = \{e | a \in \Gamma_n^+; e \in \Gamma_n^-\}$$
$$\Upsilon_a^+ = \{e | a \in \Gamma_n^-; e \in \Gamma_n^+\}$$

- $f_{ab}(t)$: upstream traffic at link b, coming from downstream traffic at link a.

$$a, b \in \mathbb{A}$$
$$a \cap b \neq \varnothing$$

- (V_a, W_a, K_a): the set of fundamental diagram parameters, e.g. free-flow speed, backward speed and maximum density, for each link $a \in \mathbb{A}$. The maximum

flow is $Q_a = \frac{K_a V_a W_a}{V_a + W_a}$, and the critical density is $\frac{K_a W_a}{V_a + W_a}$.

$$V_a \geq W_a$$

- $l_a^c(t)$: length of congested regime in link a.
- $l_a^f(t)$: length of free-flow regime in link a.
- Common indexes: a for links in $\mathbb{A}$, n for nodes in $\mathbb{V}$, i for time slots in $\mathbb{T}$, t for continuous time in $\mathbb{T}_{\mathbb{R}}$.

Network formation Two rules for source and sink nodes:

- There is only one link from a source node, called source link.
- There is only one link to a sink node, called sink link.

3.2 Overview of Two regime Transmission Model

This section briefly discusses the concept of the TTM for traffic dynamics in a link, which has been previously developed in Balijepalli et al. (2013), and explains the potential advantages of TTM in the context of the DSO traffic assignment problems. Define the fundamental relationship according to the LWR model:

$$\frac{\partial \rho}{\partial t} + \frac{\partial q}{\partial x} = 0 \tag{3.1}$$

where $\rho = \rho(x,t)$ and $q = q(x,t)$ denote, respectively, the density and flow at continuous location x and continuous time t. Here the link index a is dropped for the sake of simplicity since Eq. (3.1) holds for any link. Note, by definition, $q(0,t) \doteq u(t)$ and $q(l,t) \doteq v(t)$. Furthermore, the flow is assumed to be a function of the density, which refers to a so-called fundamental diagram. Thus we may write:

$$q(x,t) = \phi\left(\rho(x,t)\right) \tag{3.2}$$

for some function $\phi(.)$ that is independent of x and t. We shall specifically focus on the case of a triangular flow-density relationship:

$$\phi(\rho) = \begin{cases} V\rho, & \text{if } 0 \leq \rho \leq C; \\ W\left(K-\rho\right), & \text{if } C \leq \rho \leq K. \end{cases} \tag{3.3}$$

where C is the critical density, defined by $C = \dfrac{KW}{V+W}$. Together, Eqs. (3.1)-(3.3) define the LWR model for a link, for given entry and exit flow profiles.

Rewrite Eq. (3.1) as:

$$\frac{\partial \rho}{\partial t} + \frac{d\left[\phi(\rho)\right]}{d\rho}\frac{\partial \rho}{\partial x} = 0 \tag{3.4}$$

and apply the method of characteristics as described by Newell (1993). In this approach, from any given point (x,t) with density $\rho(x,t)$, we are able to trace along a curve in space-time of constant density, back in time to a boundary location: in our case, the entry-point to or exit-point from the link. For homogeneous links, this space-time curve is linear, such that for our given (x,t), we have constant density along the space-time points given by tracing back to any earlier time s along:

$$\rho(x,t) = \rho\left(x - \dot{\phi}(\rho)s, t-s\right) \qquad (0 \leq s \leq t; 0 \leq x - \dot{\phi}(\rho)s \leq l) \tag{3.5}$$

where $\dot{\phi}(\rho) \doteq \dfrac{d\left[\phi(\rho)\right]}{d\rho}$. Adopting the particular relationship Eq. (3.3) gives:

$$\rho(x,t) = \begin{cases} \rho(x - Vs, t-s), & \text{if } 0 \leq \rho \leq C; \\ \rho(x + Ws, t-s), & \text{if } C \leq \rho \leq K. \end{cases} \tag{3.6}$$

We then trace to the link-entry (in free-flow conditions) or link-exit (in congested conditions) by the particular choices:

$$s \begin{cases} \dfrac{x}{V}, & \text{if } 0 \leq \rho \leq C; \\ \dfrac{l-x}{W}, & \text{if } C \leq \rho \leq K. \end{cases} \tag{3.7}$$

which is substituted into Eq. (3.6):

$$\rho(x,t) = \begin{cases} \rho(0, t - \frac{x}{V}), & \text{if } 0 \leq \rho \leq C; \\ \rho(l, t - \frac{l-x}{W}), & \text{if } C \leq \rho \leq K. \end{cases} \tag{3.8}$$

and we are then able to relate the flows corresponding to the densities in Eq. (3.8) to the boundary conditions:

$$q(x,t) = \phi(\rho) = \begin{cases} \phi\left(\rho(0, t - \frac{x}{V})\right) = u\left(t - \frac{x}{V}\right), & \text{if } 0 \leq \rho \leq C; \\ \phi\left(\rho(l, t - \frac{l-x}{W})\right) = v\left(t - \frac{l-x}{W}\right), & \text{if } C \leq \rho \leq K. \end{cases} \tag{3.9}$$

The novel feature of the method we shall now describe is that it will utilise Eqs. (3.8) and (3.9) in a simplified way that neither requires us to treat the state of the link as homogeneous along its length (as is implicitly done in whole-link models, for example), nor requires a fixed, fine discretisation scheme to be applied (as in the CTM, for example). It is also worth noticing that the proposed method can capture the traffic states in time and space (i.e., the formation and dissipation of traffic congestions) along the link as CTM while the LTM cannot.

3.3 Linear-constraint formulation of the TTM

This section will show how it is possible to formulate the continuous TTM for a single link equivalently as a system of (discrete) linear constraints and how to determine the time and space evolution of the length of the congested regime which will be used in finding the DSO solutions in Section 3.4.

3.3.1 Definition of the discrete time variables

Let us define the following discrete time variables from the continuous TTM described in Section 3.2:

$$v(i) = \int_i^{i+1^-} v(t)\, \mathrm{d}t$$
$$u(i) = \int_i^{i+1^-} u(t)\, \mathrm{d}t$$

$$
\begin{aligned}
f(i) &= \int_{i}^{i+1^{-}} f(t)\,\mathrm{d}t \\
D(i) &= \int_{i}^{i+1^{-}} D(t)\,\mathrm{d}t \\
S(i) &= \int_{i}^{i+1^{-}} S(t)\,\mathrm{d}t \qquad\qquad \forall i \in \mathbb{T}.
\end{aligned} \tag{3.10}
$$

For the sake of simplicity without loss of generality let us assume that

$$
\begin{aligned}
\frac{l_a}{V_a} &\in \mathbb{N} \\
\frac{l_a}{V_a} &\in \mathbb{N}
\end{aligned}
$$

$\forall a \in \mathbb{A}$. In general, we could ignore this assumption but the model becomes more complex and may bring numerical issues when working with the non-integer values. As suggested in Ban et al. (2012), the mesh size should be properly chosen to avoid such non-integer problems.

3.3.2 Definition of the system variables

Basically, we will consider the number of vehicles in a link a major variable for the DNL problem. Nevertheless, the computation of such variables will take into account the dynamics of the length of the congested regimes described by the TTM. First, let us define the the number of vehicles in a link at time instant $t \in \mathbb{T}_{\mathbb{R}}$ as the number of vehicles currently staying in this link at time t, mathematically represented by:

$$
n(t) = \int_{0}^{t} \left(u(h) - v(h)\right) \mathrm{d}h \tag{3.11}
$$

In discrete time step, the number of vehicles in a link at time step $i \in \mathbb{T}$, means:

$$
n(i) = \lim_{t \to (i+1)^{-}} \int_{0}^{t} \left(u(h) - v_a(h)\right) \mathrm{d}h = \sum_{k=0}^{i} \left[u(k) - v(k)\right] \tag{3.12}
$$

The number of vehicles in a link can also be defined as:

$$
n(t) \doteq \int_{0}^{l} \rho(x,t)\,\mathrm{d}x. \tag{3.13}
$$

From the definition of the TTM in Section 3.2, the number of vehicles in a link consists of the total number of vehicles staying in the non-congested regime and in the congested regime at time instant t. Substituting Eq. (3.8) into Eq. (3.13) gives:

$$\begin{aligned} n(t) &= \underbrace{\frac{1}{V}\int_{0}^{l^f(t)} u\left(t-\frac{x}{V}\right)\,\mathrm{d}x}_{\text{Free-flow regime}} + \underbrace{\int_{l^f(t)}^{l}\left(K-\frac{1}{W}v(t-\frac{l-x}{W})\right)\,\mathrm{d}x}_{\text{Congested regime}} \\ &= \underbrace{\frac{1}{V}\int_{0}^{l^f(t)} u\left(t-\frac{x}{V}\right)\,\mathrm{d}x}_{\text{Free-flow regime}} + \underbrace{\int_{0}^{l^c(t)}\left(K-\frac{1}{W}v\left(t-\frac{x}{W}\right)\right)\,\mathrm{d}x}_{\text{Congested regime}}. \end{aligned} \tag{3.14}$$

This equation indicates that the number of vehicles in a link determined in each traffic regime by Eq. (3.8) has been implicitly incorporated and will be used later to formulate the linear discrete link constraints.

3.3.3 Linear constraints

For the sake of simplicity, we drop out the link index in this section but the derivation holds for every link. From Eq. (3.14), we can compute two specific numbers of vehicles, $n^f(i)$ for $l^f(i) = l$ and $n^c(i)$ for $l^c(i) = l$, to define the states of link (i.e., either fully non-congested or fully congested):

$$n^c(i) = K.l - \sum_{k=i+1-\frac{l}{W}}^{i} v(k) \tag{3.15}$$

$$n^f(i) = \sum_{k=i-\frac{l}{V}+1}^{i} u(k) \tag{3.16}$$

These specific numbers of vehicles are the bounds of the number of vehicles in a link as shown below.

Lemma 1 (Upper bound)**.**

$$n^c(i) \geq n(i) = \sum_{k=0}^{i} \left[u(k) - v(k)\right]; \forall i \in \mathbb{T} \tag{3.17}$$

Proof. Since the inflow and outflow are constrained by the link capacity:

$$u(t) \leq \frac{KVW}{V+W}$$
$$v(t) \leq \frac{KVW}{V+W}$$

we infer that:

$$\frac{1}{V}\int_0^{l^f(t)} u\left(t-\frac{x}{V}\right)\,\mathrm{d}x + \frac{1}{W}\int_0^{l^f(t)} v\left(t-\frac{l-x}{W}\right)\,\mathrm{d}x \leq \qquad Kl^f(t)$$

which is then substituted into Eq. (3.14) to obtain:

$$\begin{aligned} n(t) &\leq K(l^f(t)+l^c(t)) - \frac{1}{W}\int_0^{l^f(t)} v\left(t-\frac{l-x}{W}\right)\,\mathrm{d}x - \frac{1}{W}\int_0^{l^c(t)} v\left(t-\frac{x}{W}\right)\,\mathrm{d}x \\ &= Kl - \frac{1}{W}\int_{l^c(t)}^{l} v\left(t-\frac{x}{W}\right)\,\mathrm{d}x - \frac{1}{W}\int_0^{l^c(t)} v\left(t-\frac{x}{W}\right)\,\mathrm{d}x \\ &= Kl - \frac{1}{W}\int_0^{l} v\left(t-\frac{x}{W}\right)\,\mathrm{d}x \end{aligned}$$

Let $z = t - x/W$ we have: $\int_0^l v(t-\frac{x}{W})\,\mathrm{d}x = W\int_{t-l/W}^{t} v(z)\,\mathrm{d}z$. The above inequality becomes:

$$n(t) \leq Kl - \int_{t-\frac{l}{W}}^{t} v(z)\,\mathrm{d}z$$

which is then converted into discrete time step as:

$$n(i) = \lim_{t\to(i+1)^-} n(t) \leq Kl - \sum_{k=i+1-\frac{l}{W}}^{i} v(k) = n^c(i).$$

□

Lemma 2 (Lower bound)**.**

$$n^f(i) \leq n(i) = \sum_{k=0}^{i} \left[u(k) - v(k)\right]; \forall i \in \mathbb{T} \tag{3.18}$$

Proof. From the constraints of the inflow and outflow, we infer that:

$$\frac{1}{V}\int_0^{l^c(t)} u(t-\frac{l-x}{V})\,\mathrm{d}x + \frac{1}{W}\int_0^{l^c(t)} v(t-\frac{x}{W})\,\mathrm{d}x \le Kl^c(t)$$

which is then substituted into Eq. (3.14) to obtain:

$$n(t) \ge \frac{1}{V}\int_0^{l^f(t)} u(t-\frac{x}{V})\,\mathrm{d}x + \frac{1}{V}\int_0^{l^c(t)} u(t-\frac{l-x}{V})\,\mathrm{d}x$$

which can be converted to:

$$n(t) \ge \qquad \frac{1}{V}\int_0^{l} u(t-\frac{x}{V})\,\mathrm{d}x = \int_{t-\frac{l}{V}}^{t} u(z)\,\mathrm{d}z$$

or in the discrete time step:

$$n(i) = \lim_{t\to(i+1)^-} n(t) \ge \lim_{t\to(i+1)^-} \int_{t-\frac{l}{V}}^{t} u(z)\,\mathrm{d}z = \sum_{k=i+1-\frac{l}{V}}^{i} u(k) = n^f(i)$$

□

In sum, the number of vehicles in a link has two linear constraints:

$$\sum_{k=i+1-\frac{l}{V}}^{i} u(k) \le n(i) \le Kl - \sum_{k=i+1-\frac{l}{W}}^{i} v(k). \tag{3.19}$$

3.3.4 Computation of the length of the congested regime

It can be seen from the TTM described in Section 3.2 that if the length of the congested regime is given, we can compute the phase-space density $\rho(x,t)$ and consequently the number of vehicles per link $n(t)$ via Eq. (3.14). In this section, we will solve an inverse problem: if the (discrete) number of vehicles per link $n(i)$ is given, we need to compute the regime length $l^c(i)$ or $l^f(i)$. To this end, we will solve a *fixed point* problem below via a system of linear equations.

Theorem 1. *Given l and a feasible value of $n(i)$, there exists a solution of $l^c(i)$ and $l^f(i)$ satisfying:*

$$\begin{aligned} l &= l^c(i) + l^f(i) \\ n(i) &= \lim_{t\to(i+1)^-} \frac{1}{V} \int_0^{l^f(t)} u\left(t - \frac{x}{V}\right)\, \mathrm{d}x + \int_0^{l^c(t)} (K - \frac{1}{W} v\left(t - \frac{x}{W})\right)\, \mathrm{d}x \end{aligned} \tag{3.20}$$

Proof. We define $F(x)$ as:

$$F(x) = \frac{1}{V} \int_0^{l-x} u\left(t - \frac{y}{V}\right)\, \mathrm{d}y + \int_0^{x} \left(K - \frac{1}{W} v\left(t - \frac{y}{W}\right)\right)\, \mathrm{d}y.$$

From the definition of the congested link density and free-flow link density:

$$n^c(i) = \lim_{t\to(i+1)^-} \int_0^{l} K - \frac{1}{W} v\left(t - \frac{x}{W}\right)\, \mathrm{d}x = \lim_{t\to(i+1)^-} F(l)$$

$$n^f(i) = \lim_{t\to(i+1)^-} \int_0^{l} \frac{1}{V} v\left(t - \frac{x}{V}\right)\, \mathrm{d}x = \lim_{t\to(i+1)^-} F(0)$$

It is straightforward to obtain:

$$\frac{dF(x)}{dx} = \frac{-1}{V} u\left(t - \frac{l-x}{V}\right) + K - \frac{1}{W} v\left(t - \frac{x}{W}\right) \geq 0.$$

Since,

$$u\left(t - \frac{l-x}{V}\right) \leq \frac{KVW}{V+W}$$
$$v\left(t - \frac{x}{W}\right) \leq \frac{KVW}{V+W}$$

hence $\frac{\mathrm{d}F(x)}{\mathrm{d}x} \geq 0$. Therefore, $F(x)$ is monotonically increasing in $(0 \leq x \leq l)$, which leads to:

$$F(0) \leq F(x) \leq F(l).$$

From the definition of number of vehicles in the congested and non-congested regime:

$$n^c(i) = \lim_{t\to(i+1)^-} \int_0^l K - \frac{1}{W} v\left(t - \frac{x}{W}\right)\, \mathrm{d}x = \lim_{t\to(i+1)^-} F(l)$$
$$n^f(i) = \lim_{t\to(i+1)^-} \int_0^l \frac{1}{V} v\left(t - \frac{x}{V}\right)\, \mathrm{d}x = \lim_{t\to(i+1)^-} F(0)$$

Since $F(x)$ is continuous on x, using the *fixed point* principle, for any value $n(i) \in (n^f(i), n^c(i))$ we could find x^* satisfying: $F(x^*) = f(i)$. This means that we can compute the dynamics of the length of the congested regime $l^c(i)$ given the (discrete) number of vehicles per link $n(i)$. □

Theorem 2. *Given l and a feasible value of $n(i)$. The solution of the length of the congested regime in Theorem 1 is unique if the following conditions hold:*

$$u(h) < Q \qquad \forall h \in \left[t - \frac{l}{V}, t\right]$$
$$v(h) < Q \qquad \forall h \in \left[t - \frac{l}{W}, t\right]$$

where $Q = \dfrac{KVW}{V+W}$.

Proof. As in the proof of Theorem 1, we use:

$$F(x) = \frac{1}{V}\int_0^{l-x} u\left(t - \frac{y}{V}\right)\, \mathrm{d}y + \int_0^x \left(K - \frac{1}{W} v\left(t - \frac{y}{W}\right)\right)\, \mathrm{d}y$$

and:

$$\frac{dF(x)}{dx} = K - \frac{1}{V} u\left(t - \frac{l-x}{V}\right) - \frac{1}{W} v\left(t - \frac{x}{W}\right) \geq 0$$

With the above upper bound constraints on upstream flow and downstream flow, we derive that:

$$u\left(t - \frac{l-x}{V}\right) < Q \qquad \forall x \in [0, l]$$
$$v\left(t - \frac{x}{W}\right) < Q \qquad \forall x \in [0, l]$$
$$\Rightarrow \frac{dF(x)}{dx} > K - \frac{Q}{V} - \frac{Q}{W} = 0 \qquad \forall x \in [0, l].$$

It guarantees the positive value $\frac{dF(x)}{dx} > 0 \; \forall x \in [0, l]$ at a given time t. This implies that for a feasible value of $n(t)$, there is a unique solution of x so that $F(x) = n(t)$. □

The condition of unique length of the congested regime will avoid the critical point which is related to critical density and maximum flow in the triangular fundamental diagram. If this condition happens, it assures that maximum flow does not appear at any location on the link at a given time and there is unique solution of TTM at that time. A simple way to achieve this condition is that the dynamic maximum flow $q(t)$ is bounded to be strictly less than Q. Furthermore, even if the flow can reach Q at some times, but two bounds of the number of vehicles per link satisfying the strict inequality at given time t, e.g., $n^f(t) < n(t) < n^c(t)$, it also implies the condition in Theorem 2: there is a unique solution of the length of the congested regime at this time.

Theorem 3 (Non-congested condition)**.**

$$l^c(i) = 0 \Leftrightarrow n^f(i) = \sum_{k=0}^{i} [u_a(k) - v_a(k)] \tag{3.21}$$

Proof.

- Necessary condition: If $l^c(i) = 0$, traffic along the whole link is in the non-congested state, we have:

$$\sum_{k=0}^{i} [u(k) - v(k)] = \lim_{t \to (i+1)^-} \frac{1}{V} \int_0^l u\left(t - \frac{x}{V}\right) \, \mathrm{d}x = \lim_{t \to (i+1)^-} \int_{t - \frac{l}{V}}^{t} u(x) \, \mathrm{d}x$$
$$= \sum_{k=i+1-\frac{l}{V}}^{i} u(k) = n^f(i)$$

- Sufficient condition: If $n^f(i) = \sum_{k=0}^{i} [u(k) - v(k)]$: From Lemma 2, we infer that the equality happens when:
  - $l^c(t) = 0$, or

– $u(t-\frac{l-x}{V}) = v\left(t-\frac{x}{W}\right) = \frac{KVW}{V+W}; \forall x \in [0, l^c(t)]$. This condition happens when traffic is operating at the critical density, at which $l^c(t) = 0$ by definition.

□

Theorem 4 (Fully-congested condition)**.**

$$l^c(i) = l \Leftrightarrow n^c(i) = \sum_{k=0}^{i} [u(k) - v(k)] \tag{3.22}$$

Proof. Based on Lemma 1, with similar method in Theorem 3. □

These two theorems indicate that when the number of vehicles per link is equal to its bounds, the link is either fully non-congested or congested. The algorithm below shows how to compute the dynamics of the length of the congested regime in case the number of vehicles per link is within its bounds. Let us define,

$$\begin{aligned} j^f &= \lceil i+1-\frac{l^f(i)}{V} \rceil \\ j^c &= \lceil i+1-\frac{l^c(i)}{W} \rceil \\ \Delta j^f &= j^f - \left(i+1-\frac{l^f(i)}{V}\right) \\ \Delta j^c &= j^c - \left(i+1-\frac{l^c(i)}{W}\right) \end{aligned}$$

then we get:

$$\begin{aligned} n(t) &= \frac{1}{V}\int_0^{l^f(t)} u\left(t-\frac{x}{V}\right)\,\mathrm{d}x + \int_0^{l^c(t)} \left(K - \frac{1}{W}v(t-\frac{x}{W})\right)\,\mathrm{d}x \\ &= K.l^c(t) + \int_{t-\frac{l^f(t)}{V}}^{t} u(h)\mathrm{d}h - \int_{t-\frac{l^c(t)}{W}}^{t} v(h)\mathrm{d}h \\ n(i) = \lim_{t\to i+1^-} n(t) &= K.l^c(i) + \sum_{j^f}^{i} u(k) - \sum_{j^c}^{i} v(k) \\ &\quad + \Delta j^f.u(j^f-1) - \Delta j^c.v(j^c-1) \end{aligned}$$

It is worth mentioning that the following constraints are imposed for each regime length:

$$\begin{aligned} l^f(t) + l^c(t) &= l \\ l^f(t) &\geq 0 \\ l^c(t) &\geq 0 \end{aligned} \tag{3.23}$$

From the Theorems 3 and 4, we adopt the following algorithm to compute the length of the congested regime using the number of vehicles per link.

- If $n(i) = n^f(i)$, then $l^c(i) = 0$ (Theorem 3).
- If $n(i) = n^c(i)$, then $l^c(i) = l$ (Theorem 4).
- If $n^f(i) < n(i) < n^c(i)$, then let's set up $i^c = 1$ and $i^f = 0$.
  - Increase i^c until these conditions happen:

$$n(i) \geq K.W.i^c - \sum_{i+1-i^c}^{i} v(k) + \sum_{j^f}^{i} u(k) + \Delta j^f.u(j^f - 1)$$

$$n(i) < K.W.(i^c + 1) - \sum_{i-i^c}^{i} v(k) + \sum_{j^f}^{i} u(k) + \Delta j^f.u(j^f - 1)$$

  - Similarly, increase i^f until these conditions happen:

$$n(i) \leq K.l^c(i) - \sum_{j^c}^{i} v(k) - \Delta j^c.v(j^c - 1) + \sum_{i+1-i^f}^{i} u(k)$$

$$n(i) > K.l^c(i) - \sum_{j^c}^{i} v(k) - \Delta j^c.v(j^c - 1) + \sum_{i-i^f}^{i} u(k)$$

  - Solve the linear system of equations below to find $l^c(i)$ and $l^f(i)$.

$$\begin{aligned} l^c(i) &= i^c W + \Delta l^c \\ l^f(i) &= i^f V + \Delta l^f \\ \Delta l^c + \Delta l^f &= l - i^f V - i^c W \qquad \text{(to satisfy Eq. (3.23))} \\ n(i) &= K.l^c(i) + \sum_{i+1-i^f}^{i} u(k) - \sum_{i+1-i^c}^{i} v(k) \\ &\quad + \frac{\Delta i^f}{V}.u(i - i^f) - \frac{\Delta i^c}{W}.v(i - i^c) \end{aligned}$$

The above method is used as a benchmarking value for the linearised congested regime for the DSO problem presented in Section 3.4.

3.3.5 Approximation of the congested regime

Let us assume that the time step is selected at at least l/W (Ban et al., 2012) so that $u(i)$ and $v(i)$ are steady in $i \in [0, \frac{l}{W} - 1]$, that is:

$$u(t_0) = u(t_0 + i) = u(i) \tag{3.24}$$

$$v(t_0) = v(t_0 + i) = v(i) \tag{3.25}$$

then, based on Theorem 1, we can obtain:

$$n(i) \approx l^c(i)\,(K - u(i)/V - v(i)/W) + \frac{u(i)l}{V}.$$

So the approximation of l^c read:

$$l^c(i) \approx \frac{n(i) - \frac{u(i)l}{V}}{K - \frac{u(i)}{V} - \frac{v(i)}{W}}. \tag{3.26}$$

Using Eqs. (3.24) and (3.25) for Lemmas 1 and 2 leads to:

$$n^c(i) \approx Kl - \int_{i-\frac{l}{W}}^{i} v(k)\,\mathrm{d}h = Kl - \frac{v(i)l}{W} \tag{3.27}$$

$$n^f(i) \approx \int_{i-\frac{l}{V}}^{i} u(k)\,\mathrm{d}h = \frac{u(i)l}{V}. \tag{3.28}$$

To substitute Eqs. (3.27) and (3.28) to Eq. (3.26) we obtain:

$$\tilde{l}^c(i) \approx \frac{n(i) - n^f(i)}{n^c(i) - n^f(i)}l. \tag{3.29}$$

Eq. (3.29) will be used in Section 3.4 for the new optimisation framework, which optimally redistributes the vehicle queues (i.e., the dynamics of l^c over links).

3.4 Formulation of TTM-based DSO-DTA model

The DSO-DTA problem deals with a directed network $\mathbb{G}(\mathbb{V}, \mathbb{A})$ where $\mathbb{V}$ is the set of nodes and $\mathbb{A}$ is the set of the directed links. During each time instant i in the study period $\mathbb{T}$, the single origin node R has time-dependent demand $D(i)$ heading toward the destination S. The network is assumed to be empty at $i = 0$ and will be cleared at $i = T$. Suppose vehicles travel along each link according to the TTM principle and across any immediate node according to the demand-supply scheme (Daganzo, 1995). The DSO-DTA problem is to find an optimal dynamic traffic assignment pattern which results in the minimal total system travel time. Herein we only consider the DSO-DTA problem for a single origin-destination network. Nevertheless, the modelling and solution methods underpinned by an efficient implicit in-link congestion tracking approach for single origin-destination DSO problems proposed in this chapter will still contribute to the state-of-the-art. Recent research by Shen et al. (2014) have indicated that DSO may have multiple solutions in which all the solutions share the same objective while the difference is where the queues are formed and dissipated in the network. This section formulates a problem to find DSO solutions which can evenly distribute the queues (i.e., the length of the congested regime l^c over links in the network so that the spillbacks can be avoided. To this end, we present a new optimisation problem by adding a new step to the original DSO problem to optimally distribute the congestion regime over links in the network. The new step relies on how the congested length l^c is utilised if the front shock stays within a link (note that this step cannot be modelled with the LTM or DQM).

3.4.1 Objective functions

1. **Stage 1:** find DSO solutions via minimising the total travel times or the total system cost. The total system cost that all the vehicles spent in the network can be calculated as the total number of vehicles existing in the network for each time interval $i \in \mathbb{T}$, which is defined by the difference between the cumulative departure at the origin and the cumulative arrival at the destination for time

interval i as below:

$$TT = \sum_{i \in \mathbb{T}} \left(\sum_{k=0}^{i} U(k) - \sum_{k=0}^{i} \sum_{a \in \mathbb{A}_S} u_a(k) \right). \tag{3.30}$$

Since the demand $U(i)$ is given $\forall i \in \mathbb{T}$, the objective to minimise the total travel cost TT is equivalent to maximise the total outflow Shen et al. (2008):

$$F = \sum_{a \in \mathbb{A}_S} \sum_{i \in \mathbb{T}} \sum_{k=0}^{i} u_a(k) = \sum_{a \in \mathbb{A}_S} \sum_{i \in \mathbb{T}} (T + 1 - i) u_a(i). \tag{3.31}$$

Let F^* denote the optimal value of F to achieve the DSO solutions. That is:

$$F^* = \max \sum_{a \in \mathbb{A}_S} \sum_{i \in \mathbb{T}} (T + 1 - i) u_a(i) \tag{3.32}$$

subject to the node and link constraints in Sections 3.4.3 and 3.4.4.

2. **Stage 2:** find traffic patterns which optimally distribute the congestion in links.

   Objective: G, defined as the function of the queue length $l_a^c(i)$.

   Constraints: $\sum_{a \in \mathbb{A}_S} \sum_{i \in \mathbb{T}} (T + 1 - i) u_a(i) = F^*$, (i.e., to achieve the DSO solutions), and the node and link constraints in Sections 3.4.3 and 3.4.4.

   This step restricts the solution domain to system optimal solutions via the first constraint. As reported in Shen et al. (2014), there exists a solution without changing the system optimal objective value. It indirectly shows that Stage 2 provides a feasible solution domain. By optimising G we can find a DSO solution that optimally distributes the congestion over links which can lead to minimal spillbacks or even free-flow situations as shown later in our numerical study.

Remark 1. *To account for the non-holding back issue, the new constraints in Eqs.* (3.39) *and* (3.40) *below can be added into the optimisation framework above. Nevertheless, such constraints are non-linear and we will show later in our example how they affect the nature and solutions of our framework*

Remark 2. *In this chapter, we consider two specific objective functions G below for the illustration purposes:*

*(a) Optimising the heterogeneity of the queue lengths (**MDCL**) in the network:*

$$G = \min \sum_{a \in \mathbb{A}} \Big(- \bar{l}^c + \sum_{i \in \mathbb{T}} \frac{l_a^c(i)}{l_a} \Big)^2 \tag{3.33}$$

where,

$$\bar{l}^c = \frac{1}{|\mathbb{A}|} \sum_{a \in \mathbb{A}} \sum_{i \in \mathbb{T}} \frac{l_a^c(i)}{l_a} = \frac{1}{|\mathbb{A}|} \sum_{a \in \mathbb{A}} \sum_{i \in \mathbb{T}} \frac{n_a(i) - n_a^f(i)}{n_a^c(i) - n_a^f(i)}. \tag{3.34}$$

*(b) Optimising the total queue lengths (**MCL**) in the network:*

$$G = \min \sum_{a \in \mathbb{A}} \sum_{i \in \mathbb{T}} \frac{l_a^c(i)}{l_a}. \tag{3.35}$$

3.4.2 Non-holding back problem

The non-holding back (NHB) solutions require the system to discharge flow as much as it can and have been widely discussed before in the literature (Shen et al., 2014; Shen et al., 2007). For our particular TMM-DSO problem, the non-holding back condition is that, for any direct connection from link a to link b at any time $i \in \mathbb{T}$, at least one of the following conditions must hold:

$$v_a(i) = \underbrace{\min \Big\{ Q_a(i); \sum_{k=0}^{i-l_a/V_a} u_a(k) - \sum_{k=0}^{i-1} v_a(k) \Big\}}_{D_a(i)} \tag{3.36}$$

$$u_b(i) = \underbrace{\min \Big\{ Q_b(i); K_b l_b + \sum_{k=0}^{i-l_b/W_b} v_b(k) - \sum_{k=0}^{i-1} u_b(k) \Big\}}_{S_b(i)} \tag{3.37}$$

$$\sum_{k \geq i+1} f_{ab}(k) = 0. \tag{3.38}$$

These conditions indicate that, at any time i, if there is some traffic from link a to link b (i.e., $f_{ab}(i) > 0$), then the downstream flow of link a has to reach link-a demand (Eq. (3.36)) or the upstream flow of link b has to reach link-b supply (Eq. (3.37)). Otherwise, there will not be any traffic from link a to link b after time i (Eq. (3.38)). In fact, conditions Eqs. (3.36) and (3.37) follow the demand and supply principle of traffic flow at nodes. Based on Lemmas 1 and 2, these non-holding back conditions

are represented by the following constraint:

$$\big(n_a(i) - n_a^f(i)\big)\big(n_b^c(i) - n_b(i)\big)\big(Q_a(i) - v_a(i)\big)\big(Q_b(i) - u_b(i)\big)\sum_{k\geq i+1} f_{ab}(k) = 0 \quad (3.39)$$

$$\forall a, b \in \mathbb{A} : a \notin \mathbb{A}_R, b \notin \mathbb{A}_S, b \in \mathrm{Y}_a^+.$$

$$n_a(i)\big(n_b^c(i) - n_b(i)\big)\big(Q_b(i) - u_b(i)\big)\sum_{k\geq i+1} f_{ab}(k) = 0 \quad (3.40)$$

$$\forall a \in \mathbb{A}_R, b \in \mathrm{Y}_a^+.$$

$$\big(n_a(i) - n_a^f(i)\big)\big(Q_a(i) - v_a(i)\big)\sum_{k\geq i+1} f_{ab}(k) = 0 \quad (3.41)$$

$$\forall b \in \mathbb{A}_S, a \in \mathrm{Y}_b^-.$$

Eq. (3.39) guarantees the non-holding back traffic at any pair of normal links in the network (i.e., holding-free at very nodes inside the network), while Eqs. (3.40) and (3.41) are used if one of these links is source or sink respectively.

Proposition 1. *System optimal solutions guarantee the holding-free traffic to sinks.*

Proof. Assume that there exists a SO solution such that it does not satisfy non-holding back condition at sinks. It means that there is a flow from normal link a to sink link $b \in \mathbb{A}_S$ at a particular time i so that:

$$\big(n_a(i) - n_a^f(i)\big)\big(Q_a(i) - v_a(i)\big)\sum_{k\geq i+1} f_{ab}(k) > 0.$$

Let's consider flow from link a to link b. With the above condition, there exists time $j > i$ so that $f_{ab}(j) > 0$ while there is still a space for traffic to move from link a to destination link b because: $n_a(i) > n_a^f(i)$ and $Q_a(i) > v_a(i)$. For this reason, we can extract a small traffic δ at time j to move sooner at time i ($\delta < f_{ab}(j)$) without changing other flows (and also no violating to any other constraints):

$$f_{ab}^{\text{new}}(i) = f_{ab}(i) + \delta \Rightarrow u_b^{\text{new}}(i) = u_b(i) + \delta$$
$$f_{ab}^{\text{new}}(j) = f_{ab}(j) - \delta \Rightarrow u_b^{\text{new}}(j) = u_b(j) - \delta.$$

This modification helps to improve the objective F:

$$F^{\text{new}} = F^* + (T + 1 - i)\delta - (T + 1 - j)\delta = F^* + (j - i)\delta > F^*$$

which violates the assumption that F^* is the maximum value of F. For this reason, the traffic to sink is always holding-free. That is Eq. (3.41) can be guaranteed by the SO objective. Note that, we still need to take care of holding-free traffic from sources and other intermediate nodes. □

3.4.3 Link constraints

From the derivations in Section 3.3, the link constraints for the TTM are summarised as below:

$$n_a(i) = \sum_{k=0}^{i} (u_a(k) - v_a(k)) \tag{3.42}$$

$$n_a^f(i) = \sum_{k=i-\frac{l_a}{V_a}+1}^{i} u_a(k) \tag{3.43}$$

$$n_a^c(i) = K_a l_a \sum_{k=i-\frac{l_a}{W_a}+1}^{i} v_a(k) \tag{3.44}$$

$$n_a(i) \geq n_a^f(i) \tag{3.45}$$

$$n_a^c(i) \geq n_a(i) \tag{3.46}$$

for all $i \in \mathbb{T}$.

3.4.4 Node constraints

For all $i \in \mathbb{T}$:

$$\sum_{a \in \Gamma_n^-} v_a(i) = \sum_{a \in \Gamma_n^+} u_a(i) \tag{3.47}$$

$$u_a(i) = \sum_{b \in \Upsilon_a^-} f_{ba}(i) \tag{3.48}$$

$$v_a(i) = \sum_{b \in \Upsilon_a^+} f_{ab}(i) \tag{3.49}$$

$$f_{ab}(i) \geq 0. \tag{3.50}$$

Constraint Eq. (3.47) represents the conservation of the vehicles at an immediate node while Eq. (3.48) determines the total number of vehicles entering a node from all upstream links of that node and Eq. (3.49) defines the total number of vehicles

exiting a node from all downstream links of that node. In fact, Eq. (3.47) and pair of Eqs. (3.48) and (3.49) are equivalent, and we only need one of them in the final model. Eq. (3.50) guarantees the non-negative flow on link. It is worth noticing that no explicit node model is considered in this chapter as we only focus on the DSO problem at this stage. We refer to Flötteröd et al. (2011), Ngoduy (2010), Ngoduy et al. (2005), Smits et al. (2015), and Tampère et al. (2011) for node models in the network.

Source node constraints

$$u_a(i) = U_a(i) \tag{3.51}$$

$$\{Q_a(i)\} \to \infty \tag{3.52}$$

for all $a \in \mathbb{A}_R, i \in \mathbb{T}$, where $U_a(i)$ denotes the demand to source link a at time step i.

Sink node constraints

$$v_a(i) = 0 \tag{3.53}$$

$$Q_a(i) \to \infty \tag{3.54}$$

for all $a \in \mathbb{A}_R, i \in \mathbb{T}$.

The above linear optimisation problem can be solved by any simple optimisation solver. This work uses the CBC solver, one of the popular open-source software in solving optimisation problems (`http://www.coin-or.org/projects/Cbc.xml`), for our case studies.

3.5 Integration of signal control in the DTA model

The increasing availability of information, e.g., obtained from sensors, wireless devices or camera surveillance, enables not only travellers but also network operators to utilize the available data for the purposes of mobility management. In this context, this section aims to study the dynamic interaction between traffic signalling and routing by extending the dynamic traffic assignment (DTA) framework with

a new model of traffic signal at intersections taking into account the transition between signal phases. Note that, a solution for the DTA problem is well-known in the literature to achieve the global optimal route choices, however, it is usually done without realistic traffic signal timings.

Let $\mathcal{C}$ define the fixed signal cycle time parameter, and Π_j define a set of signalling phases at the intersection $j \in \mathbb{J}$ where $\mathbb{J}$ is the intersection set and $\mathbb{J} \subset \mathbb{V}$. Note that, in discrete-time models, we assume the value of $\mathcal{C}$ as the integer number of system time steps. The time domain $\mathbb{T}_c$ corresponding to the c^{th} cycle is defined as $\mathbb{T}_c = \{\mathcal{C}(c-1)+1, \mathcal{C}(c-1)+2, \ldots, \mathcal{C}c\}$ $(c \geq 1)$. The decision variable of green split for each phase $\pi \in \Pi_j$ at the c^{th} cycle is called $g_{\pi,c}$. To identify a flow from link a to link b belonging to phase π, we use the indicator $\mathcal{I}_{ab,\pi}$, where

$$\mathcal{I}_{ab,\pi} = \begin{cases} 1 & \text{if the flow } a \rightarrow b \text{ belongs to phase } \pi \\ 0 & \text{otherwise.} \end{cases} \tag{3.55}$$

In general, the total of green splits in any cycle does not exceed one unit, followed by

$$\sum_{\pi \in \Pi_j} g_{\pi,c} \leq 1 \tag{3.56}$$

$$g_{\pi,c} \geq G_0 \geq 0 \tag{3.57}$$

for all $c \geq 1$ where G_0 is the parameter of minimum green split for each phase at intersections.

At any time $t \in \mathbb{T}_c$, let $f_{\pi_n,t}$ denote an amount of signalling time assigned to phase π_n within time t. By definition, $0 \leq f_{\pi_n,t} \leq 1$. The green split $g_{\pi_n,c}$ becomes the average of $f_{\pi_n,t}$ in the c^{th} cycle, i.e.,

$$g_{\pi_n,c} = \frac{1}{\mathcal{C}} \sum_{t \in \mathbb{T}_c} f_{\pi_n,t}. \tag{3.58}$$

The traffic flow exiting link a in phase π_n is constrained by the allocated time $f_{\pi_n,t}$ in any time t:

$$\sum_{b \in \Upsilon_a^+} \mathcal{I}_{ab,\pi_n} f_{ab,t} \leq f_{\pi_n,t} Q_a \qquad \forall a \in \mathbb{A}, t \in \mathbb{T}. \tag{3.59}$$

Note that, the same split $f_{\pi_n,t}$ is applied to any link a having flow in phase π_n. Up to this point, we still allow the crossover of traffic among different phases.

Assume that the order of phases is predefined in the sequence of $\pi_1, \pi_2, \ldots, \pi_k$, i.e., phase π_m is performed before phase π_n ($m < n$) in any cycles. This order is guaranteed by the following constraint

$$f_{\pi_n,t} \sum_{h \in \mathbb{T}_c : h \geq t} f_{\pi_m,h} \leq 0 \qquad \forall t \in \mathbb{T}_c, \forall m < n. \tag{3.60}$$

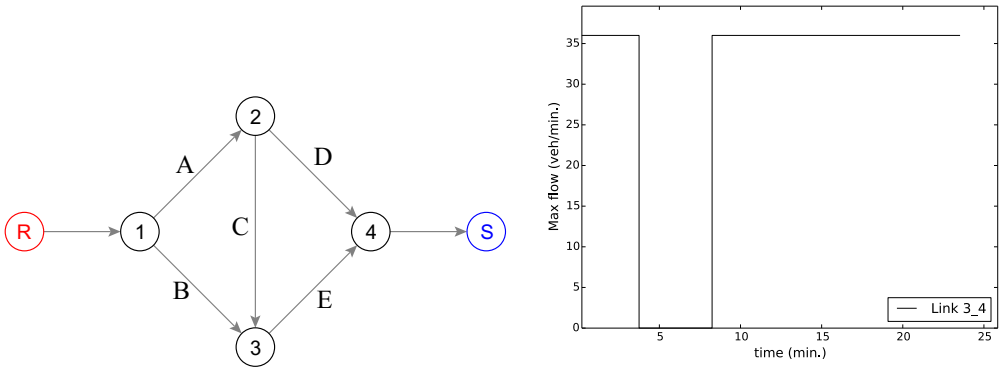
Due to the satisfaction of Eq. (3.60), if time t is used by phase π_n ($f_{\pi_n,t} > 0$), then any phase π_m ($m < n$) is not assigned to any time after t in the corresponding cycle, i.e., $f_{\pi_m,h} = 0$ for all $h \geq t$ and $h \in \mathbb{T}_c$. This rule, applied to any time t, ensures the predefined order of phases. For this, the on-and-off (or red-time) effect of traffic follows as its consequence.

In the more advanced system where we allow flexibility of phase ordering, but still require the on-and-off effect, the following constraint is used instead,

$$f_{\pi_n,t} \sum_{h \in \mathbb{T}_c : h \geq t} f_{\pi_m,h} \sum_{h \in \mathbb{T}_c : h \leq t} f_{\pi_m,h} \leq 0 \qquad \forall t \in \mathbb{T}_c, m \neq n. \tag{3.61}$$

Coming from the same idea as in Eq. (3.60), if time t is used by phase π_n ($f_{\pi_n,t} > 0$), then for any other phase π_m, its green period has to be in one of two cases: (i) finishing at or before time t, (ii) starting at or after time t. Note that, this rule is also applied to any time t, therefore it helps to separate the allocation of green period among different phases. With this constraint, it enables us to study not only signal timing but also the dynamic optimal phase arrangement in any cycles.

The continuity and convexity of Eqs. (3.60) and (3.61) are important to the solution analysis in DTA models. In the numerical results, we show their impacts on the optimal signal setting and traffic propagation.



(A) Braess network (B) Maximum flow setting

FIGURE 3.1: Braess network with the maximum flow setting at link (3,4).

3.6 Numerical results

This section provides some numerical results showing the application of our proposed approach for the SO-DTA problem in terms of accuracy and computational demand. To this end, we carry out two case studies: (1) using a simple Braess network to demonstrate in details the abilities of the TTM-DSO to avoid the queue spill-back problems, and (2) using a medium-sized grid network to compare the computational performance between TTM-DSO and CTM-DSO. For the signal control in SO-DTA model, the simple example is used to illustrate the impact of phase orders on both route choice and signal timing.

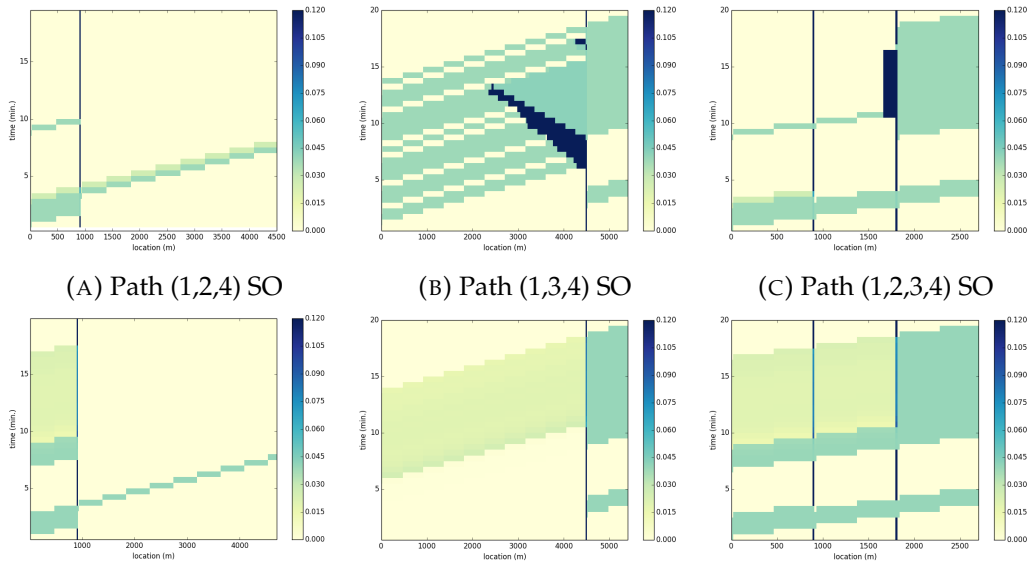
3.6.1 Case study 1: Braess network

Fig. 3.1a shows the Braess network topology with a single origin-destination ($R - S$). Table 3.1 describes the network parameters for our study. According to the topology in Fig. 3.1a, there are three paths with increasing free-flow travel time: (1,2,3,4), (1,3,4) and (1,2,4). Without disruption of link capacity, people would choose the path (1,2,3,4) for their travelling. However, since link (3,4) is closed for some period of time (meaning that traffic is not permitted to enter this link, but its downstream flow is still allowed) as described in Fig. 3.1b, it is essential to choose other paths to reach destination S as soon as possible.

Fig. 3.2 shows different DSO solutions without using the NHB constraints, including the 1-step (original) DSO solution (Figs. 3.2a–3.2c) and the proposed 2-step

TABLE 3.1: Networking parameters for the Braess network in Fig. 3.1.

Traffic parameters (applied for all links):		
$V_a = 900$ (m/min.) $W_a = 450$ (m/min.)		
$K_a = 0.12$ (veh/m)		
Demand: 72 (veh/min.), time step: 0.25 (min).		
Links	l_a (m)	Q_a (veh/min.)
$(1,2)$	900	36
$(1,3)$	4500	36
$(2,3)$	900	36
$(2,4)$	14400	36
$(3,4)$	900	See Fig. 3.1b



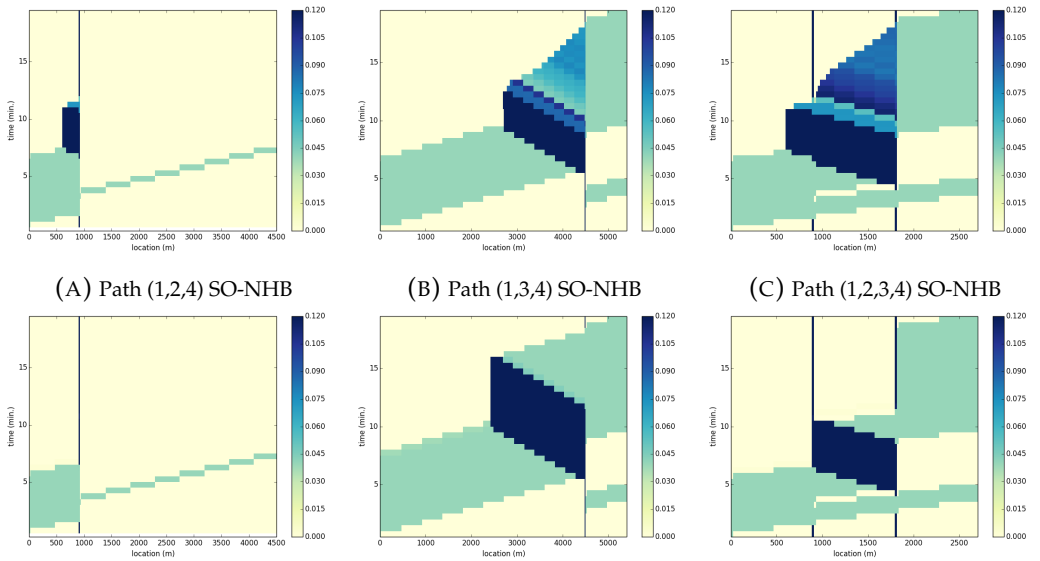
(A) Path (1,2,4) SO (B) Path (1,3,4) SO (C) Path (1,2,3,4) SO

(D) Path (1,2,4) SO-MCL (E) Path (1,3,4) SO-MCL (F) Path (1,2,3,4) SO-MCL

FIGURE 3.2: Space-time density at three paths in DSO solutions without NHB constraints.

DSO solution (Figs. 3.2d–3.2f). It is clear that the 2-step framework finds DSO solution which is free-flow in link $1-3$ and link $2-3$ at node 3 due to the capacity drop in link $3-4$. This is obtained by minimising the total queue lengths (MCL). Our 2-step DSO solution means that some vehicles are held at the source so that travellers will not have to wait in the congestion inside the network. Obviously delay at the source (i.e., stay home and start the journey later) can be much better than waiting in congestion inside the network.

Fig. 3.3 describes different DSO solutions using the NHB constraints. Since the nature of the NHB constraints is to push traffic going through the network as much



(A) Path (1,2,4) SO-NHB (B) Path (1,3,4) SO-NHB (C) Path (1,2,3,4) SO-NHB

(D) Path (1,2,4) SO-NHB-MDCL (E) Path (1,3,4) SO-NHB-MDCL (F) Path (1,2,3,4) SO-NHB-MDCL

FIGURE 3.3: Space-time density at three paths in DSO solutions with NHB constraints.

as possible so there definitely heavy congestion inside the network, as seen in Figs. 3.3a–3.3c where the 1-step framework is used. With such NHB constraints, there are queue spillbacks from link $2-3$ to link $1-2$. However, our 2-step framework finds a DSO solution which can eliminate such spillback problem by reducing the queue in link $2-3$ and increasing the queue in link $1-3$ while making link $1-2$ free-flowing (Figs. 3.3d–3.3f), which is obtained by minimising the total difference of the queue lengths (MDCL).

The difference between DSO solutions can be further seen in Fig. 3.4 describing the inflow-outflow profiles for different DSO solutions without NHB conditions. In Fig. 3.4:

- The outflow of link $R-1$ shows how travellers enter the network while the outflows of link $3-4$ and link $2-4$ show how they exit the network.
- The inflow of link $R-1$ shows the demand function which is the same in both solutions.

Fig. 3.4 clearly shows the distinct difference of the inflow-outflow profiles between DSO solutions when the congestion is inside the network (i.e., in link $1-3$ and link $2-3$) is eliminated (in the SO-MCL solution). More specially, given the same demand (i.e., the inflow of link $R-1$), SO-MCL travellers do not suffer from

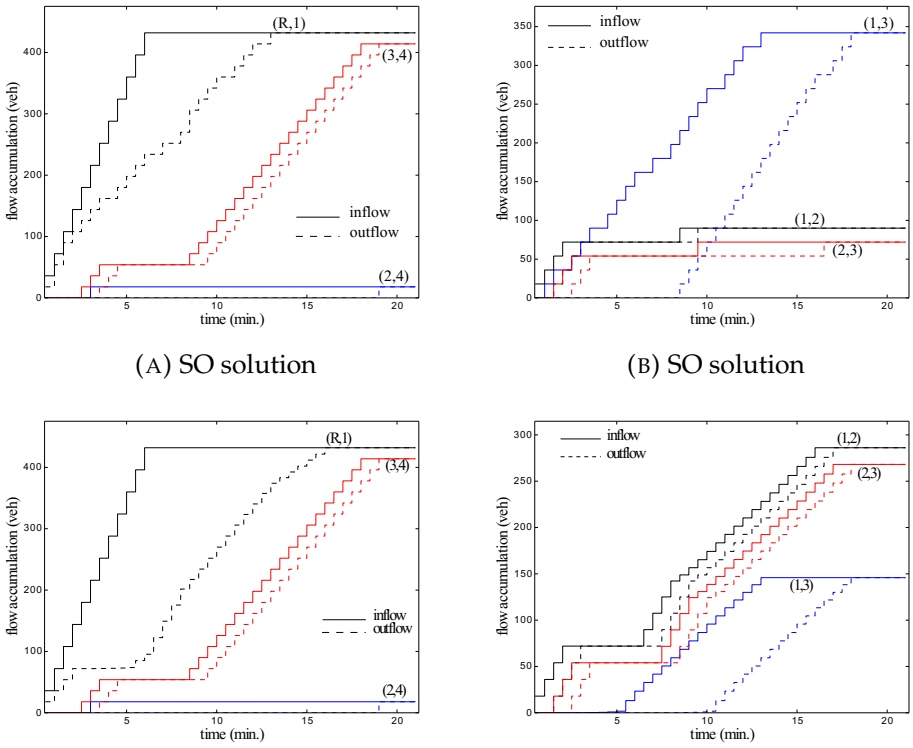


FIGURE 3.4: Inflow-outflow profiles for different DSO solutions.

the congestion in the network except the waiting time at source R, while SO travellers experience the congestion at link $1-3$ and link $2-3$.

Fig. 3.5 shows the accuracy of our approximation method for the congested length on link $2-3$. While the solid line shows the evolution of the length of the congested regime which is computed from the algorithm in Section 3.3.4, the dashed line describes the evolution of the length of the congested regime using an approximation algorithm in Section 3.3.5. Note that, while the algorithm in Section 3.3.4 is used after the DSO solution is obtained, the approximation algorithm in Section 3.3.5 is embedded in the DSO problem.

3.6.2 Case study 2: Grid network

This section compares the complexity and computational demand of the TTM-DSO against CTM-DSO in a larger network. To this end, we generate a scalable network based on grid topology in Fig. 3.6. Table 3.2 describes four network configurations for the same physical network with the only difference of time units (or the sampling

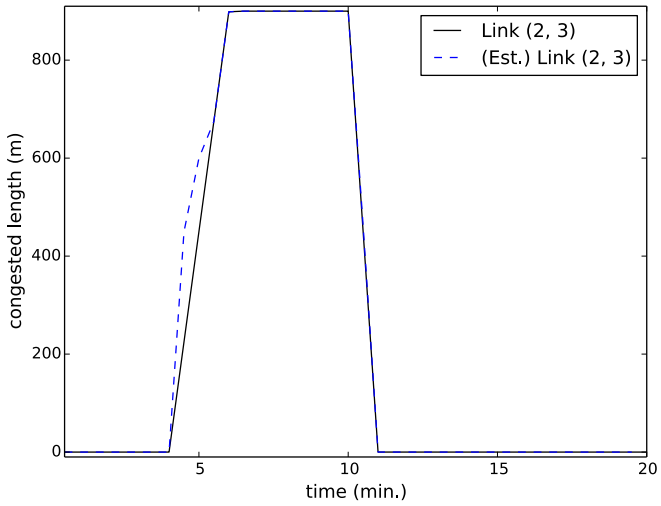


FIGURE 3.5: The evaluation of congested length in MDCL solution in Figs. 3.3d–3.3f.

Constant parameters (applied for all links):			
$V_a = 15$ (m/s)		$W_a = 7.5$ (m/s)	
$l_a = 600$ (m), $\omega = \frac{W_a}{V_a} = 0.5$		$D^{rs} = 90$ (veh)	
Time units (s): $\{2, 4, 8, 20, 40\} \approx$ Conf. $\{C20, C10, C5, C2, C1\}$			
Link types	K_a (veh/m)	Q_a (veh/sec.)	N (per cell) (veh/s)
Normal	0.12	0.6	1.8
Bottleneck	0.6	0.3	0.9

TABLE 3.2: Grid network with parameter sets for CTM-DSO and TTM-DSO in Fig. 3.6.

time intervals). Note that smaller time step results in more accuracy of traffic dynamics but requires more computational time (additional number of cells are used). The number in configuration name indicates a number of cells per link, for example, C5 produces CTM topology with 5 cells/link (equivalently 8s time step). There are two types of link: normal and bottleneck links. The number of bottleneck links accounts for 20% of the total links. We examine the impact of time and space domain on CTM-DSO and TTM-DSO models by two different scenarios, both with one O-D pair from the top-left node (R) to the bottom-right node (S). The first scenario evaluates all configurations, where the time domain also changes to produce equivalent results. The second scenario focuses on some specific configurations with varying sizes of the time domain.

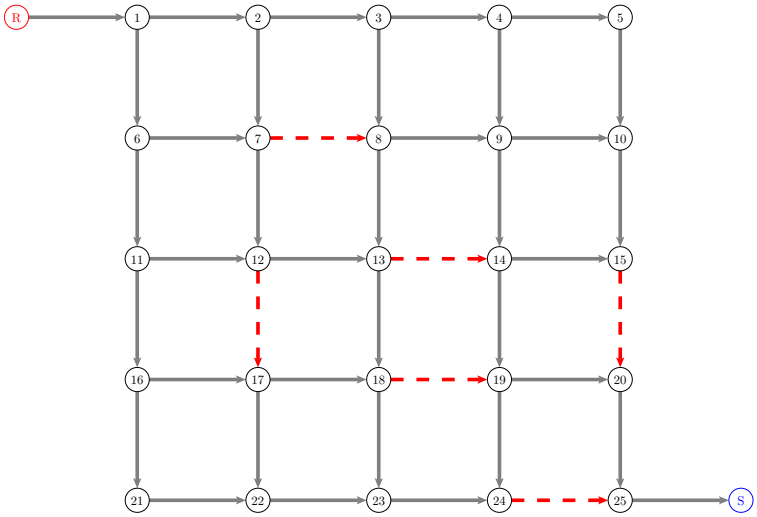


FIGURE 3.6: Grid network topology for TTM-DSO (the dashed lines show bottleneck links).

Estimation of complexity

We consider complexity of models in three aspects: (i) number of constraints ($\mathcal{C}$), (ii) number of variables ($\mathcal{V}$), and (iii) computational time ($\mathcal{T}$). For TTM-DSO: $\mathcal{C}_{TTM} = O(|\mathbb{E}.\mathbb{T}|), \mathcal{V}_{TTM} = O(|\mathbb{E}.\mathbb{T}|)$. For CTM-DSO: $\mathcal{C}_{CTM} = O(|\mathbb{C}.\mathbb{T}|), \mathcal{V}_{CTM} = O(|\mathbb{C}.\mathbb{T}|)$, where $\mathbb{C}$ is the set of all cells in the CTM-DSO network.

In TTM, the variables are the sets of $u_a(i), v_a(i)$ and $f_{ab}(i), \forall a, b \in \mathbb{A}, i \in \mathbb{T}$. In our grid example, apart from boundary nodes, all nodes have two incoming links and two outgoing links. For this reason, the number of inflow-outflow combinations is about $2|\mathbb{A}|$. From that, the estimation of $\mathcal{V}_{TTM}$ is approximately $4|\mathbb{E}\mathbb{T}|$. The number of TTM-DSO constraints is $\mathcal{C}_{TTM} \approx 8|\mathbb{E}\mathbb{T}|$.

To estimate the complexity of CTM-DSO, we first take a look at its main constraints:

$$n_i(t+1) = n_i(t) + \sum_{ji \in \mathbb{A}} y_{ji}(t) - \sum_{ij \in \mathbb{A}} y_{ij}(t)$$

$$\sum_{ji \in \mathbb{A}} y_{ji}(t) \leq \min\{Q_i(t), \omega_i(N_i(t) - n_i(t))\}$$

$$\sum_{ij \in \mathbb{A}} y_{ij}(t) \leq \min\{Q_i(t), n_i(t)\}$$

for all $i \in \mathbb{C}, t \in \mathbb{T}$, where $\mathbb{A}$ is the set of links in the cell network. The number of constraints is $\mathcal{C}_{CTM} \approx 5|\mathbb{C}\mathbb{T}|$. To evaluate $\mathcal{V}_{CTM}$, we observer that the number

TABLE 3.3: Estimation of complexity in terms of constraints and variables (grid topology)

$\|\mathbb{A}\| = 40$ (grid network in Fig. 3.6) k cells per link		
	$\mathcal{C}$	$\mathcal{V}$
TTM	$8\|\mathbb{ET}\| = 320\|\mathbb{T}\|$	$4\|\mathbb{ET}\| = 160\|\mathbb{T}\|$
CTM	$5k\|\mathbb{ET}\| = 200k\|\mathbb{T}\|$	$(2k+1)\|\mathbb{ET}\| = (80k+40)\|\mathbb{T}\|$

of $n_i(t)$ is equal to $|\mathbb{CT}|$, while the number of $y_{ij}(t)$ is equal to $|\mathbb{AT}|$. Assume that each link is divided to k cells, then $|\mathbb{C}| \approx k|\mathbb{A}|$. In the grid topology, the estimation of the number of arcs in cell network is $|\mathbb{A}| \approx (k+1)|\mathbb{A}|$ ($|\mathbb{A}|$ is large enough). This analysis leads to the evaluation of $\mathcal{V}_{CTM} \approx (2k+1)|\mathbb{ET}|$. These estimations, summarised in Table 3.3, help us to have a quick calculation of the complexity of both models. We expect a good performance of TTM-DSO for any $k \geq 2$.

Preparation time

It is the amount of time necessarily for creating a real model of grid network (case study 2) which is solved by the CBC solver. Since all models are implemented in an abstract language, like AMPL, GAMS or modern programming languages (for example Python), this time is necessary to transform these abstract models into concrete models, which are understood by solvers. It includes processing data input file and generating an instance of the abstract model with variables, constraints, and parameters. The cost of reading input file is really small because we optimise these files by reducing duplicated values, and no file exceeds 2 MB in our experiments.

Comparison results

We compare the performance of TTM-DSO and CTM-DSO with respect to the complexity and preparation time defined above in two following scenarios:

Scenario 1: Our target is to find the (system optimal) traffic pattern for a fixed demand (90 vehicles) from R to S. In our experiments, all demands need around 400s to travel in the grid. Obviously, decreasing time step leads to the increasing computational time by increasing the time domain for each configuration, see Table 3.4. Especially, for CTM, not only time but also space domain are increased. We use two

TABLE 3.4: Performance comparison between TTM (gray column) and CTM in different traffic details.

Conf.	$\|\mathbb{T}\|$	$\mathcal{C}$		$\mathcal{V}$		$\mathcal{T}_{prepare}$ (s)		$\mathcal{T}_{solve}(s)$		Obj/$\|\mathbb{T}\|$ (veh)
C2	28	9104	11690	4204	5455	0.75	2.14	0.57	0.57	41.8
C5	70	22754	71216	10504	30391	2.90	45.59	2.20	5.33	43.7
C10	140	45504	282426	21004	116751	9.60	1186.34	8.00	50.12	44.4
C20	280	91004	1124846	42004	457471	33.61	24297.18	27.48	693.44	44.7

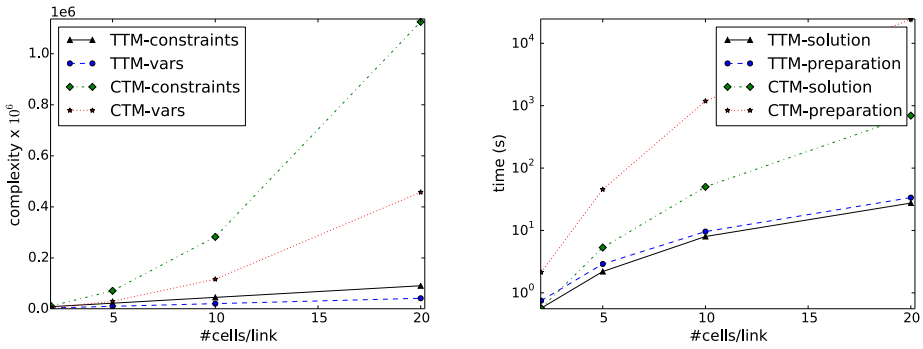


FIGURE 3.7: Performance comparison between TTM and CTM in different traffic details

cell per link in $C2$ up to 20 cells per link in $C20$. While TTM-DSO model changes only link parameters in the new configurations, CTM-DSO model changes the whole network topology including the cell parameters. The results are shown in Table 3.4 and Fig. 3.7. It is worth noticing that both models produce the same objective values in any configurations (the last column in Table 3.4).

For complexity, the result is consistent with the estimation in Table 3.3. We observe that the number of constraints and variables are increased linearly in TTM-DSO, while they are increased polynomially in CTM-DSO. Particularly, in the most complex configuration $C20$, TTM-DSO solves the problem 20 times faster than CTM-DSO. Looking at the trend of all complexity aspects in Fig. 3.7, CTM-DSO complexity is always increased at a higher rate than TTM-DSO, which means that TTM-DSO can provide more accurate traffic dynamics pattern at the equivalent cost.

Scenario 2: Instead of increasing the complexity in both time and space as in the first scenario, we use the same configuration for all test cases to ignore the impact of space and explore the influence of time domain on these models. Because of this setting, both $\mathcal{C}$ and $\mathcal{V}$ of TTM-DSO and CTM-DSO are increased linearly in Fig. 3.8. We use the most complex configuration, $C20$, with $|\mathbb{T}|$ from 160 to 300. Since the

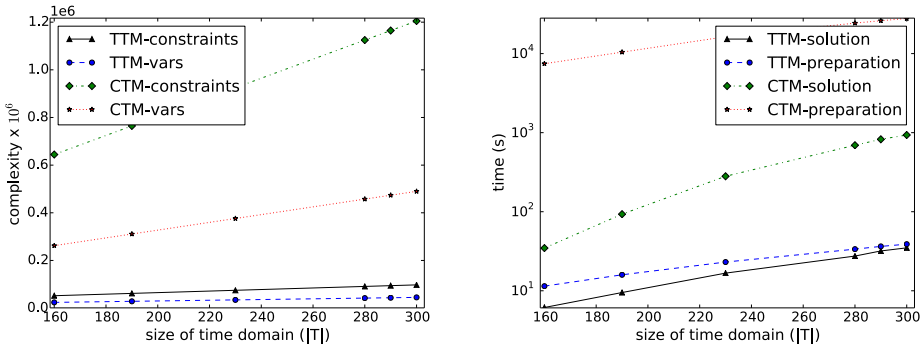


FIGURE 3.8: Performance comparison between TTM and CTM with different size of time domain (Conf. C20 $\sim$ 20 cells/link)

fastest travel time is more than 320s, there is no movement of traffic with $|\mathbb{T}|$ less than 160. With the initial demand of 90 vehicles, all travellers will finish their trips in about 280 time-slots in $C20$ (9 minutes equivalently). From Fig. 3.8, TTM-DSO outperforms CTM-DSO where all criteria of complexity show its benefit, especially, their increasing rates are also smaller.

In both scenarios, the preparation time is described in Figs. 3.7 and 3.8. It shows that TTM-DSO always finishes the preparation period quickly. On the other hand, CTM-DSO preparation time dominates the solution time in all cases. Although we only verify this observation in our programming environment, it shows that TTM-DSO outperforms CTM-DSO in terms of efficiency.

3.6.3 Numerical example of signal control model

In this section, we study a simple network with only one intersection and two paths, shown in Fig. 3.9. The network configuration for each link is listed in Table 3.5. Travelers start from source node R and move on two paths to the intersection J before reaching the destination S. Two phases, called R-J and A-J, are controlled by the traffic signal at J, with the minimum green split $G_0 = 0.25$ and 60-second cycle time. In the example, we consider two possibilities of phase ordering, named 'R-J first' and 'A-J first' to indicate the first phase of each cycle. Demand profile is shown in Fig. 3.10.

In this example, we also compare our solutions with the continuum approximation of signal control model based on Han et al. (2014). Note that, our model tends to split a signal cycle in time (as it happens in practice) by separating the phases

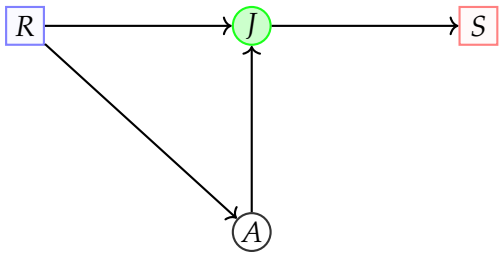


FIGURE 3.9: Two-path-single-intersection topology.

TABLE 3.5: Configuration for each link in Fig. 3.9.

L (m)	**Q** (veh/s)	**V** (m/s)	**W** (m/s)	**K** (veh/m)
450	0.6	15	7.5	0.12

in each cycle while the approximation applies green split to flow capacity without considering the phase ordering (therefore there are no explicit constraints to avoid the cross-over of traffic in the solution of approximation).

Different solutions due to phase ordering

Table 3.6 provides an overview of three solutions in this example. Due to the pre-defined phase ordering, our solutions for R-J first and A-J first are different in all aspects, including the objective value, average travel time and green splits. In A-J first, traffic splits equally into two paths (Fig. 3.11) while major traffic travels on the longer path R-A-J-S in R-J first (Fig. 3.12).

According to the network settings, traffic on link R-J will arrive at the intersection J in the middle of the first cycle, and traffic on link R-A-J will arrive J in the next cycle. Because each phase is preserved at least 25% of the cycle time (even there is

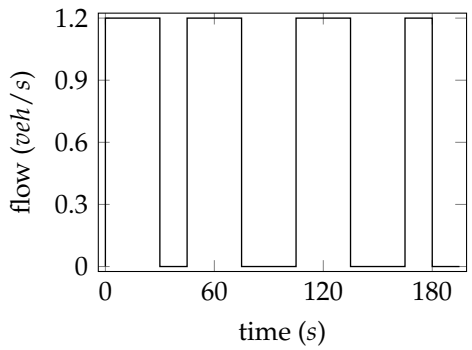


FIGURE 3.10: The O-D demand profile for network in Fig. 3.9.

TABLE 3.6: Solution analysis

Solution	**Obj. value** (veh)	**Avg. travel time** (s)	**Green split** $g_{R-J} : g_{A-J}$
A-J first	1197.00	98.9	1:1
R-J first	1080.00	112.9	1:3
Approximation	1136.25	106.2	1:3

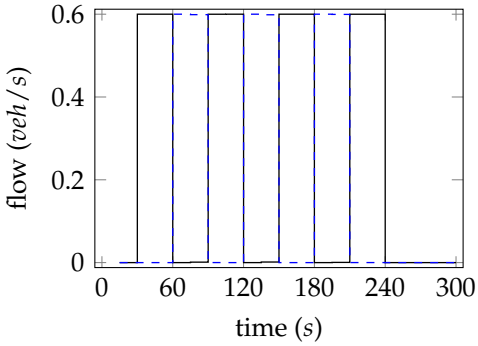


FIGURE 3.11: The downstream flow on link R-J (smooth line) and A-J (dashed line) (in A-J first solution).

no traffic), letting phase A-J performs first is the best option for traffic on link R-J to fully use the remaining time in the first cycle (time 0 to 60, Fig. 3.11). Otherwise, if phase R-J is considered first, the last 15 seconds, preserved for phase A-J, in the first cycle (see Fig. 3.12) is wasted due to no arrival during this period. With the difference in the initial cycle, traffic spreads differently in the network. The result in Table 3.6 also shows that we should choose A-J first for the best optimal travel time.

Furthermore, the on-and-off effect is certainly represented in the Figs. 3.11 and 3.12, i.e., at any time, positive traffic in a phase causes zero traffic in other. This feature distinguishes our model from the approximation presented below.

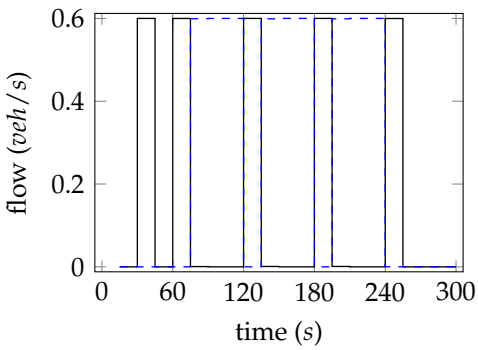


FIGURE 3.12: The downstream flow on link R-J (smooth line) and A-J (dashed line) (in R-J first solution).

Comparison with the approximation

Applying the same network and inputs, the overall result for the approximation of signal control is shown in Table 3.6. In particular, the objective value and average travel time of the approximation are in between our solutions, however, the green split is similar to R-J first.

Without information of phase ordering, the traffic pattern on two links R-J and A-J, shown in Fig. 3.13, could be seen as the average flow over the corresponding cycle, and it is impossible to show the on-and-off effect as in our model. The green split could be observed vertically at any time in Fig. 3.13 (and equal to the fraction of flow at that time over the maximum flow capacity) while it is the fraction of signalling time over the cycle time in our model (observed horizontally). Therefore, approximation method yields in the first cycle, $g_{R-J} = 0.75$ and $g_{A-J} = 0.25$, then $g_{R-J} = 0.25$ and $g_{A-J} = 0.75$ for any later cycle time.

Although the approximated green split looks similar to the case R-J first, its traffic pattern is either impossible or non-optimal in the view of phases. Particularly, if phase A-J is considered first, then the throughput of the approximated solution in the first cycle is not optimal. The reason is that 25% of the flow capacity, reserved for flows from link A-J, is unused. Otherwise, if phase R-J is greened first, it is impossible to send as much traffic as in the approximated solution.

In summary, the existing approximation method can provide the green splits, but it does not guarantee the existence, nor does it reveal the traffic patterns and the corresponding order of phases of the obtained solution. In contrast, our proposed framework provides both the optimal green splits and the exact traffic patterns associating with a particular order of phases.

3.7 Summary

This chapter proposes a novel optimisation framework that contains two steps: i) find DSO solutions via minimising the total travel times or the total system cost, ii) find traffic patterns which optimally distribute the congestion over links in the network. To fulfil Step 2, there is a need to determine the time and (horizontal) space evolution of the queues in the network. To this end, we have formulated the Two

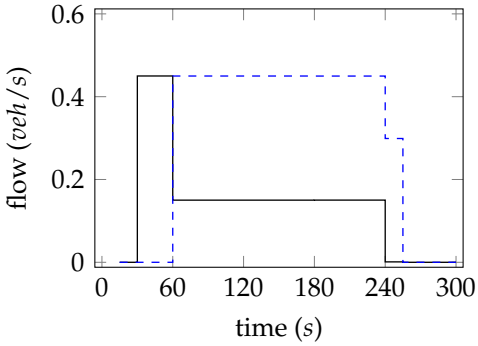


FIGURE 3.13: The downstream flow on link R-J (smooth line) and A-J (dashed line) (approximation).

regime Transmission Model (TTM)-an extended version of the Link Transmission Model (LTM)-as a set of side linear constraints for a DSO problem. Basically, TTM models the dynamics of the two regimes of traffic flow namely, non-congested and congested states using a simplified Newell's theory as LTM does and describes the traffic states in time and space using just two regimes of the link as opposed to denser spatial discretisation adopted in CTM. In the proposed optimisation framework, the TTM utilises the same variables at the two ends of the link as the LTM (i.e., the entry and exit flow) in the DSO. However, the evolution of the length of the congested regime within link tracked by the TTM is embedded in Step 2.

We have carried out case studies to illustrate the performance of our framework. On the one hand, they show that the proposed framework can optimally distribute the congestion over links in the network to achieve minimal spillbacks or free-flow situations. On the other hand, we have shown that the complexity of the TTM-DSO including the number of the side constraints, the number of the variables as well as the computational time is only increased linearly with time steps. The numerical results have indicated that the TTM-DSO model outperforms the CTM-DSO model in terms of the complexity.

This chapter also presents a convex signal control model, embedded in the TTM-based SO-DTA model, that helps us to study the optimal signal settings and route choices. The numerical results show the clear impact of phase ordering to both route choice and signal timing.

Chapter 4

An UE-DTA model for single O-D capacitated networks

The dynamic traffic assignment (DTA) problem has been studied intensively in the literature. However, there is no existing linear framework to solve the user equilibrium (UE) DTA problem. This chapter aims to fill this gap by studying the objective-based mathematical programming framework for UE-DTA problem in a single O-D network. It provides the analytical study that enables the development of incremental solution method by utilising the relation between UE and SO solutions. In contrast to the related iterative methods, such as Frank-Wolfe algorithm, successive average (MSA) or projection and their extensions where the purpose of iteration is to seek solution convergence, whereas ours is to solve a linear problem over multiple iterations but only for a single unit of demand in each iteration.

The organization of this chapter follows. Section 4.1 formulates the route choice problem in the dynamic traffic assignment framework. In Section 4.2, we show the relation between UE and SO problems which is utilized later in Section 4.3. We develop an incremental solution method in Section 4.3 to solve the UE-DTA problem, based on the study of UE objective in the previous section. Section 4.4 illustrates our method through two examples: the first simple network for the verification of results and the second medium-size network for the evaluation of computational performance. Finally, we conclude this chapter in Section 4.5.

4.1 Problem formulation

In this section, we discuss our LP formulation of route choice problem in the dynamic traffic assignment framework. We consider a road network with a single origin-destination (O-D) pair where a number of selfish but rational users travel from the origin to the destination. We shall first introduce the notations that will be used in this chapter. We then present the formulation of link transmission models (LTM), especially for the link based and path based models. The FIFO constraint and estimation of individual travel time will be discussed and derived. Finally, the complete formulation will be presented.

4.1.1 Notations

A transportation network is generally represented by a directed graph $\mathcal{G} = (\mathbb{V}, \mathbb{A})$, where $\mathbb{V}$ is the set of nodes, and $\mathbb{A}$ is the set of directed arcs connecting certain pairs of nodes. Without losing generality, there are three disjoint sets: source nodes $\mathbb{V}_R$, sink nodes $\mathbb{V}_S$, and intermediate nodes $\mathbb{V}_I$. In single-destination network, $\mathbb{V}_S = \{s\}$. Furthermore, each source node only connects to one intermediate node, and there is no incoming link to any sources. Similarly, each sink node only receives flow from one intermediate node, and there is no outgoing link from any sinks. An intermediate node could have any number of links coming to or going out of them.

Let $\mathbb{A}_R$ denote the set of arcs from sources, $\mathbb{A}_S$ denote the set of arcs to sinks, and $\mathbb{A}_I$ denote the remaining arcs. Note that, the arcs in $\mathbb{A}_R$ and $\mathbb{A}_S$ are the virtual links that store traffic at sources and destinations to maintain the rule of flow conservation. The set of all possible paths, connecting source r and destination s in this network, is denoted by $\mathbb{P}^{(rs)}$. Let $\mathbb{T}$ denote a set of discrete times where T is the maximum time horizon, i.e., $\mathbb{T} = \{1, 2, \ldots, T\}$.

Let Υ_a^- denote the set of inflow links to link a, and Υ_a^+ denote the set of outflow links from link a for any $a \in \mathbb{A}$. For any network topology in our work, we have $\Upsilon_a^- = \varnothing \; \forall a \in \mathbb{A}_R$, and $\Upsilon_a^+ = \varnothing \; \forall a \in \mathbb{A}_S$.

4.1.2 The link-transmission-based models

In this section, we describe a discrete-time kinematic-wave-based link transmission model (LTM) (Yperman, 2007) for the network loading as a set of side constraints in a DTA problem.

Link parameters For each link $a \in \mathbb{A}$, the following parameters are defined:

K_a : jam density (a number of vehicles per distance unit) at link a.

V_a : unsaturated speed in link a.

W_a : backward speed in link a.

$Q_{a,t}$: maximum number of vehicles to or out of link a at time t.

L_a : length of link a.

The parameters (K_a, V_a, W_a) helps to describe the triangular fundamental relationship between flow and density in each link. For sources and sinks, they have infinite capacity: $Q_{a,t} \to \infty \quad \forall a \in \mathbb{A}_S \cup \mathbb{A}_R, t \in \mathbb{T}$. Furthermore, we also assume integer values of $\frac{L_a}{V_a}$ and $\frac{L_a}{W_a}$ to simplify the model description. For O-D demand, let $D_t^{(rs)}$ denote an amount of traffic demand departing at time t from source r to sink s.

Link variables To capture the dynamic traffic in each link $a \in \mathbb{A}$, the following variables are defined:

$u_{a,t}$: a number of vehicles moving to link a at time t.

$v_{a,t}$: a number of vehicles moving out of link a at time t.

$f_{ab,t}$: a number of vehicles moving from link a to link b at time t.

The link-based traffic model

The link-based traffic model will consist of link constraints and node constraints as below.

Link constraints: For each link $a \in \mathbb{A}$, the demand capacity $\mathcal{D}_{a,t}$ defines a maximum amount of traffic exiting link a at time t, and the supply capacity $\mathcal{S}_{a,t}$ defines a maximum amount of traffic entering link a at time t. Their mathematical definitions are follows.

Source links ($a \in \mathbb{A}_R$) At source links, there is no limitation on demand and supply capacity.

$$\mathcal{D}_{a,t} \to \infty \tag{4.1}$$

$$\mathcal{S}_{a,t} \to \infty. \tag{4.2}$$

Sink links ($a \in \mathbb{A}_S$) At sink links, there is infinite supply but zero demand capacity to store all traffic arriving the destination.

$$\mathcal{D}_{a,t} = 0 \tag{4.3}$$

$$\mathcal{S}_{a,t} \to \infty. \tag{4.4}$$

Other links ($a \in \mathbb{A}_I$) The below equations formulate $\mathcal{D}_{a,t}$ and $\mathcal{S}_{a,t}$ which are dynamically depended on the upstream and downstream flows as follows,

$$\mathcal{D}_{a,t} = \min \left\{ Q_{a,t}; \sum_{h \le t - \frac{L_a}{V_a}} u_{a,h} - \sum_{h \le t-1} v_{a,h} \right\} \tag{4.5}$$

$$\mathcal{S}_{a,t} = \min \left\{ Q_{a,t}; K_a L_a + \sum_{h \le t - \frac{L_a}{W_a}} v_{a,h} - \sum_{h \le t-1} u_{a,h} \right\}. \tag{4.6}$$

In Eq. (4.5), the second part shows the remaining traffic at downstream point that is ready to exit link a at time t. Note that, due to the free-flow travel time $\frac{L_a}{V_a}$ on link a, any traffic entering link a after time $\left(t - \frac{L_a}{V_a}\right)$ is still travelling on this link, and will exit link a after time t. Similarly, for the supply capacity $\mathcal{S}_{a,t}$ at the upstream point of link a, it shows the vacancies made by the outgoing flow (equal to $\sum_{h \le t - \frac{L_a}{W_a}} v_{a,h}$) and the entering flow (equal to $\sum_{h \le t-1} u_{a,h}$) of link a. These supply and demand capacity in each link are used in node model below to restrict the upstream and downstream flow at each node.

Node constraints: For any node in the network, the node constraints guarantee the flow conservation and restriction of demand and supply.

$$u_{a,t} = \sum_{b \in \Upsilon_a^-} f_{ba,t} \qquad \forall a \in \mathbb{A}, t \in \mathbb{T} \tag{4.7}$$

$$v_{a,t} = \sum_{b \in \Upsilon_a^+} f_{ab,t} \qquad \forall a \in \mathbb{A}, t \in \mathbb{T} \tag{4.8}$$

$$u_{a,t} \leq \mathcal{S}_{a,t} \qquad \forall a \in \mathbb{A}, t \in \mathbb{T} \tag{4.9}$$

$$v_{a,t} \leq \mathcal{D}_{a,t} \qquad \forall a \in \mathbb{A}, t \in \mathbb{T} \tag{4.10}$$

$$f_{ab,t} \geq 0 \qquad \forall a, b \in \mathbb{A}, t \in \mathbb{T} : b \in \Upsilon_a^+. \tag{4.11}$$

For source nodes, the upstream flow is given and identical to the demand at time t, i.e.,

$$u_{r,t} = \sum_{s \in \mathbb{A}_S} D_t^{(rs)} \qquad \forall r \in \mathbb{A}_R, t \in \mathbb{T}. \tag{4.12}$$

The above constraints, i.e., Eqs. (4.1)–(4.12), formulate the discrete-time link-based network loading model for a single-destination network. In the next section, we describe the path-based model to track traffic flows on any paths.

The path-based traffic model

To trace flows on any paths, let $f_{a,p,t}$ represent the flow on path p arriving the upstream of link a at time t, and let $f_{p,t}$ denote an amount of traffic departing from source r at time t and choosing path $p \in \mathbb{P}^{(rs)}$. In the relation to the link variables, the definite constraints follow:

$$u_{a,t} = \sum_{p \in \mathbb{P}: a \in p} f_{a,p,t} \qquad \forall a \in \mathbb{A}, t \in \mathbb{T} \tag{4.13}$$

$$v_{a,t} = \sum_{b \in \Upsilon_a^+} \sum_{p \in \mathbb{P}: a,b \in p} f_{b,p,t} \qquad \forall a \in \mathbb{A}, t \in \mathbb{T} \tag{4.14}$$

$$f_{ab,t} = \sum_{p \in \mathbb{P}: a,b \in p} f_{b,p,t} \qquad \forall a, b \in \mathbb{A}, t \in \mathbb{T} \tag{4.15}$$

$$f_{p,t} = f_{r,p,t} \qquad \forall p \in \mathbb{P}, t \in \mathbb{T}. \tag{4.16}$$

In addition to the above equations, all link and node constraints in Section 4.1.2 are reused in the path-based model. In fact, with the detail of traffic in a path, the

constraint Eq. (4.12) is replaced by the following equation,

$$\sum_{p \in \mathbb{P}^{(rs)}} f_{p,t} = D_t^{(rs)}. \tag{4.17}$$

Similar to the demand capacity on link a due to the free-flow travel time L_a/V_a (see Eq. (4.5)), the additional constraint of free-moving path flows is included, i.e.,

$$\sum_{h \le t} f_{b,p,h} - \sum_{h \le t - \frac{L_a}{V_a}} f_{a,p,h} \le 0 \qquad \forall a \in \mathbb{A}_I, p \in \mathbb{P}, t \in \mathbb{T} : b \in \Upsilon_a^+, b \in p. \tag{4.18}$$

The First-In-First-Out (FIFO) constraints

In the link-based traffic model in Section 4.1.2, the FIFO principle is satisfied on any links (Yperman, 2007). However, by using the path-based model, it is required to represent FIFO among path flows in explicit constraints. The principle of FIFO describes an order of movement where a flow departs earlier at the origin would arrive sooner at the destination. In this section, we describe three types of FIFO, i.e., FIFO over links, paths and O-D pairs.

FIFO over links: Let $f_{a,p,t,h}$ denote an amount of traffic flow that enters link a on path p at time t and goes out of this link at time h. The FIFO condition on link a, based on Carey (1992), follows

$$\sum_{p \in \mathbb{P}} f_{a,p,t^+,h} \sum_{p \in \mathbb{P}} f_{a,p,t,h^+} \le 0 \qquad \forall a \in \mathbb{A}, \quad \forall h^+ > h, \quad \forall t^+ > t. \tag{4.19}$$

By Eq. (4.19), the later arriving flow $f_{a,p,t^+,h}$ (at time $t^+ > t$) could exit link a at time h only when there is no flow f_{a,p,t,h^+} that arrives sooner at time t and exits link a after time h. The relation between $f_{a,p,t,h}$ and $f_{a,p,t}$ (and to other introduced variables) reads,

$$f_{a,p,t} = \sum_{h \ge t+1} f_{a,p,t,h} \quad \forall a \in \mathbb{A}, p \in \mathbb{P}, t \in \mathbb{T}. \tag{4.20}$$

FIFO over paths: Let $f_{p,t,h}$ denote an amount of traffic flow that departs from the source on path p at time t and arrives the destination at time h. Similarly, the FIFO

condition over path p follows,

$$f_{p,t^+,h}\, f_{p,t,h^+} \leq 0 \qquad \forall p \in \mathbb{P}, \quad \forall t^+ > t, \quad \forall h^+ > h. \tag{4.21}$$

In Lo et al. (2002), the authors prove that FIFO over links lead to FIFO over paths.

FIFO over O-D: For a given O-D pair, the order of flows departing at a source is hold at a destination.

$$\Big(\sum_{p\in\mathbb{P}} f_{p,t^+,h}\Big)\Big(\sum_{p\in\mathbb{P}}\sum_{k\geq h+1} f_{p,t,k}\Big) \leq 0 \qquad \forall t, t^+, h \in \mathbb{T}, t^+ > t. \tag{4.22}$$

Note that, the FIFO over O-D (also known as O-D FIFO) is not necessary to be satisfied in a general DTA problem. For example, in SO solution that optimizes the total travel cost, the flow entering the network sooner might be directed (by the system operator) on a longer path, therefore it arrives the destination later than other flows. However, in the UE-DTA problem, FIFO over paths implies FIFO over O-D (Lo et al., 2002; Carey et al., 2014) and thus the O-D FIFO is a property of the UE solution.

4.1.3 Travel time estimation

In this chapter, we take *the travel time to be the travel cost* and consider two types of estimation: the marginal and average travel time. Let $f_{p,t,h}$ denote a partial flow of $f_{p,t}$ (the O-D pair $r, s \in p$), arriving the destination s at time h ($h > t$), where

$$f_{p,t} = \sum_{h\geq t} f_{p,t,h}. \tag{4.23}$$

We assume that, each flow has to take at least one time unit from source to destination, i.e., $f_{p,t,h} = 0$ for all $h \leq t$.

The marginal travel time function

For a given path flow $f_{p,t}$, the marginal travel time function $\tilde{T}_{p,t}(u)$ is the marginal travel time for an amount u of traffic belonging to $f_{p,t}$, formulated as

$$\tilde{T}_{p,t}(u) = h - t \qquad \text{if } u \in \left(\sum_{k\leq h-1} f_{p,t,k}, \sum_{k\leq h} f_{p,t,k}\right] \tag{4.24}$$

for all $u \in (0, f_{p,t}]$. We call $\tilde{T}_{p,t}$ the marginal travel time for the path flow $f_{p,t}$, where $\tilde{T}_{p,t} = \tilde{T}_{p,t}(f_{p,t})$.

The average travel time function

For a given path flow $f_{p,t}$, the average travel time function $T_{p,t}(u)$ is defined as the average of travel time over an amount u of flow belonging to $f_{p,t}$, i.e.,

$$T_{p,t}(u) = \frac{\int_0^u \tilde{T}_{p,t}(u)\, \mathrm{d}u}{u} \tag{4.25}$$

for all $u \in (0, f_{p,t}]$. The average travel time $T_{p,t}$ for the path flow $f_{p,t}$, averaged over all flow departing at time t on path p according to demand $D_t^{(rs)}$, is formulated as follows,

$$T_{p,t} = T_{p,t}(f_{p,t}) = \sum_{h \geq t} \frac{(h-t) f_{p,t,h}}{f_{p,t}}. \tag{4.26}$$

4.1.4 The overall model

Below we formulate the UE-DTA problem with the capacitated dynamic user equilibrium condition. We have previously defined link-based and path-based variables in the LTM-based models, however, they are inseparable and thus we cannot identify the impact of a path flow $f_{p,t}$ on the link flow $f_{ab,t}$. Hence, it is difficult to analyse these models under the Karush-Kuhn-Tucker (KKT) conditions for the optimal solution. For this reason, we introduce the new concept of trip flow below.

Trips and trip flow variables During a trip from source to destination, each vehicle travels through several links at different times. To differentiate trips in time and space, we define a *trip* φ as a collection of all entering times and the associated arrival links for an individual (or a group of travelers) during their movement from a source to a destination, i.e.,

$$\varphi = \{(h, a) : h \in \mathbb{T}, a \in \mathbb{A}\}. \tag{4.27}$$

Each element $(h, a) \in \varphi$ means that the flow following trip φ enters link a at time h. For traffic belonging to demand $D_t^{(rs)}$ and travelling on path p, let t_a denote the

arrival time to any link $a \in p$, then the trip $\varphi = \{(t_a, a) : \forall a \in p \text{ where } t_r = t, t_s > t\}$. The amount of flow following trip φ is called f_φ. The most important property of trip flow is *separability*, that any individual flow (or infinitesimal flow) has to belong to one and only one trip flow. A set of all (possible) trips is called Φ. Furthermore, for convenience, let Φ_p^t denote the set of trip flows belonging to demand $D_t^{(rs)}$ and following path p ($r, s \in p$), and let $\Phi_{a,t}$ denote the set of trip flows entering link a at time t. Their mathematical definitions are follows,

$$\Phi_{a,t} = \{\varphi \mid \forall \varphi \in \Phi : (t, a) \in \varphi\}$$
$$\Phi_p^t = \{\varphi \mid \forall \varphi \in \Phi, \forall (t, r), (h, a) \in \varphi : r \in \mathbb{A}_R, a \in \mathbb{A}, h > t, a \text{ and } r \in p\}.$$

A set of all trip flows is called $\boldsymbol{f} = \{f_\varphi : \forall \varphi \in \Phi\}$. As a result, other flow variables are constructed from trip flows as shown below,

$$f_{a,p,t,h} = \sum_{k \le t} \sum_{\varphi \in \Phi_{a,t} \cap \Phi_{b,h} \cap \Phi_p^k} f_\varphi \qquad \forall a \in \mathbb{A}, p \in \mathbb{P}; t, h \in \mathbb{T} : b \in \Upsilon_a^+, b \in p \tag{4.28}$$

$$f_{p,t,h} = \sum_{\varphi \in \Phi_p^t \bigcap \Phi_{s,h}} f_\varphi \qquad \forall p \in \mathbb{P}; t, h \in \mathbb{T} \tag{4.29}$$

$$f_{p,t} = \sum_{\varphi \in \Phi_p^t} f_\varphi \qquad \forall p \in \mathbb{P}, t \in \mathbb{T} \tag{4.30}$$

$$f_{ab,t} = \sum_{h < t} \sum_{\varphi \in \Phi_{b,t} \bigcap \Phi_{a,h}} f_\varphi \qquad \forall a, b \in \mathbb{A}, t \in \mathbb{T} \tag{4.31}$$

$$u_{a,t} = \sum_{\varphi \in \Phi_{a,t}} f_\varphi \qquad \forall a \in \mathbb{A}, t \in \mathbb{T} \tag{4.32}$$

$$v_{a,t} = \sum_{b \in \Upsilon_a^+} \sum_{\varphi \in \Phi_{b,t}} f_\varphi \qquad \forall a \in \mathbb{A}, t \in \mathbb{T}. \tag{4.33}$$

One might relate the concept of trip flow to platoon, however, they are different in the aspect of dispersion, i.e., a platoon is allowed to break up while trip flow is not. In fact, trip flow is seen as a part of the platoon that the whole group of travellers stays together in a whole trip. Note that, it is unnecessary to implement trip flow variables in any DTA models, but only in some aggregated variables, for example, the link-based model introduced in Section 4.1.2. However, trip flow variables are useful to analyse the DTA models in the rest of this chapter because of its separability.

The UE-DTA problem

The general side constrained DTA model is defined as the optimization problem $\mathcal{P}_0$ below:

$$[\mathcal{P}_0] \qquad \min_{\boldsymbol{f}} F(\boldsymbol{f}) = \min_{\boldsymbol{f}} \sum_{p\in\mathbb{P}}\sum_{t\in\mathbb{T}} \int_0^{f_{p,t}} T_{p,t}(u)\,\mathrm{d}u \tag{4.34}$$

$$\text{s.t.} \qquad D_t^{(rs)} - \sum_{p\in\mathbb{P}}\sum_{t\in\mathbb{T}} f_{p,t} = 0 \qquad \forall r\in\mathbb{A}_R, s\in\mathbb{A}_S, t\in\mathbb{T}$$

$$c_i(\boldsymbol{f}) \le 0 \qquad \forall i\in\mathbb{I}$$

$$f_\varphi \ge 0 \qquad \forall f_\varphi\in\boldsymbol{f}$$

$$\sum_{\varphi\in\Phi_p^t} f_\varphi = f_{p,t} \qquad \forall p\in\mathbb{P}, t\in\mathbb{T}$$

where $\mathbb{I}$ is the set of constraint indexes, and $\boldsymbol{f}$ is the set of separable trip flow variables $\boldsymbol{f} = \{f_\varphi : \forall\varphi\in\Phi\}$. Note that, all constraints Eqs. (4.1)–(4.19) are standardised in the form $c_i(\boldsymbol{f}) \le 0$, and indexed in set $\mathbb{I}$. Nevertheless, the study in this section is extendible to the class of non-decreasing functions, stated in the following assumption.

Assumption (Non-decreasing constraint functions)**.**

A1. All constraints $c_i(\boldsymbol{f})$ are non-decreasing functions with respect to any $f_\varphi \in \boldsymbol{f}$, i.e., $\dfrac{\partial c_i(\boldsymbol{f})}{\partial f_\varphi} \ge 0\ \forall f_\varphi\in\boldsymbol{f}, \forall i\in\mathbb{I}$.

The assumption A1 implies that, the constraint function $c_i(\boldsymbol{f})$ increases with respect to the increment of flow f_φ. However, due to the condition $c_i(\boldsymbol{f}) \le 0$, the flow is only able to increase until $c_i(\boldsymbol{f}) = 0$. Because of this property, they are also called *capacitated* constraints. Note that, all constraints used in LTM-based DTA models in Section 4.1.2 satisfy this assumption (see in A.1). With this, the definition of unsaturated route follows.

Definition 1 (The unsaturated route)**.** *A path $p \in \mathbb{P}^{(rs)}$ is said to be unsaturated for demand at time t if*

$$\exists\varphi\in\Phi_p^t \qquad \forall i\in\mathbb{I} : \frac{\partial c_i(\boldsymbol{f})}{\partial f_\varphi} > 0 \Rightarrow c_i(\boldsymbol{f}) < 0. \tag{4.35}$$

Equivalently, path p is unsaturated for demand at time t if there exists a trip flow $f_\varphi \in \Phi_p^t$ such that for all constraints $c_i(\boldsymbol{f})$, either $\dfrac{\partial c_i(\boldsymbol{f})}{\partial f_\varphi} = 0$ or $c_i(\boldsymbol{f}) < 0$.

The condition Eq. (4.35) shows that for a given $\boldsymbol{f}$, if $c_i(\boldsymbol{f})$ is strictly increasing with respect to flow f_φ, then in an unsaturated route, more flow could be added to this route without violating $c_i(\boldsymbol{f}) \leq 0$. Conversely, a route p is *saturated* for demand at time t if there exists a constraint $c_i(\boldsymbol{f}) = 0$ while the above antecedent is hold, i.e.,

$$\forall \varphi \in \Phi_p^t \qquad \exists i \in \mathbb{I} : c_i(\boldsymbol{f}) = 0 \text{ and } \frac{\partial c_i(\boldsymbol{f})}{\partial f_\varphi} > 0. \tag{4.36}$$

Based on the concept of unsaturated route, the capacitated user equilibrium for a dynamic traffic network is presented in the next section.

Capacitated dynamic user equilibrium condition

The Wardrop's first principle of route choice (Wardrop, 1952) is widely used in the study of user equilibrium solutions, i.e., the journey times in all routes actually used are equal and less than those that would be experienced by a single vehicle on any unused route. This, however, is not always held in a capacitated network where the amount of flows passing through a road is restricted by the maximum flow capacity or the capacity of downstream flow is dynamically constrained by the spill-back phenomenon (Larsson et al., 1999; Correa et al., 2004).

In this kind of network, due to the selfish routing behaviour, it might happen the case that not every used path has the same journey times, especially when the shortest paths are fully filled by travellers, then alternative routes with higher costs are considered by the remains. In this type of network, the capacitated UE definition was first introduced for a static network in Correa et al. (2004). We extend its definition for a dynamic capacitated network as follows.

Definition 2 (Capacitated user equilibrium - CUE)**.** *A flow $\boldsymbol{f}$ represents a capacitated user equilibrium if no O-D pair has an unsaturated route with strictly smaller cost than any used path for that pair. Mathematically, for all $p \in \mathbb{P}^{(rs)}$,*

$$f_{p,t} > 0 \Rightarrow T_{p,t} \leq \min \left\{ T_{p',t} : \forall p' \in \mathbb{P}^{(rs)} \textit{ unsaturated for demand } D_t^{(rs)} \right\}. \tag{4.37}$$

The CUE implies that all unsaturated paths have the same cost, which is not smaller than any cost of saturated paths. Therefore, in an *incapacitated* network where every route is unsaturated, CUE becomes Wardrop equilibrium. It is also worth to note that, not all unused paths are more costly than used routes, because they might be saturated at zero flow, e.g., path p is unused for flow $f_{p,t}$ but used for flow $f_{p',t'}$ causing saturation at the shared links between them (Larsson et al., 1999). However, this situation is avoidable in the transportation network, as sooner or later, traffic in any roads is cleared (i.e. we are able to find a trip flow f_φ that does not violate any constraints).

Proposition 2. *Suppose problem $\mathcal{P}_0$, satisfying the assumptions A1, has an optimal solution $\boldsymbol{f}^*$. Then, $\boldsymbol{f}^*$ is a CUE solution with the equilibrium cost $T^*_{p,t} = T_{p,t}(f^*_{p,t})$.*

Proof. See A.2. □

Remark 3. *The proof of Proposition 2 does not require any specific definition of $T_{p,t}(u)$, that means $\mathcal{P}_0$ can describe a more general UE-DTA problem with any increasing-continuous travel cost functions, including the average travel time function shown in Eq.* (4.25).

Furthermore, the objective function $F(\boldsymbol{f})$ in Eq. (4.34) looks similar to the one in Beckmann et al. (1956), however they are different in two aspects: (i) $F(\boldsymbol{f})$ considers the whole-path travel cost explicitly instead of assembling it into the sum of relevant arc costs, (ii) it differentiates path costs according to the physical paths and time-dependent travel demands.

In the literature, Larsson et al. (1999) shows that the CUE solutions also satisfy the Wardrop equilibrium in terms of generalized route travel costs. For the problem $\mathcal{P}_0$ in this chapter, we are able to get the same result as follows.

Let $\mathcal{L}(\boldsymbol{f}, \boldsymbol{\pi}, \boldsymbol{\mu})$ define the Lagrangian function, i.e.,

$$\mathcal{L}(\boldsymbol{f}, \boldsymbol{\pi}, \boldsymbol{\mu}) = F(\boldsymbol{f}) + \sum_{r,s,t} \pi_{rs,t}\big(D_t^{(rs)} - \sum_{p,t} f_{p,t}\big) + \sum_{i \in \mathbb{I}} \mu_i c_i(\boldsymbol{f}) \tag{4.38}$$

where $\boldsymbol{\pi}$ and $\boldsymbol{\mu}$ are multipliers associated to the constraints in the problem $\mathcal{P}_0$.

Definition 3 (Generalized route travel cost)**.** *For a given solution $\boldsymbol{f}^*$ and multiplier $\boldsymbol{\mu}^*$, let's denote*

$$\mathcal{F}_{p,t}^{\varphi} = T_{p,t}(f_{p,t}^*) + \sum_{i \in \mathbb{I}} \mu_i^* \frac{\partial c_i(\boldsymbol{f}^*)}{\partial f_{\varphi}} \qquad \forall \varphi \in \Phi_p^t \tag{4.39}$$

$$\mathcal{F}_{p,t} = \min_{\varphi \in \Phi_p^t} \mathcal{F}_{p,t}^{\varphi}. \tag{4.40}$$

$\mathcal{F}_{p,t}^{\varphi}$ is called the generalized route travel cost for $f_{\varphi} \in \Phi_p^t$, and $\mathcal{F}_{p,t}$ is called the generalized route travel cost for $f_{p,t}$ regarding to demand $D_t^{(rs)}$.

Proposition 3 (The generalized Wardrop equilibrium (GWE) for DTA)**.** *The solutions to the problem $\mathcal{P}_0$ are Wardrop equilibrium flows for the generalized route travel cost, that means*

$$f_{p,t}^* > 0 \Rightarrow \mathcal{F}_{p,t} = \pi_{rs,t}^* \qquad \forall p \in \mathbb{P}^{(rs)} \tag{4.41}$$

$$f_{p,t}^* = 0 \Rightarrow \mathcal{F}_{p,t} \geq \pi_{rs,t}^* \qquad \forall p \in \mathbb{P}^{(rs)}. \tag{4.42}$$

Within each $f_{p,t}$, the trip flow $f_{\varphi} \in \boldsymbol{\phi}_{p,t}$ is also equilibrium, meaning that

$$f_{\varphi}^* > 0 \Rightarrow \mathcal{F}_{p,t}^{\varphi} = \mathcal{F}_{p,t} = \pi_{rs,t}^* \qquad \forall \varphi \in \Phi_p^t \tag{4.43}$$

$$f_{\varphi}^* = 0 \Rightarrow \mathcal{F}_{p,t}^{\varphi} \geq \mathcal{F}_{p,t} \geq \pi_{rs,t}^* \qquad \forall \varphi \in \Phi_p^t. \tag{4.44}$$

Proof. See A.3. □

Note that, GWE does not require the assumption A1, meaning that CUE is GWE but not vice versa. By definition, a flow at a particular time and space is related to a trip flow f_{φ}, therefore, the Eqs. (4.43) and (4.44) show the *temporal* aspect of equilibrium, i.e., departure time choice (where space or route choice is fixed at path p), while the Eqs. (4.41) and (4.42) show the *spatial* aspect of equilibrium, i.e., route choice (where the departure time is fixed at time t).

In practice, each side constraint is typically related to a road segment, and the multiplier $\boldsymbol{\mu}$ is the price for travellers to choose a route. Therefore, GWE is suitable to

study the travel cost that involves the side effect during the movement, for example, the impact of dynamic link tolls (Altman et al., 2004; Yang et al., 2004b).

In summary, $CUE \Rightarrow GWE \overset{\text{A1}}{\Rightarrow} CUE$. In the next section, we study in detail both UE and SO objective functions and shows the relation between them (the term 'UE' means 'CUE' in the following sections).

4.2 The relation between UE and SO

In this section, we expand the expression of UE objective as the function of route travel time in the problem $\mathcal{P}_0$ and derive the relation between UE and SO objectives. Furthermore, we show that by using the marginal travel cost function $\tilde{T}_{p,t}(u)$ instead of the average travel cost function $T_{p,t}(u)$, $\mathcal{P}_0$ becomes the SO-DTA problem.

4.2.1 Evaluation of UE objective

On each path, for a given demand starting at time t, the relation between flow and travel time forms a stair function as shown in Fig. 4.1. For a particular path flow $f_{p,t}$, let g_i denote the accumulation of $f_{p,t,h}$ up to time $(t+i)$, i.e., $g_i = \sum_{h=t}^{t+i} f_{p,t,h}\ \forall i \in \mathbb{N}, i \geq 0$. Initially, we have $g_0 = 0$. Let $m = \max_{i \geq 0}\{i : g_i = 0\}$ denote the time up to which none of the traffic departing at time t has arrived at the destination. For any positive path flow $f_{p,t} > 0$, from Eq. (4.25), the average travel time function $T_{p,t}(u)$ for an amount u of flow is computed by

$$\begin{aligned} T_{p,t}(u) &= \frac{(i+1)(u-g_i) + \sum_{j=1}^{i} j f_{p,t,t+j}}{u} \\ &= \frac{(i+1)(u-g_i) + \sum_{j=1}^{i} j(g_j - g_{j-1})}{u} \end{aligned} \tag{4.45}$$

for any $u \in [g_i, g_{i+1}]$ and $i \geq m$. The function $T_{p,t}(u)$ is also called a running average of the travel time of path flow $f_{p,t}$. If $i \geq m$, the integration of $T_{p,t}(u)$ in the range $[g_i, g_{i+1}]$ becomes:

$$\int_{g_i}^{g_{i+1}} T_{p,t}(u)\mathrm{d}u = \int_{g_i}^{g_{i+1}} \frac{(i+1)(u-g_i) + \sum_{j=1}^{i} j f_{p,t,t+j}}{u}\mathrm{d}u$$

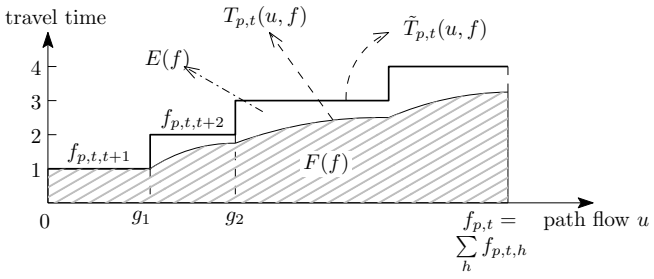


FIGURE 4.1: Stair function shows the travel time of a given flow on a path.

$$= \begin{cases} (i+1)(g_{i+1}-g_i) - \sum_{j=1}^{i} g_j \ln \frac{g_{i+1}}{g_i} & \text{if } i > m \\ (i+1)(g_{i+1}-g_i) & \text{if } i = m. \end{cases}$$

From that, for the whole range $(0, f_{p,t})$, after some simple transformations, we get:

$$\begin{aligned} \int_0^{f_{p,t}} T_{p,t}(u)\mathrm{d}u &= \sum_{i\geq m} \int_{g_i}^{g_{i+1}} T_{p,t}(u)\mathrm{d}u \\ &= (m+1)(g_{m+1}-g_m) + \sum_{i\geq m+1} \Big((i+1)(g_{i+1}-g_i) - \sum_{j=1}^{i} g_j \ln \frac{g_{i+1}}{g_i} \Big) \\ &= \sum_{i\geq 1} i f_{p,t,t+i} - \sum_{i\geq m+1} \sum_{j=1}^{i} g_j \ln \frac{g_{i+1}}{g_i} \\ &= \sum_{h>t} (h-t) f_{p,t,h} - \sum_{i\geq 0} g_i \big(\ln f_{p,t} - \ln g_i \big). \end{aligned}$$

The second sum ignores the role of m because $\lim_{g_i \to 0} g_i \ln g_i = 0$, and $\lim_{g_i \to 0} g_i \ln f_{p,t} = 0$ for a given $f_{p,t}$. To distinguish g_i from different path flows, it is replaced by the notation $g_{p,t,t+i}$. The computation of UE objective function follows:

$$\begin{aligned} F(\boldsymbol{f}) &= \sum_{p\in\mathbb{P}} \sum_{t\in\mathbb{T}} \int_0^{f_{p,t}} T_{p,t}(u)\mathrm{d}u \\ &= \underbrace{\sum_{p\in\mathbb{P}} \sum_{t\in\mathbb{T}} \sum_{h>t} (h-t) f_{p,t,h}}_{F_T(\boldsymbol{f})} - \underbrace{\sum_{p\in\mathbb{P}} \sum_{t\in\mathbb{T}} \sum_{i\geq 0} g_{p,t,t+i} \big(\ln f_{p,t} - \ln g_{p,t,t+i} \big)}_{E(\boldsymbol{f})} \end{aligned} \tag{4.46}$$

Note that, $F_T(\boldsymbol{f})$ represents the total travel time, which is minimized in the system optimal DTA problem.

4.2.2 SO objective as the marginal equilibrium

The system optimal objective is defined by minimizing the total travel time $F_T(\boldsymbol{f})$, which is defined below:

$$F_T(\boldsymbol{f}) = \sum_{p\in\mathbb{P}}\sum_{t\in\mathbb{T}}\sum_{h>t}(h-t)f_{p,t,h}. \tag{4.47}$$

In Proposition 4, we show that, Eq. (4.34) could be used to define both UE and SO objectives by using different definitions of travel cost, i.e., marginal travel time Eq. (4.24) for SO and the average travel time Eq. (4.25) for UE.

Proposition 4. *The function Eq.* (4.34) *in problem $\mathcal{P}_0$ represents the SO objective by considering the marginal travel time function, i.e.,*

$$\tilde{F}(\boldsymbol{f}) = \sum_{p\in\mathbb{P}}\sum_{t\in\mathbb{T}}\int_0^{f_{p,t}}\tilde{T}_{p,t}(u)\mathrm{d}u = F_T(\boldsymbol{f}). \tag{4.48}$$

Proof. In the previous section, we have proved that, problem $\mathcal{P}_0$ provides UE solutions. Similarly, by applying Eq. (4.24) into this problem, we get:

$$\begin{aligned}
&\int_{g_{i-1}}^{g_i}\tilde{T}_{p,t}(u)\,\mathrm{d}u &&= i(g_i - g_{i-1}) = if_{p,t,t+i} && \forall i \geq 1\\
\Rightarrow\quad &\int_0^{f_{p,t}}\tilde{T}_{p,t}(u)\,\mathrm{d}u &&= \sum_{i\geq 1}\int_{g_{i-1}}^{g_i}\tilde{T}_{p,t}(u)\,\mathrm{d}u = \sum_{i\in\mathbb{T}} if_{p,t,t+i}\\
\Rightarrow\quad &\tilde{F}(\boldsymbol{f}) &&= \sum_{p,t}\int_0^{f_{p,t}}\tilde{T}_{p,t}(u)\mathrm{d}u = \sum_{p\in\mathbb{P}}\sum_{t\in\mathbb{T}}\sum_{h\geq t}(h-t)f_{p,t,h} = F_T(\boldsymbol{f})
\end{aligned}$$

Therefore, solving problem $\mathcal{P}_0$ with the definition of travel cost in Eq. (4.24) gives SO solutions. □

Applying the CUE definition to SO solutions, the marginal travel cost of any used paths are not smaller than the marginal cost of any unsaturated paths. In the case that all traffic exits the network at the end of time horizon, the SO objective is equivalent to the maximum aggregation of throughput at the destination, proved in Proposition 5.

Proposition 5. *If all demands exit the network at the end of time horizon, then*

$$\min_{f} F_T(\boldsymbol{f}) \Leftrightarrow \max_{f} TF(\boldsymbol{f}) = \max_{f} \sum_{t \in \mathbb{T}} (T + 1 - t) u_{s,t} \tag{4.49}$$

where s is the destination.

Proof. According to Eq.(1) in Shen et al. (2008), we deduct that,

$$\min_{f} F_T(\boldsymbol{f}) \Leftrightarrow \max_{f} \sum_{t \in \mathbb{T}} \sum_{h=0}^{t} u_{s,h} \Leftrightarrow \max_{f} \sum_{t \in \mathbb{T}} (T + 1 - t) u_{s,t}.$$

□

The Proposition 5 suggests an efficient way to achieve SO solutions by only considering the traffic flow entering the destination. Note that, these variables are also supported by the linked-based traffic model described in Section 4.1.2. Moreover, using the original SO objective Eq. (4.47) requires the implicit assumption (or constraint) that all demands finish at destinations, otherwise, the minimization of $F_T(\boldsymbol{f})$ would come to zero total travel time. For these reasons, we use $TF(\boldsymbol{f})$ in the SO-DTA model in this work.

4.2.3 The solution existence

According to the Kuhn-Tucker conditions sufficient for generalized convex programs (Friesz, 2010, Theorem 2.19), $\boldsymbol{f}^*$ is the global optimal solution of problem $\mathcal{P}_0$ if the following conditions are satisfied:

- The feasible domain is non-empty and open convex set.
- Objective function $F(\boldsymbol{f})$ is pseudoconvex.
- Capacitated constraints $c_i(\boldsymbol{f})$ are quasiconvex, and other equality constraints are quasilinear.

- f^* is a Kuhn-Tucker (KT) point, i.e., there exist multipliers μ and π satisfying the KKT conditions:

$$f^{*T}\nabla\mathcal{L}(f^*,\pi,\mu)=0$$
$$\mu^T c(f^*)=0.$$

Table 4.1 summarises the proposed DTA models and includes the complexity for each model. For the SO-DTA model, problem $\mathcal{P}_0$ is transformed to the link-based program described in Table 4.1. Due to its linearity, the above requirements of the objective function and constraints are simply satisfied. However, for the UE-DTA model, the need of FIFO constraints will lead to a non-convex feasible domain which is difficult to identify whether a KT point is a local minimum or not. In the next section, we propose a method for UE solutions by solving a series of SO-DTA problems. This method is able to achieve an approximated UE solution if there is a SO solution for each SO-DTA sub-problem.

4.2.4 Model complexity

All materials above help to describe DTA models, typically for SO and UE solutions. Table 4.1 summarizes the proposed DTA models and includes the complexity for each model. Note that, the UE-DTA problem, called $\mathcal{P}_{UE}$, requires the path-flow variables $f_{p,t,h}$ with double time indexes. Due to this extra variables and constraints, the complexity of $\mathcal{P}_{UE}$ in terms of the number of constraints and/or variables is proportional to $O(|\mathbb{A}||\mathbb{P}|T^2)$. In contrast, the SO-DTA problem, called $\mathcal{P}_{SO}$, only needs the link-based traffic model described in Section 4.1.2, therefore its complexity is significantly smaller than $\mathcal{P}_{UE}$ and proportional to $O(|\mathbb{A}|T)$.

Furthermore, the problem $\mathcal{P}_{SO}$ is linear in comparison with the non-linearity of $\mathcal{P}_{UE}$, therefore, it is more efficient to solve $\mathcal{P}_{SO}$ than $\mathcal{P}_{UE}$. In Eq. (4.46), we show the relation between UE and SO objectives. In the next section, we utilize this relation to develop a so-called incremental solution method to compute UE solutions based on solving the related SO-DTA problems instead of directly solving $\mathcal{P}_{UE}$.

TABLE 4.1: Summary of DTA models (Notation: $\surd$ for Used, and 'X' for Unused).

Objective or constraints	UE-DTA ($\mathcal{P}_{UE}$)	SO-DTA ($\mathcal{P}_{SO}$)	Type	# constraints or # variables
UE (Eq. (4.46))	$\surd$	X	Non-linear	-
SO (Eq. (4.49))	X	$\surd$	Linear	-
Link-based LTM model				
Eqs. (4.1)–(4.12)	X	$\surd$	Linear	$O(\lvert\mathbb{A}\rvert T)$
Path-based LTM model				
Eqs. (4.1)–(4.19)	$\surd$	X	Non-linear	$O(\lvert\mathbb{A}\rvert\lvert\mathbb{P}\rvert T^2)$

4.3 The solution method

In the previous sections, we develop DTA models which are represented in the discrete time. In this section, we first show that the difference between SO and UE objective functions, i.e., $E(\boldsymbol{f})$, is able to reduce according to the length of a time step. Based on this observation, we develop the ISM method to solve these models effectively.

In fact, the time domain in the proposed models are discretized from the continuous time horizon $[0, \mathcal{T}]$ with the time interval δ, i.e., $\mathbb{T} = \{1, 2, 3, ..., T\}$ where $T = \dfrac{\mathcal{T}}{\delta}$. In continuous time, let $d^{(rs)}(\tau)$ denote a time-dependent demand function, showing the arrival rate at source r to destination s at time $\tau \in [0, \mathcal{T}]$. The relation between $D_t^{(rs)}$ (in discrete time domain) and $d^{(rs)}(\tau)$ (in continuous time domain) reads,

$$D_t^{(rs)} = \int_{\delta(t-1)}^{\delta t} d^{(rs)}(\tau)\mathrm{d}\tau. \tag{4.50}$$

Therefore, there is the same total traffic demands D for both time domains, i.e.,

$$D = \sum_{t\in\mathbb{T}} D_t^{(rs)} = \int_0^{\mathcal{T}} d^{(rs)}(\tau)\mathrm{d}\tau. \tag{4.51}$$

Note that, the t index is for discrete time, and τ index is for continuous time.

Similarly, in a capacitated network, the link flow capacity $Q_{a,t}$ is also discretized from the continuous-time link capacity $q_a(\tau)$ where

$$Q_{a,t} = \int_{\delta(t-1)}^{\delta t} q_a(\tau)\mathrm{d}\tau \tag{4.52}$$

for all $a \in \mathbb{A}$ and $t \in \mathbb{T}$. Given functions $d^{(rs)}(\tau)$ and $q_a(\tau)$, the parameters $D_t^{(rs)}$ and $Q_{a,t}$ are function of the time interval δ. We rewrite in this section the relation between UE and SO objective in Eq. (4.46) with the appearance of the time interval δ as follows,

$$F(\boldsymbol{f}) = \underbrace{\sum_{p\in\mathbb{P}}\sum_{t\in\mathbb{T}}\sum_{h>t}\delta(h-t)f_{p,t,h}}_{F_T(\boldsymbol{f})} - \underbrace{\sum_{p\in\mathbb{P}}\sum_{t\in\mathbb{T}}\sum_{i\geq 0}\delta g_{p,t,t+i}\big(\ln f_{p,t} - \ln g_{p,t,t+i}\big)}_{E(\boldsymbol{f})}. \tag{4.53}$$

Note that, the value of $F(\boldsymbol{f})$ is expected to be consistent with the continuous-time solution when the time step δ approaches zero.

4.3.1 Approximation of UE objective function

While computing $F(\boldsymbol{f})$, we observe that $E(\boldsymbol{f})$ in Eq. (4.53) has little impact on the UE objective value when the time step δ goes to zero (see Proposition 6 and Remark 4 below).

Proposition 6. *Let $N_{p,t}$ denote the number of positive elements $f_{p,t,h}$ belonging to path flow $f_{p,t}$. If $N_{p,t}$ is bounded by a constant N (independent from δ, p and t), then*

(i) $0 \leq E(\boldsymbol{f}) \leq \dfrac{\delta D N}{e}$,

(ii) $\lim\limits_{\delta\to 0} E(\boldsymbol{f}) = 0$

where D is the total traffic demands, i.e., $D = \sum\limits_{t\in\mathbb{T}} D_t^{(rs)}$.

Proof. See A.4. □

Remark 4. *Note that, from Eqs.* (4.50) *and* (4.52)*, both the discrete demand $D_t^{(rs)}$ and the link flow capacity $Q_{a,t}$ are approximately proportional to the value of time interval δ. The value of $N_{p,t}$ is normally proportional to $D_t^{(rs)}$ and inversely proportional to $Q_{a,t}$, then the impact of δ to $N_{p,t}$ is destructive, that means N is likely independent from δ (and p or t). While we only need an upper-bound of $N_{p,t}$ for the convergence, Carey (2009) further observed the time-independence of $N_{p,t}$, where each time halving the length of the time intervals, the number of time-space links in the fans (i.e., $N_{p,t}$ in our study) remains much the same in the study of UE solutions. From this, by applying Proposition 6, we have $\lim\limits_{\delta\to 0} E(\boldsymbol{f}^*) = 0$, where $\boldsymbol{f}^*$ is an UE solution.*

Furthermore, $F_T(\boldsymbol{f})$ is linear while $E(\boldsymbol{f})$ is non-linear. By reducing the time step and ignoring the term $E(\boldsymbol{f})$ in the approximation of $F(\boldsymbol{f})$, we achieve two important goals: (i) more detail of traffic propagation in the network, and (ii) improving the computational complexity by avoiding the non-linear formulation of $E(\boldsymbol{f})$.

4.3.2 Difference between UE and SO models

In the previous section, we show that for a given solution, e.g., UE or SO solutions, $E(\boldsymbol{f})$ tends to zero at the infinitesimal time scale. However, note that it is not equivalent to achieve the UE solutions by simply solving the SO-DTA problem at this small time step. The reason is that UE and SO DTA problems require the different set of constraints, i.e., FIFO. As discussed in Section 4.1.2, the SO solution only satisfies FIFO over links (and over paths) but not over O-D pair (i.e., O-D FIFO is not satisfied) as in the UE solution. We demonstrate this difference in the simple example below.

Let's consider a simple network with two links connecting source R and destination S as in Fig. 4.2. The travel time is 10 seconds on the first path, and 12 seconds on the second one. All paths have the identical flow capacity of one vehicle per second. Let d denote the amount of traffic demand arriving R uniformly during the time interval $(0, 2)$. Assume that, $d = 6$ (or 3 veh/s in two seconds).

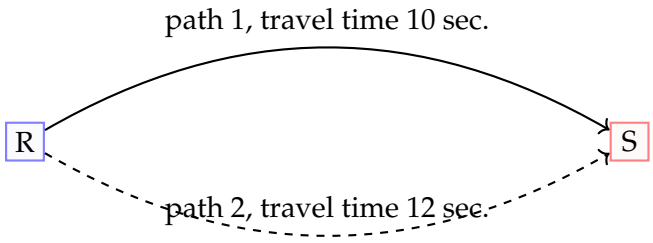


FIGURE 4.2: The simple network with two paths from source R to sink S.

Table 4.2 results the UE and SO solutions with the time step $\delta = 1$ second. For SO solution, the traffic demand starting at time $\tau = 0$ exits the network at time $\tau = 10$ (on path 1) and $\tau = 12$, $\tau = 13$ (on path 2) while the later demand only selects the first path and arrive the destination S from time $\tau = 11$ to $\tau = 13$. It clearly shows that the SO solution does not guarantee the O-D FIFO.

TABLE 4.2: SO and UE traffic patterns with the time step $\delta = 1$ second, where the vector (x, y) shows x vehicles on path 1, y vehicles on path 2.

		Demand at R			Flows at S on (path1, path2)				
Time (sec.)	τ	0	1	...	10	11	12	13	14
UE solution		3	-	-	$(1,0)$	$(1,0)$	$(1,0)$	-	-
		-	3	-	-	-	-	$(1,1)$	$(0.5,0.5)$
SO solution		3	-	-	$(1,0)$	-	$(0,1)$	$(0,1)$	-
		-	3	-	-	$(1,0)$	$(1,0)$	$(1,0)$	-

In contrast, for UE solution, the demand at time $\tau = 0$ choose path 1 as the average travel time is equal to 11 seconds, which is also smaller than the marginal cost on path 2 (12 seconds). Due to the delay on the first path to release this demand, the next traffic at time $\tau = 1$ observes the identical travel costs on both paths, therefore splits equally on them. The O-D FIFO is satisfied in the UE solution as all traffic starting at time $\tau = 0$ finishes sooner than the later demand at $\tau = 1$, shown in Table 4.2.

Furthermore, this UE solution is also achieved by solving sequentially the SO problems with the additional O-D FIFO requirement, as follows.

- We first consider only the demand at time $\tau = 0$ and ignore the others. The SO and UE solutions are identical together as they both select the smallest travel cost.

- We continue solving the SO solution for demand at time $\tau = 1$, given the assignment of path flow traffic for demand at time $\tau = 0$. By doing this, we guarantee the O-D FIFO. The SO and UE solutions result the same for demand at $\tau = 1$.

Remark 5. *In the infinitesimal time scale, both UE and SO objective functions find the smallest cost route for the infinitesimal flow, however, UE and SO solutions differ in regard to the O-D FIFO property. Using the simple example, we show that, in a small time scale, by imposing the O-D FIFO in the SO-DTA problem, we are able to reach the UE solution.*

4.3.3 The ISM framework

Based on the insights from the previous section, we propose herein an incremental approach, called the incremental solution method (ISM), to obtain the UE solution

for the UE-DTA problem.

The incremental procedure: By using a sufficiently small time step and based on the Proposition 6 and Remarks 4 and 5, the $E(\boldsymbol{f})$ term in the UE objective function (4.46) can be omitted which results in a linear system optimal objective. The incremental procedure also guarantees the FIFO relation by imposing the O-D FIFO requirement at any links. The overall method is described below.

Step 0 Choose the time step δ, and set $t = 1$.

Step 1 Find a SO solution for the demand $D_t^{(rs)}$, given the solution for previous demands ($\{D_h^{(rs)} : \forall h \in \mathbb{T}, h < t\}$): $\hat{\boldsymbol{f}} = \{\hat{f}_{ab,t}, \forall a, b \in \mathbb{A}, t \in \mathbb{T}\}$, by solving the following SO-DTA problem:

$$[\mathcal{P}_{SO}^{+}] \qquad \max_{f} TF(\boldsymbol{f}) = \max_{f} \sum_{t \in \mathbb{T}} (T + 1 - t) u_{s,t}$$

s.t. Link-based LTM model: *Eqs.* (4.1) $-$ $-$(4.12)

$$(f_{ab,t} - \hat{f}_{ab,t}) \sum_{h \geq t+1} \hat{f}_{ab,h} = 0 \qquad a, b \in \mathbb{A}, t \in \mathbb{T}. \tag{4.54}$$

In fact, the problem $\mathcal{P}_{SO}^{+}$ is the problem $\mathcal{P}_{SO}$ (in Table 4.1) with the additional *linear* constraint Eq. (4.54) (for a given $\hat{\boldsymbol{f}}$), which shows that at any time t, there is new traffic from link a to b only if no flow from previous demand continues on this direction after time t.

Step 2 If there is no more demand after t, then the last solution is the final one. Otherwise, update $t = t + 1$, and jump to step 1.

Study of convergence: Given a solution $\boldsymbol{f}^*$ generated from this framework, we could evaluate $E(\boldsymbol{f}^*)$ in Eq. (4.46) to know how close to the UE solution. Based on Proposition 6, the smaller value of $E(\boldsymbol{f}^*)$, the better approximation of UE solution $\boldsymbol{f}^*$.

4.3.4 Performance of ISM method

In this section, we show the performance of ISM in a number of iterations ($\mathcal{I}$), constraints ($\mathcal{C}$) and variables ($\mathcal{V}$). At each iteration, the problem $\mathcal{P}_{SO}^{+}$ includes a

number of constraints $\mathcal{C}_I = O(|\mathbb{A}|\mathcal{T}\delta^{-1})$ and variables $\mathcal{V}_I = O(|\mathbb{A}|\mathcal{T}\delta^{-1})$ (where $|\mathbb{T}| = \mathcal{T}\delta^{-1}$). Note that, $\mathcal{P}_{UE}$ model in Table 4.1 is more complex than each iteration of ISM due to the combination of path and time indexes, i.e., $\mathcal{C}_{UE} = O(|\mathbb{A}||\mathbb{P}|\mathcal{T}^2\delta^{-2})$, and $\mathcal{V}_{UE} = O(|\mathbb{A}||\mathbb{P}|\mathcal{T}^2\delta^{-2})$.

Furthermore, by the linearity, each iteration in ISM is solved efficiently than the original non-linear UE-DTA model. As shown in our numerical result and in Wright et al. (2014), the computational time for non-linear model increases exponentially to the increment of complexity, while linear solution methods are proved to have polynomial computing time (Karmarkar, 1984). These computational evaluations show the efficiency of ISM to obtain UE solutions.

In the trade-off between the efficient performance and the approximation of the solution, ISM might require numerous iterations, which is decided by the chosen time step δ, i.e., $\mathcal{I} = O(\mathcal{T}\delta^{-1})$. We note that the ISM method could be extended to vary the time unit δ at each iteration, therefore this is another way to enhance its performance in practice. In summary, the ISM is a divide-and-conquer method that helps to improve the computational performance by sequentially solving the linear SO-DTA problems ($\mathcal{P}^{+}_{SO}$). In the next section, we demonstrate this method in two numerical examples.

4.4 Numerical results

The numerical results are studied in two different network sizes: small-size Braess network and medium-size grid network. In Braess network, our ISM method is studied in different aspects: the approximation of UE solutions via $E(\boldsymbol{f})$ and the chosen value of time step to achieve user optimal solutions. For the larger (grid) network, the performance of ISM is measured in terms of complexity (i.e., a number of iterations, constraints and variables) and computational time.

4.4.1 The Braess network

Fig. 4.3 shows the network topology with parameter settings in Table 4.3. The dynamic of O-D demand and link flow capacity are shown in Fig. 4.4. The traffic demands depart from source R at rate 2 veh/s in 120 seconds. From the supply side,

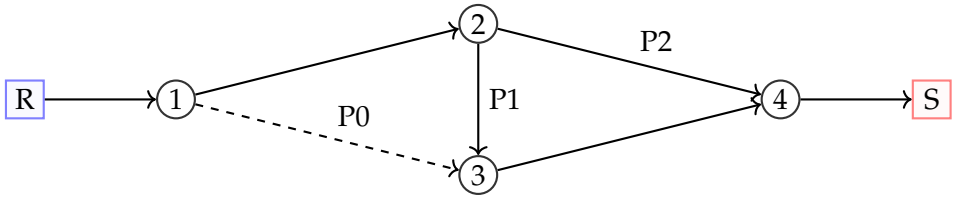


FIGURE 4.3: Braess network topology with source R and sink S.

TABLE 4.3: Configuration for each link in Fig. 4.3.

Links	**L** (m)	**V** (m/s)	**W** (m/s)	**K** (veh/m)	**Q** (veh/s)	Free-flow travel time (s)
(1,2), (3,4)	1000	16.67	8.33	0.24	1.33	60
(1,3), (2,4)	1000	16.67	8.33	0.12	0.67	60
(2,3)	500	16.67	8.33	0.12	0.67	30

link (1,3) drops its flow capacity in 30 seconds at different periods shown in Fig. 4.4. These dynamic settings impact on the route choices in one of three paths: P0 (1,3,4), P1 (1,2,3,4), and P2 (1,2,4). In the free-flow network state, paths P0 and P2 are shortest with 120-second travel time while path P1 costs 150 seconds. However, the flow capacity on P0 and P2 cannot cover heavy traffic demands entering the network, therefore causing travel delay in these paths. If the delay is large enough, traffic will split to P1.

In this example, to study the impact of the capacitated network to the UE solution, we design three scenarios, illustrated in Fig. 4.4:

- Normal case: there is no capacity drop at any links.
- Case 1: Link (1,3) drops its flow capacity in 30 seconds from the initial time.
- Case 2: Link (1,3) also drops its flow capacity in 30 seconds but from $t = 211$ to $t = 240$.

Note that, the capacity drop in case 1 and case 2 could be caused by the red signal of a traffic light or an incident. With these configurations, we apply ISM method in two different time scales, i.e., $\delta = 3s$ or $\delta = 15s$. The theoretical results in previous sections show that the solution from ISM method with small time step could lead to UE solutions.

Solution without capacity drop (normal case) The result is shown in the first-row figures in Fig. 4.5. The travel cost is steady after $t = 60$ and equal to the free-flow

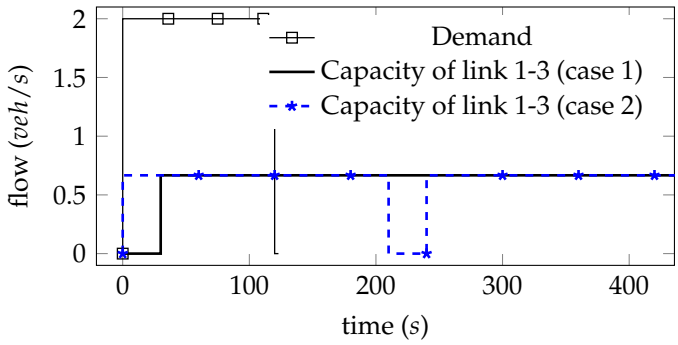


FIGURE 4.4: The dynamic setting for network in Fig. 4.3.

travel time on P1 (150 seconds) when all three paths are used. Before this time, no traffic selects P1 because its cost is still higher than the travel cost on P0 and P2. We also observe that the traditional Wardrop UE holds in the two chosen time steps.

Impact of capacity drop to UE solutions When the inflow capacity of the link (1,3) changes temporarily in a short period of 30 seconds, the results are shown as the last four figures in Fig. 4.5. These cases aim to put the off-periods before or after the traffic network is steady in the normal case. In the first case where the link (1,3) is unused in the first 30 seconds, the solution shows that travellers first choose the only P2, then P0 and P2, and finally all three paths. In comparison with the normal case, the steady state of travel cost happens sooner due to the higher increasing rate of travel cost. Therefore, the early drop on P0 does not worsen the network performance (any travellers departing later than $t = 60$ still experience the same cost as in normal case).

In the second case, since link (1,3) is still available before $t = 90$, the traffic is still loaded in both P0 and P2 until the network starts to be steady from the departure time $t = 60$. However, right after this time, P0 is temporarily unused, then the traffic splits on P1 and P2. After link (1,3) is recovered, all three paths are used, and the network reaches a new steady state with higher travel cost, i.e., travellers would take 160 seconds to reach the destination instead of 150 seconds in the previous two cases. In summary, the time and duration of a capacity drop do impact on the UE solutions as described above. If the capacity drop is planned ahead, a good system performance is achievable as in the first case of this example.

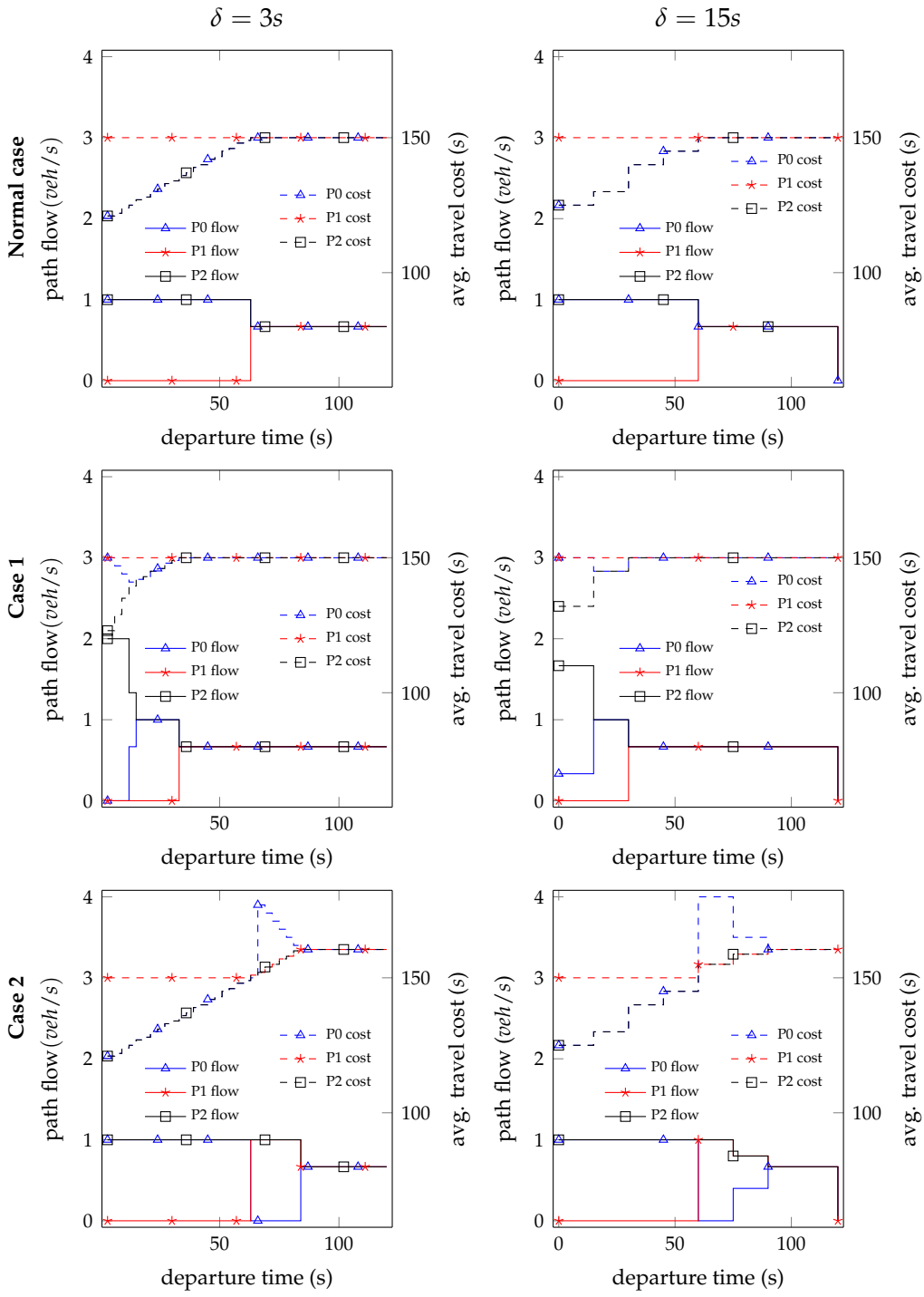


FIGURE 4.5: The path flows and travel cost in Braess network.

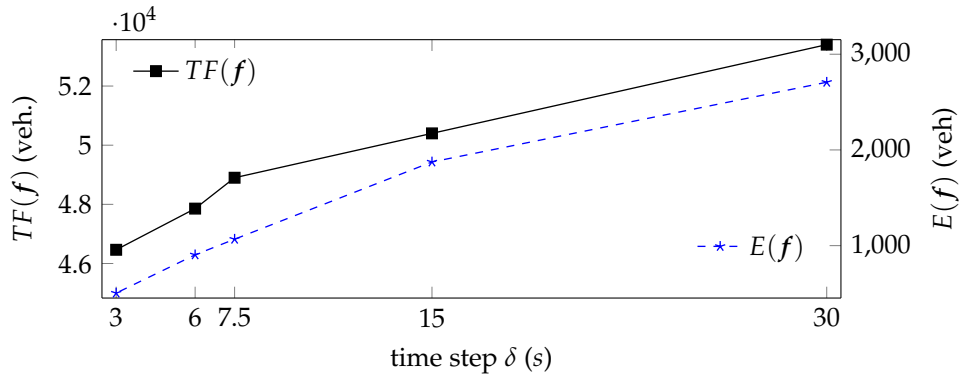


FIGURE 4.6: Evaluation of $E(f)$ with respect to different time steps.

The capacitated travel cost Fig. 4.5 shows the results in different scenarios and time steps. At small time step ($\delta = 3s$), the travel cost and path flows follow the traditional Wardrop UE. However, the solution aggregated at higher time step ($\delta = 15s$) shows the idea of CUE where the travel costs are not the same at some departure times. For example, for $t \leq 15$ in case 1 ($\delta = 15s$), flows on P0 and P2 experience different costs. The reason is that the cost is averaged over a large range of demand which is not uniformly distributed over the large time step. Via more details of the solution at $\delta = 3s$ (case 1), for $t \leq 15$, traffic before $t = 12$ is only loaded on P2 and having lower cost, thus the average of travel cost over the first 15 seconds is not the same for P0 and P2. Therefore, a higher resolution of traffic (smaller time step), a better representation of UE solutions.

Evaluation of $E(f)$ In this network, we present the difference between SO and UE objective, i.e., $E(f)$, with respect to time steps $\delta \in \{3s, 6s, 7.5s, 15s, 30s\}$ in Fig. 4.6. It is shown that $E(f)$ decreases and tends to zero with the reduction of time step, that supports the Proposition 6. Particular, with $\delta = 3$, the value of $E(f)$ is about 1% of the SO objective value $TF(f)$. In the next section, we perform our method in a medium-size network to evaluate the computational performance.

4.4.2 Medium-size grid network: computational performance

This section compares the complexity and computational performance of two methods: ISM (with different time steps) and solving $\mathcal{P}_{UE}$ directly (at the large time step).

All methods are applied in the medium-size grid network (Fig. 4.7), with the parameters for two types of links described in Table 4.4. In this network, the normal links have higher flow capacity and travel time than the saturated links. Therefore, travellers will first choose paths through the saturated links until their travel costs are risen and equal to other longer paths.

TABLE 4.4: Grid network with parameter settings for the network in Figure 4.7.

Constant parameters:			
$V_a = 15$ (m/s)		$W_a = 7.5$ (m/s)	
$\omega = \frac{W_a}{V_a} = 0.5$		Demand: 4.5 veh/s (total: 180 veh)	
Time step (sec.): $\{5, 10, 20, 40\} \approx$ Conf. $\{C5, C10, C20, C40\}$			
Link types	K_a (veh/m)	Q_a (veh/s)	**Link length** (m)
Normal	0.12	0.6	1200
Bottleneck	0.6	0.3	600

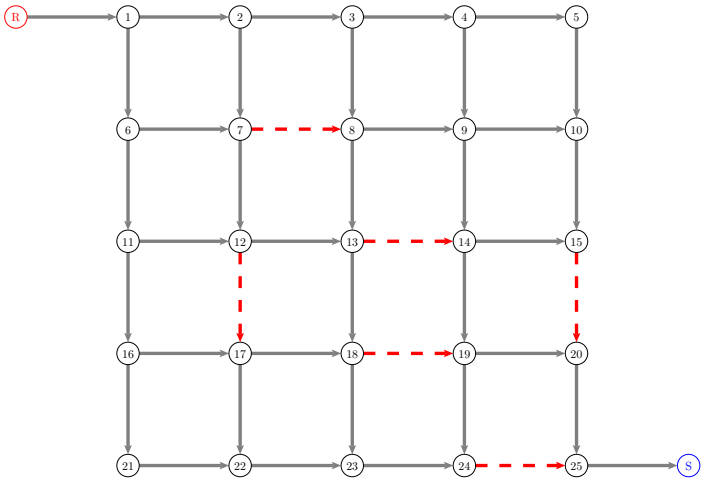


FIGURE 4.7: Grid network topology for TTM (5x5).

Estimation of complexity

In our grid network example, all nodes have two incoming links and two outgoing links except the boundary nodes. Thus, the number of inflow-outflow combinations is $2|\mathbb{A}|$. In the estimation for grid network, problem $\mathcal{P}^{+}_{SO}$ in Section 4.3.3 includes a number of variables $\mathcal{V}_{ISM} \approx 4|\mathbb{A}|T$, and constraints $\mathcal{C}_{ISM} \approx 9|\mathbb{A}|T$ in the ISM method.

For the problem $\mathcal{P}_{UE}$, their complexities increase quadratically with the size of time domain, and about $8T_D$ times in comparison with the ISM method, shown in

TABLE 4.5: Estimation of complexity in terms of constraints and variables for grid network in Fig. 4.7.

	$\lvert\mathbb{A}\rvert = 40, \lvert\mathbb{P}\rvert = 70$, T_D is the number of time steps at which demand value is positive.		
	$\mathcal{C}$	$\mathcal{V}$	Note
ISM	$9\lvert\mathbb{A}\rvert T = 360T$	$4\lvert\mathbb{A}\rvert T = 160T$	LP
$\mathcal{P}_{UE}$	$O(\lvert\mathbb{A}\rvert\lvert\mathbb{P}\rvert T_D T) = O(2800 T_D T)$	$O(\lvert\mathbb{A}\rvert\lvert\mathbb{P}\rvert T_D T) = O(2800 T_D T)$	NLP

Table 4.5. From this estimation, the performance of ISM is expected to be more efficient than solving $\mathcal{P}_{UE}$ directly.

However, to get UE traffic patterns, ISM has to use a small enough time step. According to the network, we observe that the shortest paths take 3 saturated links and 5 normal links in the total of 520 second travel time. The next short paths take 2 saturated links and 6 normal links, costing 560 seconds of travel time. For the first few travellers, they could move on two shortest paths, which go through links (15, 20) and (24, 25). To get more traffic to destination S, the only way is to find paths passing through the link (19, 20). However, the extra flow through the link (19, 20) could not utilize saturated links, which are all used in the shortest paths. It means that the alternative paths (also the longest paths) would take 640 second travel time, passing through normal links. For this particular scenario, to be more efficiently, we could predict the UE solutions as the following:

- If the total demand is less than 144 vehicles, all travellers would choose to move on shortest paths, for example, (R, 1, 2, **7, 8, 13, 14, 15, 20**, 25, S) or (R, 1, 6, 11, **12, 17, 18, 19, 24, 25**, S), where the bold numbers indicate nodes belong to the bottleneck links of the paths.

- If the demand is less than 0.6 veh/s, then they always travel on the shortest paths.

Our demand in this example is 4.5 veh/s with total of 180 vehicles. We see that the first 144 vehicles would select the two shortest paths with 640 seconds of average travel time. This large cost motivates vehicles to choose the longest path beside the shortest paths where they are all at equilibrium state. The better choice of δ could follow the below strategy:

- First, choose $\delta = 240s$, to solve the first 144 vehicles of demand.

TABLE 4.6: Performance comparison between ISM (gray-left side column) and solving directly $\mathcal{P}_{UE}$ (right side) in different traffic details.

Conf.	$\mathcal{I}$	T	$\mathcal{C}$		$\mathcal{V}$		$\mathcal{T}_{pre.}$ (s)		$\mathcal{T}_{sol.}(s)$		$TF(f)$	$F(f)$
C40	1	21	7299	31358	2941	29401	2.2	352.4	0.4	2.87e+5	37440	2595.86
C20	2	41	14269	-	5729	-	9.5	-	0.93	-	32040	
C10	4	82	28576	-	11454	-	46.6	-	2.7	-	29850	
C5	8	164	57190	-	22904	-	337.4	-	8.8	-	27450	

- Second, choose $\delta = 60s$ to solve the remaining demand.

However, to make it more general, we set up a fixed value of δ (also to avoid the transformation of solutions among different time units) and evaluate the performance in the next section.

Comparison results

The result in Table 4.6 shows the efficiency of ISM over $\mathcal{P}_{UE}$. As expected, the ISM provides UE traffic pattern in the configuration C5 where the only two shortest paths are selected. For the complexity of a number of constraints and variables, the ISM method at 5-second time unit is relatively similar to the $\mathcal{P}_{UE}$ model at 40-second time unit.

For computational performance, there are two criteria: the preparation time $\mathcal{T}_{pre.}$ and the solution time $\mathcal{T}_{sol.}$. Since all models are implemented in an abstract language, for example, AMPL, GAMS or modern programming languages (for example Python), $\mathcal{T}_{pre.}$ is necessary to transform these abstract models into concrete ones, which are understood by the solvers (cbc (Forrest, 2012) for linear model, and ipopt (Biegler et al., 2009) for non-linear models). Differently, $\mathcal{T}_{sol.}$ measures the computing time to get the solution from these solvers. In solving the $\mathcal{P}_{UE}$ model, even at the large time unit (i.e., 40 seconds), it requires a considerable large amount of computing time in comparison with the ISM as shown in Table 4.6. Therefore, while it is comfortable to perform ISM in the small time units, we only able to get solutions from solving $\mathcal{P}_{UE}$ model at 40 second time unit. The comparison between ISM (C5) and $\mathcal{P}_{UE}$ (C40) shows the advantage of our proposed method that is significantly less complexity and computational time.

4.5 Summary

This chapter develops a novel linear programming framework to solve the UE-DTA problem in a dynamic capacity network that exploits the linkage between the UE and SO solutions underpinned by a first-in-first-out (FIFO) principle. Furthermore, it provides a theoretical proof that in the limit, the SO objective can be used to obtain the UE solution as the system time step goes to zero given the satisfaction of the FIFO constraint. This important property enables us to develop an incremental loading method to obtain the UE solutions efficiently by solving a sequence of linear programs. The numerical results show the significant improvements in the obtained UE solution both in terms of accuracy and computational complexity.

Chapter 5

Information-based SO DTA problem

In using real-time information for adaptive routing, most of the existing works (with the exception of Waller et al. (2013)) focus on the user's point of view where questions regarding the improvement in travel time of the individual, or the equilibria of that of many individuals (i.e., network perspective) have been studied. In contrast to this focus, this chapter will examine the role of information from the network operators' point of view where questions regarding the long-term planning of the network with full information accessibility can be answered. In other words, given we know that road users increasingly have access to information that influences their travel behaviour, popular traffic analysis and modelling options such as dynamic traffic assignment (DTA) model (Szeto et al., 2012) should be revisited in light of this knowledge.

The remaining content is organised in six sections. The information model and choice framework are described in Section 5.1. The individual behaviour of choice is related to constraints of the DTA problem in Section 5.2. These constraints then appear in the two system optimal DTA models, considering policy choice and path choice in Section 5.3. Section 5.4 discusses the complexity of models and makes the comparison between solutions of policy and path choices. Section 5.5 presents numerical results in two different topologies: Braess network (Braess et al., 2005) for solution analysis, and Nguyen-Dupuis network (Nguyen et al., 1984) for computational evaluation. Finally, Section 5.6 summarises our work, discusses the results and gives some future directions.

5.1 The information and choice framework

5.1.1 The traffic network model

Traffic network model such as dynamic traffic assignment (DTA) models, alongside with travel forecasting and microscopic traffic simulation models, are important elements of the traffic analysis tools, frequently used by transportation planners and traffic engineers.

The inputs to the traffic network model consist of the specification of the network topology, traffic demand, supply and the route choice model. The topology of a transportation network is generally represented by a directed graph $\mathcal{G} = (\mathbb{C}, \mathbb{A})$, where $\mathbb{C}$ is the set of nodes and $\mathbb{A}$ is the set of directed arcs connecting certain pairs of nodes. The demand is defined by the flows arriving at each source node and destining to each sink node, while the supply is specified by the flow capacities (i.e., maximum allowable flows) of each road segment in the network as functions of time. Given the inputs and a so-called dynamic network loading (DNL) model, the outputs of the traffic network model are the resulting flows, occupancies or density profiles as functions of space and time. The role of the DNL model is to ensure that while network traffic is propagated forward in time, key traffic flow properties, such as flow conservation and the fundamental relationship between traffic flow variables, are preserved.

As an input to DNL model, route choice models naturally represent routing decisions from travellers. Motivated in the introduction, next we develop a novel information model and integrate it with the route choice model (in a so-called choice framework) for the above generic traffic network model.

5.1.2 A novel information model

By "information" we mean knowledge that reduces the uncertainty about some phenomenon or system. There are several sources of uncertainty in the traffic network model. The demand, the supply and the network loading model can all be stochastic (i.e., time-varying random variables). For simplicity, the topology is assumed to be

fixed in this chapter.[1]

Let D_t and S_t denote the demand and supply at time t where time is slotted over a finite time horizon $\mathbb{T}$, starting at time index 1. Note that D_t is a matrix, i.e., $D_t = [D_t^{(ij)}]$, where the (ij)th component, $D_t^{(ij)}$, represents the demand arriving at time t at node i and destined towards node j. Similarly, $S_t = [S_t^{(k)}]$ is a vector, where the kth component, $S_t^{(k)}$ represents the flow capacity of the kth road segment at time t.

Furthermore, let the joint demand and supply process be given by the stochastic process

$$\{(D_t(x), S_t(x)), t \in \mathbb{T}, x \in \mathcal{X}\},$$

where x denotes a realisation or sample path, and $\mathcal{X}$ denotes the set of all possible realisations. Note that, each realisation is a deterministic function of time. A specific realisation x defines and is defined by the traffic demand and supply pertaining to that realisation as functions of time, i.e., $\forall x \in \mathcal{X}$, we can write $x \equiv (D(x), S(x))$, where

$$(D(x), S(x)) = \{(D_t(x), S_t(x)), t \in \mathbb{T}\}.$$

Note that in this chapter we assume that the joint distribution of demand and supply is given. This joint distribution can be estimated using the distribution of the underlying causes of uncertainty such as bad weather as in Rudloff et al. (2015). In some instances, it can be approximated as a product of the marginal distributions of the demand and supply where they are largely independent. The individual distribution is then determined separately, for example as in Zito et al. (2011) for the demand and Mahmassani et al. (2012) for the supply.

Definition 4 (Scenario)**.** *The traffic network scenario is a realisation* $x \equiv (D(x), S(x))$.

We assume that *the set* $\mathcal{X}$ *is finite*, that is, the number of realisations is finite. Thus, the number of scenarios is also finite. This simplification is motivated by the fact that any stochastic process can be approximated by a sufficiently large but finite number of (deterministic) realisations (Küchler et al., 2010). Considering the scenario as a discrete random variable $\tilde{x}$, the probability distribution over the set of realisations,

[1] Temporary changes in the topology in terms of the (non)availability of certain road segments can be effectively modelled through the change in their flow capacities.

$\mathcal{X}$, is given by $\{q_x\}_{x\in\mathcal{X}}$, where

$$q_x = \Pr(\tilde{x} = x), \quad \forall x \in \mathcal{X}, \quad \text{and} \quad \sum_{x\in\mathcal{X}} q_x = 1.$$

The restriction of a realisation $x \in \mathcal{X}$ for the first t time units is called the *t-traced pattern of* x, denoted by

$$x(t) = \{(D_u(x), S_u(x)), u \leq t, u \in \mathbb{T}\}.$$

Two scenarios x and x' are said to have the same t-traced pattern, written as $x(t) = x'(t)$, if $(D_u(x), S_u(x)) = (D_u(x'), S_u(x'))$ for all $u \leq t, u \in \mathbb{T}$. The set of scenarios having the same t-traced pattern $x(t)$ is denoted by $\mathcal{X}_{x(t)}$, i.e.,

$$\mathcal{X}_{x(t)} = \{x' \in \mathcal{X} : x'(t) = x(t)\}.$$

For each t-traced pattern, $x(t)$, the random set $\mathcal{X}_{x(t)}$ is called a *t-realisable scenario set* or *hyperscene*. Note that, if $x_1(t) = x_2(t)$, then $\mathcal{X}_{x_1(t)} = \mathcal{X}_{x_2(t)}$, i.e., x_1 and x_2 belong to the same hyperscene. The power set of all possible hyperscenes at time t is called the *hyperscene set*, $\mathbb{X}_t$. As the scenario is a discrete random variable, $\tilde{x}$, so are the t-traced pattern, $\tilde{x}(t)$, and the t-realisable scenario set, $\tilde{\mathcal{X}}_{x(t)}$. Their distributions are given by

$$\Pr(\tilde{x}(t) = x(t)) = \Pr\left(\tilde{\mathcal{X}}_{x(t)} = \mathcal{X}_{x(t)}\right) = \sum_{x'\in\mathcal{X}_{x(t)}} q_{x'}.$$

Observe that the scenarios belonging to a t-realisable scenario set, $\mathcal{X}_{x(t)}$, are indistinguishable up to and including time t. Note that this requires the stochastic demand to be observed in real-time. There are several sources of information, such as data from mobile GPS, route guidance systems (e.g., Tomtom) or day-to-day data collected by road-side units (e.g., Bluetooth), etc., that would be useful in determining the near real-time O-D demand. Examples include the works by Bierlaire et al. (2004), and Barceló et al. (2015). Given the t-traced pattern $x(t)$, the probability that

the actual scenario being x' is given by

$$\Pr\left(\tilde{x} = x' | x(t)\right) = \begin{cases} 0 & \text{if} \quad x' \notin \mathcal{X}_{x(t)}, \\ \dfrac{q_{x'}}{\sum\limits_{z \in \mathcal{X}_{x(t)}} q_z} & \text{if} \quad x' \in \mathcal{X}_{x(t)}. \end{cases} \tag{5.1}$$

However, the t-realisable scenario sets are non-increasing in time, i.e.,

$$t_2 > t_1 \Rightarrow \mathcal{X}_{x(t_2)} \subseteq \mathcal{X}_{x(t_1)}. \tag{5.2}$$

Indeed, if t_2 is sufficiently larger than t_1, then two realisations x and x' ($x \neq x'$), identical up to the time t_1, would be different by the time t_2. This means that even though $\mathcal{X}_{x(t_1)} = \mathcal{X}_{x'(t_1)}$, we will have $\mathcal{X}_{x(t_2)} \neq \mathcal{X}_{x'(t_2)}$ and $|\mathcal{X}_{x(t_2)}| < |\mathcal{X}_{x(t_1)}|$ where $|\mathcal{X}|$ denotes the cardinality of set $\mathcal{X}$. This implies that the uncertainty about the scenario $\tilde{x}$ will reduce over time if one observes the realised t-traced pattern, $x(t)$. Hence, via real-time observations, the realised t-traced pattern, $x(t)$, constitutes a source of information.

Definition 5 (Global real-time information)**.** *The global real-time information $\mathcal{K}_t$ at time t is the set consisting of traced patterns up to time t, i.e., $\mathcal{K}_t = x(t)$.*

Clearly, the global real-time information $\mathcal{K}_t$ evolves over time and chronologically unfolds the actual realisation. As time goes to infinity, the actual realisation will completely be known with certainty.

In practice, the decision maker may have access to only a subset of the global information, e.g., the information pertaining to only a subset of nodes and links and/or for only a limited time window. Such *incomplete* information will be denoted by $\mathcal{K}_t(\mathcal{L}, \mathcal{T})$, where $\mathcal{L}$ and $\mathcal{T}$ denote the subset of locations (i.e., nodes and links) and time instants, respectively, for which the information is available. Furthermore, in practice, information will not be perfectly known, e.g., due to the errors in the data collection process or the lack of knowledge in all the possible realisations of the stochastic processes involved. In such cases, the decision maker will have access to only an estimate, $\hat{\mathcal{K}}_t(\mathcal{L}, \mathcal{T})$, of the actual information.

It is worth noting that in this chapter we adopt a deterministic DNL model. However, the proposed information model can be extended with another potential source of information, $y(t)$, that describes the stochasticity of the DNL model. In the latter, $y(t)$ represents the traced pattern of the traffic state, e.g., the amount of flow on all links in the network up to the current time t. The global real-time information $\mathcal{K}_t$ will then include both $x(t)$ and $y(t)$.

5.1.3 The choice framework

We develop in this section a framework (referred to as *the choice framework*) in which the route choice model is integrated with the aforementioned information model. Similar to the route choice model, the choice framework specifies the choice of paths from the source to the destination. A *path* is an alternating sequence of nodes and arcs. For example, a path starting at a source node r and ending at a destination node s, with $r, s \in \mathbb{C}$, is given by

$$(c_0, a_1, \ldots, c_k, a_k, \ldots, a_n, c_n),$$

where $c_0 = r$ and $c_n = s$ are the source and destination nodes, respectively, $c_1, \ldots, c_{n-1} \in \mathbb{C}$ are intermediate nodes and $a_1, \ldots, a_n \in \mathbb{A}$ are intermediate arcs.

The set of all possible paths in the network is denoted by $\mathbb{P}$. Each path $p \in \mathbb{P}$ is a set of connecting nodes and arcs, and we refer to a node i or arc (i, j) in path p by $i \in p, j \in p$, or $(i, j) \in p$. The set of all possible paths from node i to node j, with $i, j \in \mathbb{C}$, is denoted by $\mathbb{P}^{(ij)}$, which is a subset of $\mathbb{P}$. A *hyperpath*, $H^{(ij)}$, is defined as a nonempty set of (some) paths from node i to node j, i.e., $\forall i, j \in \mathbb{C}$, we have

$$H^{(ij)} \subseteq \mathbb{P}^{(ij)} \quad \text{and} \quad H^{(ij)} \neq \varnothing.$$

The power set of all hyperpaths from node i to node j is denoted by $\mathbb{H}^{(ij)}$, which has a cardinality of $(2^{|\mathbb{P}^{(ij)}|} - 1)$.

Definition 6 (User preference)**.** *For users that are at node i at time t and having node s as their destination, the user preference is the preferred hyperpath for each hyperscene,* $\mathcal{X}_{x(t)}$,

given by a function

$$v_t^{(is)} : \mathbb{X}_t \longrightarrow \mathbb{H}^{(is)}.$$

The above user preference function can account for different traffic conditions captured in the hyperscene. It can also be extended to different classes of users (e.g., car drivers vs. passengers on trains) and to account for user inertia or habit of some specific paths rather than a full set of the relevant paths.

Definition 7 (Real-time choice set)**.** *Given the t-traced pattern, $x(t)$, the real-time choice set $L_{x(t)}^{(is)}$ for the users that are at node i at time t and having node s as their destination is given by*

$$L_{x(t)}^{(is)} = v_t^{(is)} \left(\mathcal{X}_{x(t)} \right),$$

where the input of the function $v_t^{(is)}$ is the hyperscene $\mathcal{X}_{x(t)}$.

Note that, the above proposed real-time choice set is stochastic and dynamic instead of being deterministic and static in the existing works (Paz et al., 2009a; Waller et al., 2013). It is a central element of our *choice framework* where we incorporate the available real-time information in the decision making via the mapping between an observed *t-realisable scenario set* (i.e., *hyperscene*) of the studied traffic network scenario and the set of relevant *hyperpaths* of that network. In other words, the real-time choice set $L_{x(t)}^{(is)}$ is a collection of all possible route choices or path preferences for a given realised hyperscene, the decision maker(s) can then select a path that meets his or her own objective. For example, in the user equilibrium setting, users choose paths in $L_{x(t)}^{(is)}$ that optimise their selfish objectives. In the system optimal setting, however, the network operator makes the decisions by choosing appropriate paths in $L_{x(t)}^{(is)}$ to optimise a network-wide social objective, and the users then simply follow those chosen paths in this setting.

For simplicity, we shall discuss the route choice paradigms (or *choice framework*) in the context of complete and perfect information, $\mathcal{K}_t$. However, our route choice framework can be extended to the cases of incomplete and/or imperfect information as well. Given the real-time information, $\mathcal{K}_t$, and the real-time choice sets, $L_{x(t)}^{(is)}$, the

decision maker can make routing decisions in two possible ways: either by deciding on the next link only (policy choice) or selecting a complete path to a destination (path choice). Formally, we can describe them as follows.

- **Policy choice:**

  For users that are at node i at time t and having node s as their destination, the policy choice μ is the decision of the next node based on the real-time information, $\mathcal{K}_t$, and the real-time choice set, $L_{x(t)}^{(is)}$, at time t, given by a function

  $$\mu : (i, s, \mathcal{K}_t) \longmapsto j,$$

  such that $j \in \Gamma_i^+$ and there exists a path $p \in L_{x(t)}^{(is)}$ such that $j \in p$, where $\Gamma_i^+ = \{j : j \in \mathbb{C}, (i,j) \in \mathbb{A}\}$. For convenience, the domain of j is denoted as $\Gamma_{i,s,x(t)}^+ = \{j | j \in \Gamma_i^+ \text{and } \exists p \in L_{x(t)}^{(is)} : j \in p\}$.

- **Path choice:**

  For users that are at *source node* r at time t and having node s as their destination, the path choice π is the decision of the path to the destination s based on the real-time information, $\mathcal{K}_t$, and the real-time choice set, $L_{x(t)}^{(rs)}$, at time t, given by a function

  $$\pi : (r, s, \mathcal{K}_t) \longmapsto p,$$

  such that $p \in L_{x(t)}^{(rs)}$.

With the above generic formulation of policy and path choice, different functions π and μ can be defined depending on the cost function. For example, one can define a perception-based choice model, similar to that in Ukkusuri et al. (2007), but the evaluation of perception via flow split at intersections can take into account the real-time information, $\mathcal{K}_t$, as well as the real-time choice set, $L_{x(t)}^{(ij)}$.

In this chapter, the policy and path choice functions are implicitly defined through the system optimal solutions of the appropriate DTA problem formulations with the policy and path choice paradigms incorporated as constraints. The main difference between path and policy route choices is that in the former the flow-split over the

paths occurs only at the source, but in the latter, it can dynamically change at each intermediate node.

5.2 Properties of adaptive routing with perfect and complete information

As discussed in Section 5.1.3, we consider *perfect and complete information* in route choice paradigms. The information is perfect if its content reflects the reality about the entire past and up to the current time without any errors due to either noises (system or measurement noises) or human perception. Besides, the complete information is able to provide the current state of any location in the network, and at any time. From the user's point of view, this property means that all travellers can access the consistent global information without any restriction to the time instant, space and user classes. Particularly, the scenario set $\mathcal{X}$ contains all possible scenarios, and the real-time information, $\mathcal{K}_t$, covers the whole network at any time until the current time t. These assumptions enable us to quantify the maximum achievable gain obtained by the use of information. To enable the application of the choice frameworks uniformly among users, our models make an additional assumption of *homogeneous* travellers or single user class. Specifically, we assume that

- all the travellers observe the same real-time information $\mathcal{K}_t$ at time t,
- their preference on the hyperpath set is the same if they belong to the same origin-destination pair.

Based on the above assumptions, any travellers, residing at node i and aiming to destination s at time t, are having the same real-time information $\mathcal{K}_t$ with the same choice set $L^{(is)}_{x(t)}$. We define a *rule of consistent decisions* where travellers will make the same decisions at time t given the information $\mathcal{K}_t$ for any of the scenarios in the t-realisable set $\mathcal{X}_{x(t)}$.

Let $y(t)_x$ denote the t-pattern of traffic state in the scenario x within the hyperscene $\mathcal{X}_{x(t)}$. For the policy choice, let $f_{i,s,j|x(t),y(t)_x}$ denote the flow at node i, choosing the next node j towards the destination s, given the real-time information $\mathcal{K}_t$. Note that $f_{i,s,j|x(t),y(t)_x}$ is the result of a particular policy choice function μ, where the flow

decision at time t is denoted as $\mathbf{f}_{\mu|x(t),y(t)_x} = \{f_{i,s,j|x(t),y(t)_x} : \forall i \in \mathbb{C}, s \in \mathbb{C}_S, j \in \Gamma_i^+\}$. For a particular chosen dynamic network loading model (DNL), we able to write the network state update at time $(t+1)$ as

$$y(t+1)_x = \text{DNL}(\mathbf{f}_{\mu|x(t),y(t)_x}, y(t)_x) \qquad \forall x \in \mathcal{X}_{x(t)}. \tag{5.3}$$

Since the initial network state at time $t = 1$, $y(1) = y(1)_x \quad \forall x \in \mathcal{X}$, is the same for all possible scenarios, we observe a consistency among traffic states up to time $(t+1)$ after the routing decision at time t as a result of the above rule of consistent decisions, i.e.,

$$y(t+1)_x = y(t+1)_{x'} \qquad \forall x' \in \mathcal{X}_{x(t)}, \tag{5.4}$$

where the rule of consistent decisions implies

$$f_{i,s,t,j|x} = f_{i,s,t,j|x'} \qquad \forall x' \in \mathcal{X}_{x(t)} \tag{5.5}$$

($f_{i,s,t,j|x}$ is the simplified representation of $f_{i,s,j|x(t),y(t)_x}$ due to the consistent traffic state $y(t)_x$). In other words, the consistent decision at time t in Eq. (5.5) results in the consistency of the network traffic states up to time $(t+1)$ in Eq. (5.4).

Similarly, for the path choice, let $f_{p,t|x}$ denote the amount of flow on path p which is a result of a particular path choice function, π, for the demand arriving at the source of that path at time t. Based on the same arguments, Eq. (5.4) is also applied to path choice, while the rule of consistent decisions for the path choice is given by

$$f_{p,t|x} = f_{p,t|x'} \qquad \forall x' \in \mathcal{X}_{x(t)} \tag{5.6}$$

The above Eqs. (5.4)–(5.6) represent the properties of adaptive route choice with information, and will be utilized as the constraints in the DTA problems formulated later.

Furthermore, user preferences are accounted in the decision making by setting the flows, which correspond to the paths or next nodes not in the real-time choice

sets, equal to zero as follows:

$$f_{i,j,s,t|x} = 0 \qquad \forall x \in \mathcal{X}_{x(t)}, \forall j \notin \Gamma^{+}_{i,s,x(t)} \tag{5.7}$$

$$f_{p,t|x} = 0 \qquad \forall x \in \mathcal{X}_{x(t)}, \forall p \notin L^{(rs)}_{x(t)}, \tag{5.8}$$

and are included in the DTA problems formulated later.

5.3 Formulation of SO-DTA with information-based adaptive routing

In the previous sections, we show that through the use of real-time information, the uncertainty of scenarios reduces over time (see Eq. (5.2)). This knowledge can help travellers make more adaptive decisions during their stay in the network. This section describes the models of system optimal dynamic traffic assignment with information-based adaptive routing (policy choice and path choice) using the cell-transmission-based network loading model (CTM).

5.3.1 Network representation and notations

The CTM-based network is now represented as a directed graph $(\mathbb{C}, \mathbb{A})$ where $\mathbb{C}$ is the set of all cells, and $\mathbb{A}$ is the set of directed arcs connecting these cells. There are three non-overlapping subsets of $\mathbb{C}$: source cells $\mathbb{C}_R$, sink cells $\mathbb{C}_S$, and intermediate cells $\mathbb{C}_I$, see Fig. 5.1. Besides, each source cell only connects to intermediate cells, and there is no incoming link to any source cell. Each sink cell only receives flow from intermediate cells, and there is no outgoing link from any sink cell. Any intermediate cell could have any number of links coming to or going out of them.

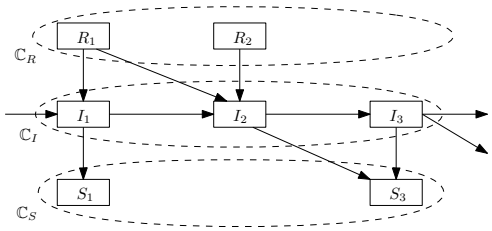


FIGURE 5.1: Cell network with three main sets of cells and their connections.

In each scenario $x \in \mathcal{X}$, for each cell $i \in \mathbb{C}$, the following parameters are defined: $Q_{i,t|x}$ for maximum flow in or out of cell i at time t, $N_{i,t|x}$ for maximum number of vehicles to be present at time t, and $\omega_{i|x}$ for backward propagation ratio (i.e., the ratio between the backward and free flow speeds). To capture the dynamic traffic in the scenario x at time t, $n_{i,t|x}$ represents a number of vehicle currently staying on cell i, and $y_{ij,t|x}$ represents a number of vehicles moving from cell i to cell j. Source cells and sink cells are used to represent the O-D pairs, and $D_{t|x}^{(rs)}$ is used to show the demand between cell r to cell s entering the network at time t in the scenario x. Especially, for source cells and sink cells, they have infinite capacity, described by: $Q_{i,t} \to \infty,\ N_{i,t} \to \infty$, for any $i \in \mathbb{C}_R \cup \mathbb{C}_S, t \in \mathbb{T}$.

The following *variables* for a given scenario x are specifically designed depending on the choice framework described in Section 5.1.3.

- Path-based CTM variables: for each cell $i \in \mathbb{C}$ at time t, $n_{i,p,t|x}$ represents a number of vehicles on this cell which follows path p, and $y_{ij,p,t|x}$ $(i, j \in p)$ is a number of vehicles moving from cell i to cell j on the same path. The variable $f_{p,t|x}$ represents a number of vehicles, belonging to demand $D_{t|x}^{(rs)}$, and choosing to move on path p.
- Policy-based CTM variables: for each cell $i \in \mathbb{C}$ at time t, $n_{i,t|x}^{(-s)}$ represents a number of vehicles on cell i at time t which has the destination $s \in \mathbb{C}_S$, and $y_{ij,t|x}^{(-s)}$ denotes a number of vehicles moving from cell i to cell j with the destination s at time t.

5.3.2 System optimal objective function

The total system cost is the total travel time that all travellers spend in the network and is formulated as follows. For any flow starting from time t_0, they could arrive at destination s at any time $t > t_0$, and hence their travel time is $(t - t_0)$. Let $y_{s,t_0,t|x}$ denote the flow entering at time t_0 and exiting at time t, aiming to destination s in the scenario x. The total system cost, averaged over all possible scenarios, is computed by:

$$TS = \sum_{x \in \mathcal{X}} \sum_{s \in \mathbb{C}_S} \sum_{t_0 \in \mathbb{T}} \sum_{t > t_0} q_x (t - t_0) y_{s,t_0,t|x} \tag{5.9}$$

In the Proposition 7 below, we state that minimizing the system cost in Eq. (5.9) is equivalent to maximising the aggregate throughput carried through the network.

Proposition 7. *Given that the network is clear (i.e. empty) at the end time, minimising total system cost in Eq.* (5.9) *is equivalent to maximising F, where*

$$F = \sum_{x \in \mathcal{X}} \sum_{s \in \mathbb{C}_S} \sum_{t \in \mathbb{T}} q_x n_{s,t|x} \tag{5.10}$$

Proof. See B.1. □

To apply this result, in the rest of this chapter, we will formulate the objective function in our DTA problems as the maximisation of F which is more suitable for CTM-based models due to the readily built-in variable $n_{s,t|x}$. The following sections describe the constraints of traffic propagation and information-based route choice for path-based and policy-based models, respectively.

5.3.3 Path-based constraints

This section describes the constraints for the Path-based Scenario-based CTM (PaSCTM). The classification of traffic into path flows in CTM will increase the complexity of the formulation but it enables us to evaluate the path-based travel cost (Lo et al., 2002) or to support the multiple O-D pairs (Waller et al., 2013). Moreover, our formulation is extended with the information framework to describe the path choices based on real-time information.

The traffic flow constraints

Based on the cell transmission model (Daganzo, 1994), the constraints of path-based traffic flow are shown below in each scenario $x \in \mathcal{X}$.

$$n_{i,p,t+1|x} = n_{i,p,t|x} + y_{bi,p,t|x} - y_{ie,p,t|x} \qquad i \in \mathbb{C}_I; t \in \mathbb{T}; b \in \Gamma_i^-; e \in \Gamma_i^+; b,i,e \in p \tag{5.11}$$

$$n_{s,p,t+1|x} = n_{s,p,t|x} + y_{bs,p,t|x} \qquad s \in \mathbb{C}_S; t \in \mathbb{T}; b \in \Gamma_i^-; b,s \in p \tag{5.12}$$

$$n_{r,p,t+1|x} = n_{r,p,t|x} + f_{p,t|x} - y_{ri,p,t|x} \qquad r \in \mathbb{C}_R; t \in \mathbb{T}; i \in \Gamma_r^+; r, i \in p \tag{5.13}$$

$$\sum_{p \in \mathbb{P}^{(rs)}} f_{p,t|x} = D_{t|x}^{(rs)} \qquad r \in \mathbb{C}_R; s \in \mathbb{C}_S; t \in \mathbb{T} \tag{5.14}$$

$$n_{i,t|x} = \sum_{p \in \mathbb{P}} n_{i,p,t|x} \qquad i \in \mathbb{C}; t \in \mathbb{T} \tag{5.15}$$

$$y_{ij,t|x} = \sum_{p \in \mathbb{P}} y_{ij,p,t|x} \qquad (i,j) \in \mathbb{A}; t \in \mathbb{T} \tag{5.16}$$

$$\sum_{j \in \Gamma_i^-} \sum_{p \in \mathbb{P}} y_{ji,p,t|x} \leq \min\{Q_{i,t|x},\ \omega_{i|x}(N_{i,t|x} - n_{i,t|x})\} \qquad i \in \mathbb{C}_I \cup \mathbb{C}_S; t \in \mathbb{T} \tag{5.17}$$

$$\sum_{j \in \Gamma_i^+} \sum_{p \in \mathbb{P}} y_{ij,p,t|x} \leq Q_{i,t|x} \qquad i \in \mathbb{C}_I; t \in \mathbb{T} \tag{5.18}$$

$$0 \leq y_{ij,p,t|x} \leq n_{i,p,t|x} \qquad (i,j) \in \mathbb{A}; t \in \mathbb{T}; p \in \mathbb{P}; i, j \in p \tag{5.19}$$

Eqs. (5.11)–(5.13) provide the path flow conservation for the sources, sinks and intermediate cells. The relationship between flow and occupancy in each cell is described by Eqs. (5.17)–(5.19), which in fact shows the upper bound of the total inflow to and outflow of a cell, assuming a triangular fundamental relationship between flow and occupancy. For each demand, Eq. (5.14) shows the conservation of demands when assigning path flow $f_{p,t|x}$. Eqs. (5.15) and (5.16) are the aggregate value of cell occupancy and link flow over paths.

The path choice consistency constraints

By observing the evolution of t-realisable scenarios $\mathcal{X}_{x(t)}$, travellers decide their adaptive path choices, and adhere to these choices till the end of their trips. Based on the consistency requirement of traffic state and path-based decision in Eqs. (5.4)

and (5.6), the following constraints are applied for path choice:

$$f_{p,t|x'} = f_{p,t|x} \qquad \forall x' \in \mathcal{X}_{x(t)}, p \in \mathbb{P} \tag{5.20}$$

$$f_{p,t|x} = 0 \qquad \forall x \in \mathcal{X}_{x(t)}, p \notin L^{(rs)}_{x(t)} \tag{5.21}$$

$$y_{ij,p,t|x'} = y_{ij,p,t|x} \qquad \forall x' \in \mathcal{X}_{x(t)}, p \in \mathbb{P} \tag{5.22}$$

Note that, in the special case where $\mathcal{X}_{x(t)} = \mathcal{X}$ and $L^{(rs)}_{x(t)} = \mathbb{P}^{(rs)}$, our path-based model reduces to the one proposed by Waller et al. (2013). In summary, the SO-DTA model for information-based adaptive path choice (SO-PaSCTM) reads:

Objective:	max F (Eq. (5.10)).
Traffic flow constraints:	Eqs. (5.11)–(5.19).
Path choice constraints:	Eqs. (5.20)–(5.22).

5.3.4 Policy-based constraints

This section describes the constraints for the Policy-based Scenario-based CTM (PoSCTM). They include the traffic flow constraints based on cell transmission model (Daganzo, 1994) and policy-based route choice constraints.

Traffic flow constraints

For each scenario $x \in \mathcal{X}$, the policy-based traffic flow constraints are shown below.

$$n^{(-s)}_{i,t+1|x} = n^{(-s)}_{i,t|x} + \sum_{b \in \Gamma^-_i} y^{(-s)}_{bi,t|x} - \sum_{e \in \Gamma^+_i} y^{(-s)}_{ie,t|x} \qquad i \in \mathbb{C}_I \cup \mathbb{C}_S; t \in \mathbb{T}; s \in \mathbb{C}_S \tag{5.23}$$

$$n^{(-s)}_{r,t+1|x} = n^{(-s)}_{r,t|x} + D^{(rs)}_{t|x} - \sum_{e \in \Gamma^+_r} y^{(-s)}_{re,t|x} \qquad r \in \mathbb{C}_R; t \in \mathbb{T}; s \in \mathbb{C}_S \tag{5.24}$$

$$n_{i,t|x} = \sum_{s \in \mathbb{C}_S} n^{(-s)}_{i,t|x} \qquad i \in \mathbb{C}; t \in \mathbb{T} \tag{5.25}$$

$$y_{ij,t|x} = \sum_{s \in \mathbb{C}_S} y^{(-s)}_{ij,t|x} \qquad (i,j) \in \mathbb{A}; t \in \mathbb{T} \tag{5.26}$$

$$\sum_{j \in \Gamma^-_i} \sum_{s \in \mathbb{C}_S} y^{(-s)}_{ji,t|x} \leq \min\{Q_{i,t|x}, \omega_{i|x}(N_{i,t|x} - n_{i,t|x})\} \qquad i \in \mathbb{C}_I \cup \mathbb{C}_S; t \in \mathbb{T} \tag{5.27}$$

$$\sum_{j \in \Gamma^+_i} y^{(-s)}_{ij,t|x} \leq n^{(-s)}_{i,t|x} \qquad i \in \mathbb{C}_I \cup \mathbb{C}_R; t \in \mathbb{T}; s \in \mathbb{C}_S \tag{5.28}$$

$$\sum_{j \in \Gamma^+_i} \sum_{s \in \mathbb{C}_S} y^{(-s)}_{ij,t|x} \leq Q_{i,t|x} \qquad i \in \mathbb{C}_I \cup \mathbb{C}_R; t \in \mathbb{T} \tag{5.29}$$

$$n_{i,t|x}^{(-s)} \geq 0 \qquad i \in \mathbb{C}; t \in \mathbb{T}; s \in \mathbb{C}_S \quad (5.30)$$

$$y_{ij,t|x}^{(-s)} \geq 0 \qquad (i,j) \in \mathbb{A}; t \in \mathbb{T}; s \in \mathbb{C}_S \quad (5.31)$$

Eqs. (5.23) and (5.24) show the conservation of policy flow at different types of cells in the network. The aggregation of cell occupancy and link flow are presented in Eqs. (5.25) and (5.26). Finally, the bounds of outflow and inflow, following the triangular fundamental relationship between flow and occupancy, are described in Eqs. (5.27)–(5.31).

The policy choice consistency constraints

Based on the analysis in Section 5.2, we have the properties of consistent traffic states in Eq. (5.4) and consistent policy decision in Eq. (5.5). These rules in PoSCTM are formulated as follows:

$$y_{ij,t|x'}^{(-s)} = y_{ij,t|x}^{(-s)} \qquad \forall x' \in \mathcal{X}_{x(t)}, s \in \mathbb{C}_S \quad (5.32)$$

$$y_{ij,t|x}^{(-s)} = 0 \qquad \forall x \in \mathcal{X}_{x(t)}, s \in \mathbb{C}_S, j \notin \Gamma_{i,s,x(t)}^{+} \quad (5.33)$$

In PoSCTM, the consistency of traffic state and that of decision are equivalent because the state and decision variables are identical. Hence, only Eq. (5.32) is added to PoSCTM to provide both the consistencies of state and decision. However, in PaSCTM the consistency of decisions is only for the source nodes and therefore the inclusion of Eq. (5.20), while Eq. (5.22) represents the consistency of traffic state involving all the nodes in the network. In summary, the SO-DTA model for information-based adaptive policy choice (SO-PoSCTM) reads:

Objective:	max F (Eq. (5.10)).
Traffic flow constraints:	Eqs. (5.23)–(5.31).
Policy choice constraints:	Eqs. (5.32) and (5.33).

5.4 Complexity and solution analysis

In the previous section, policy and path-based models are developed as two linear programming models (PoSCTM and PaSCTM) which could be solved by efficient

algorithms in general, such as simplex algorithm or interior point algorithm with the complexity of $O(n^{3.5}L^2)$ (Karmarkar, 1984) where n is the problem dimension, and L is the number of bits per variable. In this section, we study the computational complexity and compare the policy-based and path-based choice solutions. Based on the model descriptions above, we identify the advantage of PoSCTM over PaSCTM in terms of computational complexity as summarised in Table 5.1. In particular, we directly observe that $|\mathbb{C}_S| \leq |\mathbb{P}|$ (where the equality can happen in a ring-topology). Furthermore, the difference between $|\mathbb{C}_S|$ and $|\mathbb{P}|$ is usually significant, especially for multiple O-D pairs. For example, in a complete network with N nodes, a single O-D demand (or $|\mathbb{C}_S| = 1$) will lead to at least one directed path, $(N-2)$ two-link paths, $(N-2)(N-3)$ three-link paths, etc. Our numerical results will show that the execution time to solve PoSCTM among others is one of its big advantages.

TABLE 5.1: Computational complexity of path-based and policy-based models.

	Number of constraints	Number of variables
PoSCTM	$O((\lvert\mathbb{C}\rvert + \lvert\mathbb{A}\rvert)\lvert\mathcal{X}\rvert\lvert\mathbb{T}\rvert\lvert\mathbb{C}_S\rvert)$	$O((\lvert\mathbb{C}\rvert + \lvert\mathbb{A}\rvert)\lvert\mathcal{X}\rvert\lvert\mathbb{T}\rvert\lvert\mathbb{C}_S\rvert)$
PaSCTM	$O((\lvert\mathbb{C}\rvert + \lvert\mathbb{A}\rvert)\lvert\mathcal{X}\rvert\lvert\mathbb{T}\rvert\lvert\mathbb{P}\rvert)$	$O((\lvert\mathbb{C}\rvert + \lvert\mathbb{A}\rvert)\lvert\mathcal{X}\rvert\lvert\mathbb{T}\rvert\lvert\mathbb{P}\rvert)$

In particular, PoSCTM can provide not only faster computation but also a better solution than PaSCTM as stated in Theorem 5. Intuitively, for SO-PaSCTM, the choice is made only when vehicles start to enter the network using the latest information. In contrast, for SO-PoSCTM, there is a flexibility of making choice even when vehicles are inside the network.

Theorem 5. *For the same network and traffic demand, SO-PaSCTM solutions are not better than SO-PoSCTM solutions.*

Proof. We can easily see that the feasible PaSCTM solution domain is the subset of the feasible PoSCTM solution domain, e.g., any feasible PaSCTM solutions satisfy all the PoSCTM constraints. Furthermore, since the system optimal objective functions are the same, the SO solution of PaSCTM DTA problem cannot be better than that of PoSCTM DTA. □

Lemma 3. *In a non-stochastic (or a single scenario) network, for any feasible PoSCTM solution, there exists an equivalent PaSCTM solution, and vice versa.*

Proof. See B.2. □

Proposition 8. *In a non-stochastic (or a single scenario) network, if the objective function F is the same for both PoSCTM and PaSCTM, then both PoSCTM and PaSCTM solutions produce the same objective value.*

Proof. Based on Lemma 3, each PoSCTM solution can be transformed to a PaSCTM solution and vice versa. For this reason, their optimal objective values are the same with the same objective function F. In particular, this is the case of system optimal objective function in Eq. (5.10), which is used for both SO-PoSCTM and SO-PaSCTM. □

It is worth noticing that the SO-CTM linear formulation in Ziliaskopoulos (2000) itself is not enough to work with multiple destinations, because its variables are not able to differentiate two flows targeting different destinations or sinks, especially at the merging cells or nodes. For multiple O-D, the recent work (Doan et al., 2015) used a path-based model in solving SO-DTA problem. By tracing only the destinations instead of paths, Table 5.1 and Theorem 5 show that the PoSCTM formulation gives us an efficient method to solve a problem with multi-destination demands while providing the same objective values as in the path-based models.

5.5 Numerical results

In this section, we will present two case studies: (1) using Braess network to demonstrate the abilities of strategic models (PoSCTM and PaSCTM), and (2) using Nguyen-Dupuis network to compare the computational performance between two models in four criteria: number of constraints, number of variables, execution time, and objective values. The network parameters in both examples (i.e., free-flow speed, backward speed and jam density) are shown in Tables 5.2 and 5.7 respectively, following a well-accepted triangular fundamental diagram relationship. We also provide results for a real-sized problem, i.e., Fort-Worth network (Mahmassani, 2001), which includes 182 nodes, 441 links and two O-D pairs in Section 5.5.3. All results are computed in the desktop computer (Intel Core i5 3.2GHz, 8GB RAM) by the open-source solver COIN-OR CBC (Forrest, 2012) for LP problems.

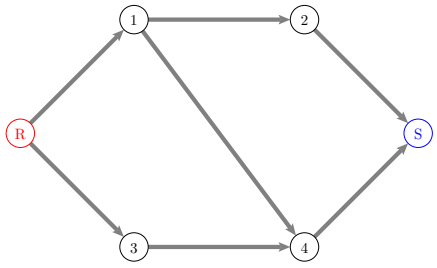


FIGURE 5.2: Braess network (bottleneck can happen at link $(2, S)$ and link $(3, 4)$).

5.5.1 Example 1: Braess network

Figure 5.2 shows the Braess topology with parameters described in Table 5.2. In this case study, three scenarios, representing good (x_0) and bad ($\{x_1, x_2\}$) situations, are considered, i.e., $\mathcal{X} = \{x_0, x_1, x_2\}$. While every link operates at full capacity at all times in scenario x_0, there are capacity drops in scenarios x_1 and x_2 (at links $(1, 2)$ and $(3, 4)$ for time $t \geq 4$). Furthermore, the duration of the capacity drop also varies, where it is much longer in scenario x_2 compared to that in x_1. Sustained capacity reduction in scenario x_2 could be as a result of bad weather, which also impacts on the demand, e.g., less demand than that in other scenarios. For convenience, let us denote three possible paths in this network as path p_0 (R-3-4-S), path p_1 (R-1-2-S) and path p_2 (R-1-4-S).

In this example, as shown in Fig. 5.3, we study one policy-based solution (PoSCTM), and two path-based solutions (PaSCTM and PaSCTM*) where PaSCTM* allows path change at sources by deactivating Eq. (5.20) in the PaSCTM model. In other words, PaSCTM* represents the temporal adaptation of path choices at the source for part of the demand arriving at time t but entering the network sometimes after t.

In Table 5.3, for each given scenario, the realisable scenario set $\mathcal{X}_{x(t)}$ as a function of time is shown. The predefined user preference of route choices tailored to each possible $\mathcal{X}_{x(t)}$ is given in Table 5.4. For example, for $t < 4$, knowing $\mathcal{X}_{x(t)} = \{x_0, x_1, x_2\}$, the user preference of route choices at the source R is the set $\{p_0, p_1, p_2\}$ (i.e., all paths are preferable or $\Gamma^{+}_{R,S,x(t)} = \{1, 3\}$ is the set of next nodes in the policy-based routing) as shown in the first column and row of Table 5.4. It can be seen from this table that if scenario x_0 is realised (i.e., by $t \geq 4$), users only prefer paths p_0 or p_1, and not p_2. In contrast, if the scenario x_2 is realised (by $t \geq 7$), the user preference

TABLE 5.2: Braess network setting.

Constant parameters (applied for all links): Free-flow speed: 1000 (m/min.), backward speed: 500 (m/min.). Jam density: 0.096 (veh/m). Link length: 500 (m). Modelling time horizon: 13 min. Time unit: 6 sec. Demands from R to S: see the right table.		
Link(s)	N (veh/cell)	Q (veh/min.)
$(R,1)$, $(R,3)$, $(1,4)$, $(4,S)$, $(1,2)$	9.6	32
$(2,S)$, $(3,4)$	9.6	see the below table

Note: $Q_{t\|x}$ is for links (2,S) and (3,4)						
Time (min.)	$x_0 : q_{x_0} = 0.2$		$x_1 : q_{x_1} = 0.3$		$x_2 : q_{x_2} = 0.5$	
	$D^{(RS)}_{t\|x_0}$ (veh/min.)	$Q_{t\|x_0}$	$D^{(RS)}_{t\|x_1}$ (veh/min.)	$Q_{t\|x_1}$	$D^{(RS)}_{t\|x_2}$ (veh/min.)	$Q_{t\|x_2}$
$1 \to 3$	64	32	64	32	64	32
$4 \to 6$	64	32	64	16	64	16
$7, 8$	64	32	64	32	0	8
$9 \to$ end	0	32	0	32	0	8

is p_0 or p_2 to avoid the long bottleneck at link $(1,2)$.

TABLE 5.3: The evolution of knowledge over time in Braess network (n^- means before the beginning of time n).

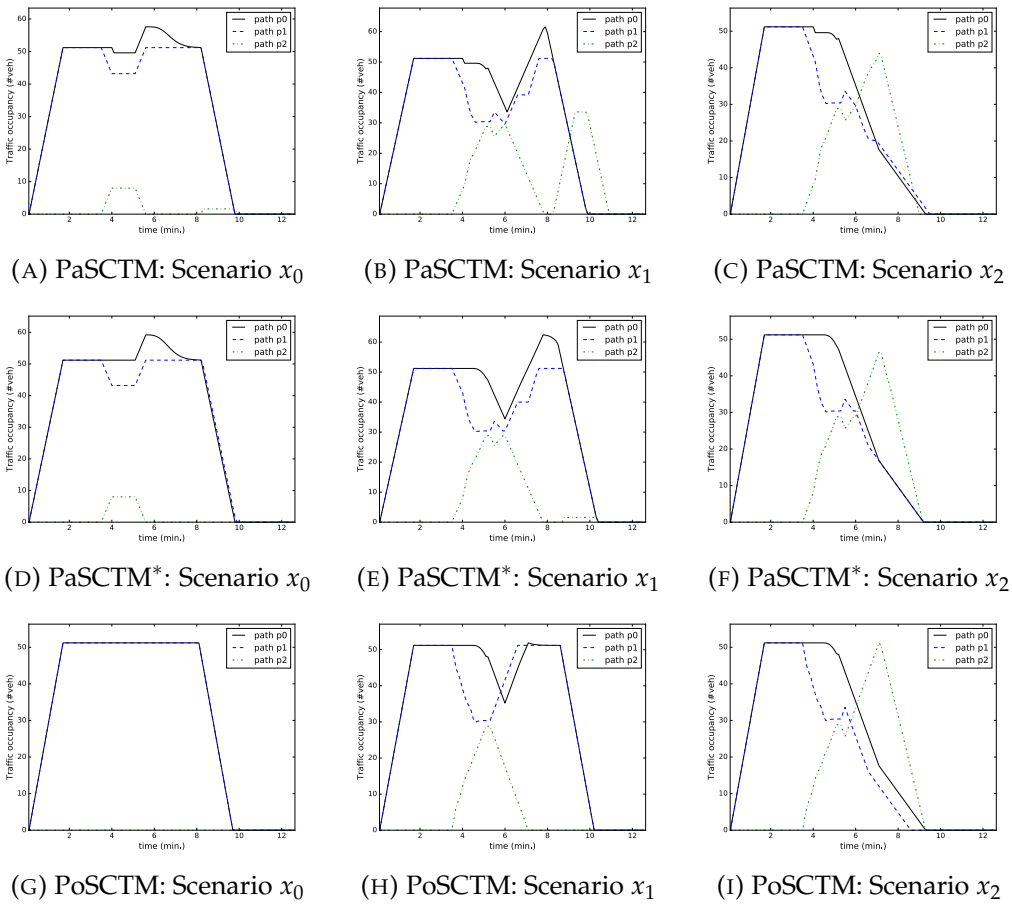
Time (in min.)	$1 \to 4^-$	$4 \to 7^-$	$7 \to$ end
$\mathcal{X}_{t\|x_0}$	$\{x_0, x_1, x_2\}$	$\{x_0\}$	$\{x_0\}$
$\mathcal{X}_{t\|x_1}$	$\{x_0, x_1, x_2\}$	$\{x_1, x_2\}$	$\{x_1\}$
$\mathcal{X}_{t\|x_2}$	$\{x_0, x_1, x_2\}$	$\{x_1, x_2\}$	$\{x_2\}$

TABLE 5.4: The user preference of route choice in Braess network.

$\mathcal{X}_{x(t)} \to$	$\{x_0, x_1, x_2\}$	$\{x_1, x_2\}$	$\{x_0\}$	$\{x_1\}$	$\{x_2\}$
Node R	$L^{(R,S)}_{x(t)} = \{p_0, p_1, p_2\}$	$L^{(R,S)}_{x(t)} = \{p_0, p_1, p_2\}$	$L^{(R,S)}_{x(t)} = \{p_0, p_1\}$	$L^{(R,S)}_{x(t)} = \{p_0, p_1, p_2\}$	$L^{(R,S)}_{x(t)} = \{p_0, p_2\}$
Node 1	$L^{(1,S)}_{x(t)} = \{p_1, p_2\}$	$L^{(1,S)}_{x(t)} = \{p_1, p_2\}$	$L^{(1,S)}_{x(t)} = \{p_1\}$	$L^{(1,S)}_{x(t)} = \{p_1, p_2\}$	$L^{(1,S)}_{x(t)} = \{p_2\}$

Impact of real-time information: the reduction in uncertainty via the use of real-time information is illustrated in Table 5.3. In particular, for the first three minutes, there is no difference amongst all three possible scenarios (both in terms of traffic demand and network capacity). In this situation, system operator assigns flows with respect to all possible scenarios. When there is a difference in network capacity at time $t = 4$, depending on the traced pattern, travellers can identify exactly if they are in the first scenario x_0 (i.e., good traffic) or in $\{x_1, x_2\}$ (bad traffic) with somewhat reduced uncertainty. Assume that the travellers are in the bad traffic situation, then

they will know for certain the realised scenario when they have passed the time instant $t = 7$, i.e., no uncertainty left. By updating the latest information and adapting to it, a better optimal solution could be achieved and provided to travellers. Particularly, for this example, the adaptation is shown at the traffic flow on path p_2: it carries a small amount of traffic in the non-congested network and more amount of traffic when congestion appears on path p_0 and path p_1 as shown in Fig. 5.3. Note that the vertical axis represents the amount of unfinished traffic over time on each of the possible paths for various scenarios and solutions.



(A) PaSCTM: Scenario x_0 (B) PaSCTM: Scenario x_1 (C) PaSCTM: Scenario x_2

(D) PaSCTM*: Scenario x_0 (E) PaSCTM*: Scenario x_1 (F) PaSCTM*: Scenario x_2

(G) PoSCTM: Scenario x_0 (H) PoSCTM: Scenario x_1 (I) PoSCTM: Scenario x_2

FIGURE 5.3: Traffic flows for all paths in Braess network, $F_{PaSCTM} = 42561.1 < F_{PaSCTM^*} = 42591.9 < F_{PoSCTM} = 42760.8$.

Spatial adaptation: Note that, only the policy-based model is able to provide this feature in comparison with path-based models. Particularly, in good traffic situation (x_0), PoSCTM only uses paths p_0 and p_1 to carry all traffic, while all three paths are used in the path-based solutions (PaSCTM or PaSCTM*). It means that the decision of travelling on path p_2 could be postponed up to the time $t = 4$ when travellers

can decide if they are in x_0 or not. At time $t = 4$, the drop in network capacity (at links $(1, 2)$ and $(3, 4)$) will affect the flow on the corresponding paths. To improve the network performance, path p_2 is then (and only then) used to avoid these bottlenecks. However, since the capacity on these bottlenecks will eventually be restored (in scenario x_1), travellers equipped with real-time information will be able to make en-route decisions to leverage on this network's improvement. In particular, if the realised scenario is x_1, we observe in Fig. 5.3h that the spreading of flow back to path p_0 and path p_1 at time $t = 7$, while in x_2, more flows keep going on path p_2 as the bottleneck does not disappear.

Furthermore, the traffic flow solution in Fig. 5.3 illustrates that travellers in PoSCTM seem to know the scenario even before the time $t = 4$ as its solution is different in comparison with that of PaSCTM during the first four minutes. Whilst the understanding of the scenarios is exactly the same in both path and policy-based models, solution with policy-choice (PoSCTM) is still better than that of path-choice (PaSCTM) because travellers entering network before the time $t = 4$ still have the ability to adjust their route after the capacity drop at time $t = 4$ if they are staying at node 1. These adjustments, performed in both time and space, make a big difference between PoSCTM and PaSCTM/PaSCTM* and improve the quality of the PoSCTM solution significantly.

We observe that policy-based and path-based solutions look similar in scenario x_1 (Figs. 5.3e and 5.3h), which means that the temporal path-based adaptation could provide a solution as good as policy-based route choice in this case. However, when the new information is obtained at time $t = 7$, the policy-based model provides better solutions because it can adapt not only in time but also in space at source node R and node 1 (Figs. 5.3f and 5.3i).

Our results clearly show the advantage of using real-time information in routing policy choice over adaptive path choice in a stochastic network in terms of capacity. We next study the temporal adaptation of different route choices where real-time information are available for both the path-choice and policy-choice models.

Temporal adaptation: For this feature, we only compare PaSCTM and PaSCTM* (policy-based model could do the same as PaSCTM* or even better). We expect better results for PaSCTM* because travellers now have chances to adapt their path

choices to the current traffic at the source R well beyond the arrival time at source. This is confirmed in Figs. 5.3d–5.3f: the objective value of PaSCTM* solution in this case increases to 42591.9, while that of PaSCTM is 42561.1.

For this network setting, the capacity drop causes congestion in the network and prevents the outgoing flow at source R. For this reason, travellers arriving at the source at time t might have to wait until they could enter the network. While PaSCTM model assigns path choices once at the arrival time t, PaSCTM* could assign path flows adaptively at the arrival time t or later. The results in Figs. 5.3a–5.3f show that, in scenarios x_0 and x_1, less flow goes on path p_2 in PaSCTM* than in PaSCTM. On the other hand, more traffic switches to path p_2 in PaSCTM* than in PaSCTM to help reduce the overall travel time in scenario x_2. This example indicates that adaptive path choice still gets the benefit of real-time information if travellers obtain it before starting their trips (and making their path choices).

The rolling horizon framework: In Table 5.5 we study the impact of real-time information to route choices both temporally and spatially (see in Fig. 5.3) using our model and compare them with results obtained from an existing rolling horizon framework by Paz et al. (2009a), which performs the en-route routing decisions via updating path choices at any nodes in the network in each so-called roll period. With the roll period being the smallest single time unit (0.1 min) and the stage length spanning over all the time horizon studied, the solution obtained by the rolling horizon method is still less optimal than our policy solution, as shown in the last two rows in Table 5.5. It is because the set of path choices in each roll period is deterministic and static which does not include possible future route changes in the next stage horizons. In contrast, our policy model optimizes the route choice decisions both in time and space within a single DTA problem which embeds the information evolution in the stochastic-dynamic real-time choice set.

Note that, the existing framework collects information synchronously over each roll period, while our work updates information asynchronously whenever new information is available. The first five rows in Table 5.5 shows our PaSCTM solution with better performance compared to that of the rolling horizon method with periodic information update. Note that in Table 5.3, the new information would be

TABLE 5.5: Solution comparison between our framework and the existing rolling horizon framework in Braess network.

Method	Roll period (min.)	Average travel time (min.)			Objective F (veh)
		scenario x_0	scenario x_1	scenario x_2	
PaSCTM		1.669	1.958	1.926	42561.08
Rolling horizon	0.1	1.669	1.956	1.929	42558.40
(without en-route	1	1.677	1.963	1.926	42545.85
routing, stage length =	3	1.880	1.944	1.985	42255.23
$3 \times$ roll period)	4	1.600	2.101	2.090	42098.29
PoSCTM		1.600	1.872	1.928	42760.80
Rolling horizon (en-route routing, all-time stage)	0.1	1.669	1.937	1.927	42591.90

received at time $t = 4$ and $t = 7$ at which the uncertainty in the demand and network capacity will be reduced. By using the small roll periods of 0.1 and 1 minute, results from rolling horizon method are close to our path-based solution as they are responsive to the critical information updates as shown in Table 5.5. Nevertheless, as mentioned above, our policy solution is still better compared to the results obtained by the smallest roll period in the rolling horizon method (see last two rows in Table 5.5). As the roll period increases, the performance of rolling horizon method degrades significantly (e.g., 3- or 4-minute roll period) as the critical updates might happen in the middle of some roll periods.

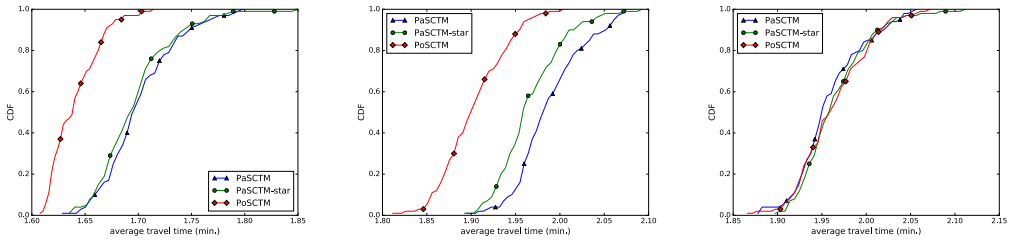
Results in Table 5.5 indicate that the rolling horizon method and its solutions are sensitive to the value of the roll period and stage horizon. In particular, one can observe the considerable difference between the solutions of 3-minute and 4-minute roll periods, respectively, as shown in Table 5.5.

Impact of demand estimation on the optimal solutions: We investigate the impact of errors in demand estimation on the optimal solutions by adding a zero mean Gaussian noise $N(0, \sigma)$ with the standard deviation σ from 10% to 30% of the expected value of the demand. The results, averaged over 100 samples for each case of σ, are summarised in Table 5.6. Observe that the standard deviation of the resulting average travel time is between 2% and 6% of its mean value in all the studied scenarios indicating the *insensitivity* of the demand estimation to the obtained SO solution. Furthermore, in Fig. 5.4, we provide the cumulative distribution (CDF) of the travel time in each scenario for the three solution methods, viz., PaSCTM, PaSCTM* and PoSCTM, with the standard deviation $\sigma = 10\%$ of the average demand. Our results

TABLE 5.6: The (mean, standard deviation) of the average travel time (in minute) caused by noised demands in Braess network.

Gaussian noise $N(0,\sigma)$	Scenario	PaSCTM	PaSCTM*	PoSCTM
$\sigma = 10\%$ mean-demand	x_0	(1.700, 0.034)	(1.696, 0.035)	(1.640, 0.023)
	x_1	(1.988, 0.039)	(1.966, 0.037)	(1.902, 0.037)
	x_2	(1.958, 0.039)	(1.964, 0.039)	(1.963, 0.040)
$\sigma = 20\%$ mean-demand	x_0	(2.028, 0.074)	(1.733, 0.064)	(1.690, 0.062)
	x_1	(1.988, 0.039)	(2.000, 0.067)	(1.954, 0.079)
	x_2	(1.999, 0.075)	(2.007, 0.073)	(2.023, 0.089)
$\sigma = 30\%$ mean-demand	x_0	(1.666, 0.044)	(1.663, 0.041)	(1.646, 0.029)
	x_1	(1.930, 0.119)	(1.884, 0.112)	(1.834, 0.106)
	x_2	(1.915, 0.106)	(1.915, 0.109)	(1.913, 0.106)

show that the travel time with PoSCTM is stochastically smaller (Shaked et al., 2007) than that with PaSCTM and PaSCTM*, especially, in scenarios x_0 and x_1, which is exactly due to the use of evolving information.



(A) Scenario x_0 (B) Scenario x_1 (C) Scenario x_2

FIGURE 5.4: Statistical cumulative distribution (CDF) of the average travel time in Braess network ($\sigma = 10\%$ of the expected demand).

In summary, the results show clearly the value of real-time information and the benefit in adapting policy-based route choice. Next, we study the computational performance of these models in a more realistic size network, referred to as the Nguyen-Dupuis network (Nguyen et al., 1984).

5.5.2 Example 2: Nguyen-Dupuis network

Let us consider the Nguyen-Dupuis network in Fig. 5.5 (Nguyen et al., 1984). With this medium scale network, we aim to demonstrate that the policy-based model provides significant gains over path-based model both in terms of complexity and optimality. Table 5.7 shows the link parameters in normal and bottleneck conditions. Similar to the previous example, there is one good scenario (x_0) when all links are in

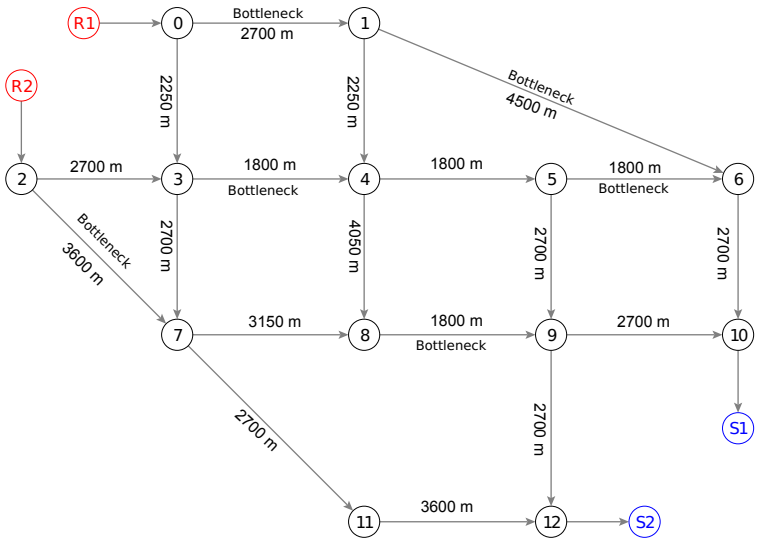


FIGURE 5.5: Nguyen-Dupuis network.

TABLE 5.7: Nguyen-Dupuis network and parameters in two scenarios.

Constant parameters (applied for all links):				
Free-flow speed: 900 (m/min.) Backward speed: 600 (m/min.) $\omega = 0.67$				
O-D demands arrive source nodes (R_1 or R_2) consecutively at 73.33 veh/min from time 1.				
Total demands (in veh.): $(R_1, S_1) : 475$ $(R_1, S_2) : 525$ $(R_2, S_1) : 575$ $(R_2, S_2) : 425$.				
Modelling time horizon: 48 minutes.				
Link type	**Scenarios**	**Time (in min.)**	**Jam density** (veh/m)	Q (veh/min.)
Normal	$\{x_0, x_1\}$	$\mathbb{T}$	0.1	36.67
Bottleneck	x_0 w.p. 0.1	$\mathbb{T}$	0.1	36.67
	x_1 w.p. 0.9	$\mathbb{T}$	0.025	9.17

normal condition, while in the other scenario (x_1) there are some bottleneck links inside the network. In particular, six links in Fig. 5.5 have smaller capacity forming the bottleneck links in scenario x_1. There are four O-D pairs, starting at R_1 and R_2, and targeting both S_1 and S_2. The computation is performed using five different time steps, from 20s to 60s. For CTM-based models, the cell networks are also prepared for each time step from the original topology. Note that smaller time steps require smaller cell lengths, hence, a larger number of cells per link due to the Courant-Friedrichs-Lewy (CFL) condition.

In what follows, we assess the computational performance (referred to as performance criteria) in three aspects: the number of constraints $\mathcal{C}$, the number of variables $\mathcal{V}$, and the execution time $\mathcal{T}$ for solving the optimisation problem. The execution time $\mathcal{T}$ consists of the preparation time $\mathcal{T}_{prepare}$ and solving time $\mathcal{T}_{solve}$. The former

TABLE 5.8: Results for Nguyen-Dupuis network (the gray columns are the results for PoSCTM, and the remains are for PaSCTM).

$\mathcal{X}$	Time unit (s)	$\mathcal{C} \div 1000$		$\mathcal{V} \div 1000$		$\mathcal{T}_{prepare}$ (s)		$\mathcal{T}_{solve}(s)$		Δ (%)
$[x_0]$	20	186.1	344.3	120.0	253.5	5.0	11.7	333.9	1430.5	0
	30	82.5	154.2	54.0	114.0	2.2	4.9	43.9	102.4	0
	40	49.4	93.2	32.7	69.1	1.3	3.0	9.1	23.7	0
	60	21.9	42.1	14.9	31.4	0.7	1.3	2.6	4.8	0
$[x_0, x_1]$	20	496.2	1221.7	319.7	686.4	18.4	44.1	1679.8	12936.9	34.3
	30	220.0	547.8	144.0	308.4	7.8	20.0	107.9	804.8	34.0
	40	131.8	331.3	87.3	186.6	4.5	11.9	39.3	165.4	32.8
	60	58.5	150.1	39.7	84.9	2.0	5.3	7.1	21.6	33.3

represents the amount of time to transform the abstract models into concrete models with a particular set of parameters, and the latter is the amount of time to find an optimal solution of the concrete model.

Results are shown in Table 7 where Δ represents the potential improvement in average travel time with policy choice utilising real-time information and is calculated as

$$\Delta = \frac{\text{path-based travel time} - \text{policy-based travel time}}{\text{path-based travel time}}.$$

In other words, the positive value of Δ shows the benefit of PoSCTM over PaSCTM. Similar to what has been noted previously, we observe here that PaSCTM solutions are less 'optimal' and have a higher computational complexity than the PoSCTM solutions.

Single scenario: we first assume that there is only x_0 in the scenario set. In this case, there is no difference in the system optimal values calculated by PoSCTM or PaSCTM, i.e., $\Delta = 0$ as seen in Table 5.8. Although the common trend is that the computational complexity increases as the time step decreases, in all cases, the PoSCTM performance criteria are much better than those of the PaSCTM (see Table 5.8).

In particular, $\mathcal{C}, \mathcal{V}$ and $\mathcal{T}_{prepare}$ of the PaSCTM are double those of the PoSCTM, while at 20-second time unit $\mathcal{T}_{solve}$ in PaSCTM is five times bigger. The significant benefit of PoSCTM over PaSCTM comes from the models with PoSCTM only traces two destinations S_1 and S_2, while PaSCTM has to trace all the possible paths (more than 20 in this network) for the four O-D pairs. Furthermore, in preparation time, finding these paths, a prerequisite in PaSCTM but not in PoSCTM, also contributes

to the cost of the execution time. In this single scenario, the results show that we can use PoSCTM solutions as a starting point to find PaSCTM solutions with less computational cost as compared to the case of finding them directly.

Two scenario set: we now consider both scenarios x_0 and x_1. Recall that the positive value of Δ obtained using different time units represents the gain of policy choice over path choice where the travel time improves by more than 30% with the perfect information as shown in Table 5.8. A double number of scenarios result in a double number of constraints and variables, but about 3-10 times of the execution time, see Table 5.8. For multiple scenarios, the same significant benefits are achieved in terms of the performance criteria for the computational complexity. However, in contrast to the single scenario set, the PoSCTM solutions no longer serve as the referenced point for the PaSCTM solutions, which must be solved at much higher computational cost.

5.5.3 Example 3: Fort-Worth network

We continue evaluating our methodology on a real-sized problem, i.e., Fort-Worth network (Mahmassani, 2001), which includes 182 nodes, 441 links and two O-D pairs, i.e., (R1, S1) and (R2, S2) shown in Fig. 5.6. By setting 30-second time step, the transformed cell network has 541 nodes and 1461 links. Similar to Nguyen-Dupuis network, we also consider two scenarios where the capacity drop happens on randomly chosen links in one of the scenarios. Due to a large number of possible paths, the path-based model could not be solved within four hours using our computing machines. For the proposed policy-based model, the results in Table 5.9 show that, for a large-sized network, solving the policy-based model requires significantly less computation than solving the path-based model to obtain the SO solution.

We also evaluate the impact of inaccuracy in the demand estimation by adding zero-mean Gaussian noise with the standard deviation 20% of the expected demand value. Fig. 5.7 shows the cumulative distribution of the average travel time where its mean and standard deviation (in minute) in each scenario are $x_0(12.30, 0.89)$ and $x_1(12.66, 0.90)$ respectively. In companion with the results in Section 5.5.1 for the small network, they certainly show that larger topology reduces the robustness of solutions.

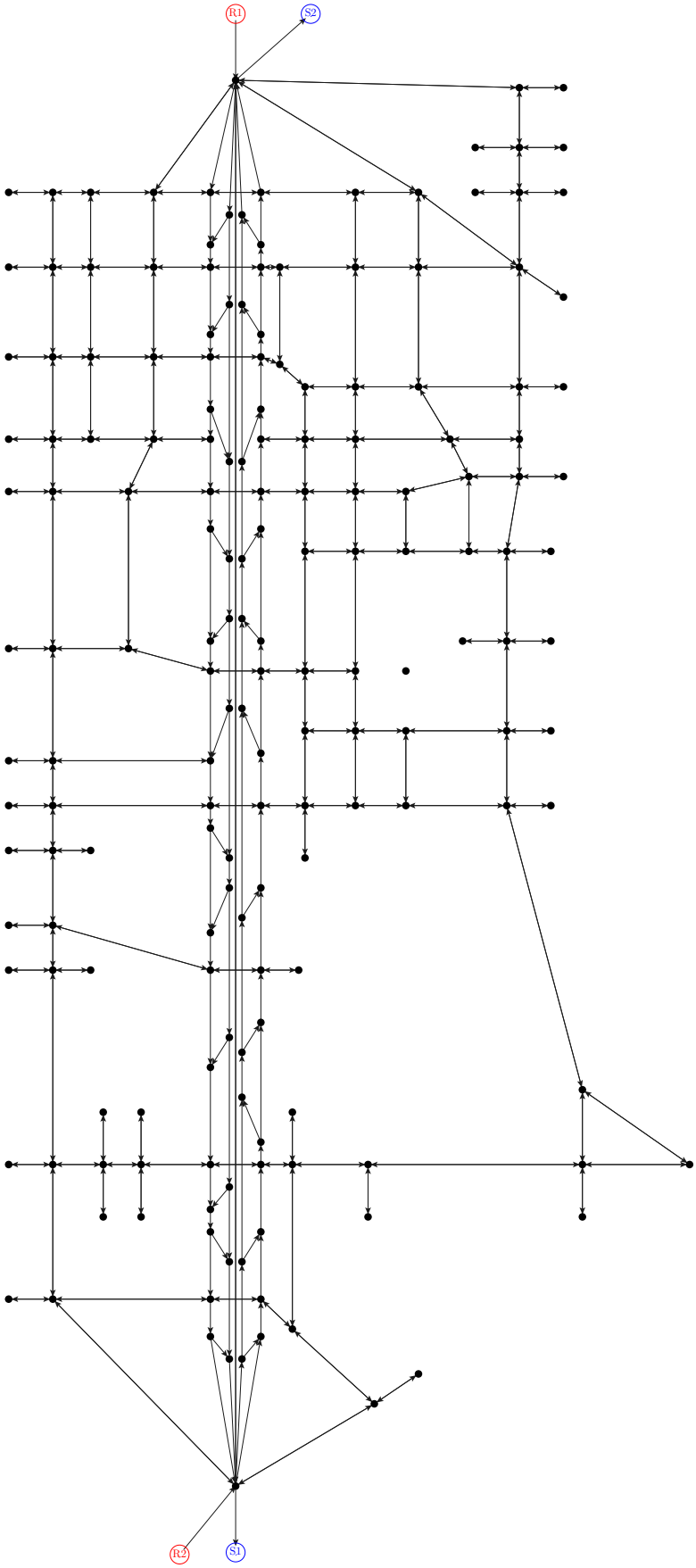


FIGURE 5.6: Fort-Worth network.

TABLE 5.9: Results of solving policy-based model in Fort-Worth network.

Demand (#veh each pair)	500	1000	1500	2000
Modelling time horizon (min.)	15	20	25	30
Average travel time (min.)	8.86	12.54	15.55	18.38
Computing time (min.)	1.0	3.7	9.4	22.8

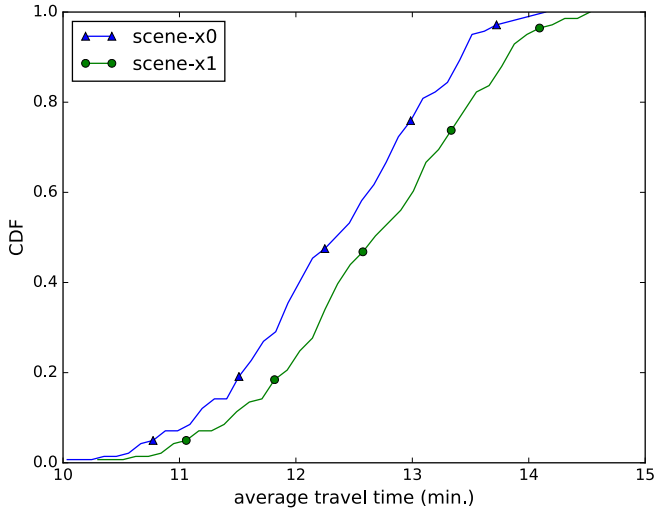


FIGURE 5.7: CDF of the average travel time in Fort-Worth network (the standard deviation 20% of the mean demand).

5.6 Summary

This chapter proposes a novel information model that caters to both the path and policy choices with real-time information in a comprehensive mathematical framework for the SO-DTA problem. This model describes an evolution of dynamic stochastic information in terms of probabilistic traffic demand and network capacity. Two linear SO-DTA models, each for path-based and policy-based route choice, formulate the DTA problem with real-time information. In this SO framework, users comply with the route advice decided by the system operator as long as the advice align with the users' preferences. The integration of information model in SO-DTA framework, therefore, enables us to study the impact of information on route choices and network performance. The numerical results show the advantage of the policy-based model over the path-based model to not only obtain a better solution in terms of system performance but also require less computational complexity and computing time.

Chapter 6

Information-based UE DTA problem

Traditional dynamic traffic assignment problem assumes the deterministic network conditions, therefore travellers know exactly their travel cost in the user equilibrium solution. Herein, we study a more general problem where users only know the most recent update of information in terms of the probabilistic traffic demand and network capacity before they enter the network. The evolution of information over time reduces the uncertainty of realisable scenarios, therefore, it enables travellers to adapt or change their route choices at the source node in real-time. This chapter proposes a mathematical programming framework to formulate the information-based stochastic user equilibrium DTA problem. The numerical examples investigate the impact of real-time information on both UE route choice and system performance.

The remaining content is organized as follows. Section 6.1 describes the information model for adaptive path-based routing and the calculation of travel time estimation. Section 6.2 proposes the LTM-based DTA model, taking into account the evolving real-time information. Section 6.3 analyses the relationship between ISUE and ISSO solutions that the ISUE objective becomes ISSO objective in the limit when the system time step comes to zero. Section 6.4 develops an incremental solution method for the proposed model and the linear approximation of ISUE objective. Section 6.5 illustrates our method in two numerical results, i.e., a simple network for solution analysis and a large-scale network for performance evaluation. Finally, our work is summarised in Section 6.6.

Notations

A transportation network is generally represented by a directed graph $\mathcal{G} = (\mathbb{V}, \mathbb{A})$, where $\mathbb{V}$ is the set of nodes, and $\mathbb{A}$ is the set of directed arcs connecting certain pairs of nodes. Without losing generality, there are three disjoint sets: source nodes $\mathbb{V}_R$, sink nodes $\mathbb{V}_S$, and intermediate nodes $\mathbb{V}_I$. Furthermore, each source node only connects to one intermediate node, and there is no incoming link to any sources. Similarly, each sink node only receives flow from one intermediate node, and there is no outgoing link from any sinks. An intermediate node could have any number of links coming to or going out of them.

Let $\mathbb{A}_R$ denote the set of arcs from sources, $\mathbb{A}_S$ denote the set of arcs to sinks, and $\mathbb{A}_I$ denote the remaining arcs. Note that, the arcs in $\mathbb{A}_R$ and $\mathbb{A}_S$ are the virtual links that store traffic at sources and destinations to maintain the rule of flow conservation. The set of all possible paths, connecting source r and destination s in this network, is denoted by $\mathbb{P}^{(rs)}$. Let $\mathbb{T}$ denote a set of discrete times where T is the maximum time horizon, i.e., $\mathbb{T} = \{1, 2, \ldots, T\}$.

Let Υ_a^- denote the set of inflow links to link a, and Υ_a^+ denote the set of outflow links from link a for any $a \in \mathbb{A}$. For any network topology in our work, we have $\Upsilon_a^- = \varnothing \; \forall a \in \mathbb{A}_R$, and $\Upsilon_a^+ = \varnothing \; \forall a \in \mathbb{A}_S$.

6.1 The information-based adaptive routing framework for path choice

In this section, we present the information-based adaptive routing framework in which travellers use real-time information of traffic network to make a decision of route choice. This decision is based on the path-based travel time estimation discussed later.

6.1.1 The framework

The scenario set Our study considers the stochastic time-dependent network characterised by the joint distribution of demand and supply (or road capacity). This joint distribution is assumed to be discrete with a finite number of realisations, also

called a *set of scenarios* $\mathcal{X}$. Each scenario $x \in \mathcal{X}$ is associated with a probability ρ_x, such that $\sum_{x \in \mathcal{X}} \rho_x = 1$. The network parameters are deterministic in each scenario that describe the traffic demand $D_{t|x}^{(rs)}$ for any O-D pair (r, s) departing at time t, and the maximum flow capacity $Q_{a,t|x}$ of any link $a \in \mathbb{A}$ at the same time t within a scenario x as

$$x = \left\{ D_{t|x}^{(rs)}, Q_{a,t|x} : \forall r \in \mathbb{A}_R, s \in \mathbb{A}_S, t \in \mathbb{T}, a \in \mathbb{A} \right\}. \tag{6.1}$$

Let y_t represent the *traced pattern* observed at time t. The group of indistinguishable scenarios sharing a same y_t is called $\mathcal{X}_{y_t}$ or the *t-realisable scenario set*. In this work, we consider the real-time information that helps to reduce the set of feasible scenarios, therefore the content of y_t helps to indicate the realisable scenario set at time t, i.e., $y_t = (\mathcal{X}_{y_t}, t)$, and $\mathcal{X}_{y_t} \subseteq \mathcal{X}$. Note that, time t in y_t is also an important information that able to differentiate dynamic network states at different times even though they realise the same $\mathcal{X}_{y_t}$. With the assumption of error-free data, *the consistency of evolving information* reads,

$$h > t \Rightarrow \mathcal{X}_{y_h} \subseteq \mathcal{X}_{y_t} \text{ or } \mathcal{X}_{y_t} \cap \mathcal{X}_{y_h} = \varnothing. \tag{6.2}$$

The Eq. (6.2) means that as the information evolves over time, the later realisable scenario set has to be a subset of the previous one. Let $\mathbb{Y}$ denote the set of possible real-time information y_t that $\mathbb{Y} = \{y_t = (\mathcal{X}_{y_t}, t) : \mathcal{X}_{y_t} \subseteq \mathcal{X}, t \in \mathbb{T}\}$. The random distribution of information at time t, called $\tilde{y}_t$, can be derived below,

$$P(\tilde{y}_t = y_t) = P_{y_t} = \sum_{x \in \mathcal{X}_{y_t}} \rho_x. \tag{6.3}$$

The decision of path choice For users departing at time t from source r to destination s, the path choice π is the decision of a path based on the real-time information y_t at time t, mathematically presented below

$$\pi : (r, s, y_t) \longmapsto p, \tag{6.4}$$

such that $p \in \mathbb{P}^{(rs)}$ for a given demand $D_{y_t}^{(rs)}$ at time t associated with y_t. According to this function, travellers choose a route depended on the realisable scenarios at time t represented by y_t. Specifically, in the case of no information update in route choice, we simply consider $y_t = (\mathcal{X}, t)$ in the above π function which means that the route choice decision takes into account all scenarios at any time. The evaluation of route cost among different paths is described in the following section.

6.1.2 Travel time estimation

In this chapter, we take *the travel time to be the travel cost* and consider two types of estimation: the marginal and average travel time. Let $f_{p,t,h|x}$ denote a flow on path p (the O-D pair $r, s \in p$), departing the source r at time t, and arriving the destination s at time h ($h > t$) in the scenario $x \in \mathcal{X}$. Let $f_{p,t|x}$ represent an amount of traffic choosing path p for demand at time t in the scenario x, which accumulates all flows $f_{p,t,h|x}$ in any time h, i.e.,

$$f_{p,t|x} = \sum_{h \geq t} f_{p,t,h|x}. \tag{6.5}$$

We assume that, each flow has to take at least one time unit from source to destination, i.e., $f_{p,t,h|x} = 0$ for all $h \leq t$.

The marginal travel time function

In a particular scenario x, for a given path flow $f_{p,t|x}$, the marginal travel time function $\tilde{T}_{p,t|x}(u)$ is the marginal travel time for an amount u of traffic belonging to $f_{p,t|x}$, formulated as

$$\tilde{T}_{p,t|x}(u) = h - t \qquad \text{if } u \in \left(\sum_{k \leq h-1} f_{p,t,k|x}, \sum_{k \leq h} f_{p,t,k|x} \right] \tag{6.6}$$

for all $u \in (0, f_{p,t|x}]$. We call $\tilde{T}_{p,t|x}$ the marginal travel time for the path flow $f_{p,t|x}$, where $\tilde{T}_{p,t|x} = \tilde{T}_{p,t|x}(f_{p,t|x})$. For a given real-time information y_t, the average of marginal travel time $\tilde{T}_{p,y_t}$ over all realisable scenario $x \in \mathcal{X}_{y_t}$ is equal to

$$\tilde{T}_{p,y_t} = \frac{1}{P_{y_t}} \sum_{x \in \mathcal{X}_{y_t}} \rho_x \tilde{T}_{p,t|x}. \tag{6.7}$$

The average travel time function

For a given path flow $f_{p,t|x}$, the average travel time function $T_{p,t|x}(u)$ for each scenario x is defined as the average of travel time over an amount u of flow belonging to $f_{p,t|x}$, i.e.,

$$T_{p,t|x}(u) = \frac{\int_0^u \tilde{T}_{p,t|x}(u)\,\mathrm{d}u}{u} \tag{6.8}$$

for all $u \in (0, f_{p,t|x}]$. The average travel time $T_{p,t|x}$ for the path flow $f_{p,t|x}$, averaged over all flow departing at time t on path p according to demand $D_{t|x}^{(rs)}$, is formulated as follows,

$$T_{p,t|x} = T_{p,t|x}(f_{p,t|x}) = \sum_{h \geq t} \frac{(h-t) f_{p,t,h|x}}{f_{p,t|x}}. \tag{6.9}$$

At any time t, for a given real-time information y_t, travellers make route choice decision based on the overall average travel time T_{p,y_t} which is averaged over all realisable scenarios in $\mathcal{X}_{y_t}$, i.e.,

$$T_{p,y_t} = \frac{1}{P_{y_t}} \sum_{x \in \mathcal{X}_{y_t}} \rho_x T_{p,t|x}. \tag{6.10}$$

6.2 The formulation for information-based DTA problem

In this section, we discuss our formulation of information-based stochastic route choice problem in the dynamic traffic assignment framework. We first present the formulation of link transmission model (LTM), the FIFO constraints, and finally, the complete formulation.

6.2.1 The link transmission model

In this part, we describe the link transmission model (LTM) in each scenario $x \in \mathcal{X}$ for the network loading as a set of side constraints in a DTA problem. The set of parameters and variables are follows.

Link parameters For each link $a \in \mathbb{A}$, the following parameters are defined:

L_a : length of link a.

K_a : jam density (a maximum number of vehicles per distance unit) at link a.

$V_{a|x}$: unsaturated speed at link a in the scenario x.

$W_{a|x}$: backward speed at link a in the scenario x.

$Q_{a,t|x}$: maximum number of vehicles to or out of link a at time t in the scenario x.

The parameters $(K_{a|x}, V_{a|x}, W_{a|x})$ helps to describe the triangular fundamental relationship between flow and density in each link. For sources and sinks, they have infinite capacity: $Q_{a,t|x} \to \infty \quad \forall a \in \mathbb{A}_S \cup \mathbb{A}_R, t \in \mathbb{T}, \forall x \in \mathcal{X}$. Furthermore, we also assume integer values of $\frac{L_a}{V_{a|x}}$ and $\frac{L_a}{W_{a|x}}$ to simplify the model description. For O-D demand, let $D_{t|x}^{(rs)}$ denote an amount of traffic demand departing at time t from source r to sink s in the scenario x.

Link variables To capture the dynamic traffic in each link $a \in \mathbb{A}$, the following variables are defined:

$u_{a,t|x}$: a number of vehicles moving to link a at time t in the scenario x.

$v_{a,t|x}$: a number of vehicles moving out of link a at time t in the scenario x.

$f_{ab,t|x}$: a number of vehicles moving from link a to link b at time t in the scenario x.

Link constraint

For each link $a \in \mathbb{A}$ in a scenario x, the demand capacity $\mathcal{D}_{a,t|x}$ defines a maximum amount of traffic exiting link a at time t, and the supply capacity $\mathcal{S}_{a,t|x}$ defines a maximum amount of traffic entering link a at time t. Their mathematical definitions are follows.

Source links ($a \in \mathbb{A}_R$) At source links, there is no limitation on demand and supply capacity.

$$\mathcal{D}_{a,t|x} \to \infty \tag{6.11}$$

$$\mathcal{S}_{a,t|x} \to \infty. \tag{6.12}$$

Sink links ($a \in \mathbb{A}_S$) At sink links, there is infinite supply but zero demand capacity to store all traffic arriving the destination.

$$\mathcal{D}_{a,t|x} = 0 \tag{6.13}$$

$$\mathcal{S}_{a,t|x} \to \infty. \tag{6.14}$$

Other links ($a \in \mathbb{A}_I$) The below equations formulate $\mathcal{D}_{a,t}$ and $\mathcal{S}_{a,t}$ which are dynamically depended on the upstream and downstream flows as follows,

$$\mathcal{D}_{a,t|x} = \min \left\{ Q_{a,t|x}; \sum_{h \le t - \frac{L_a}{V_{a|x}}} u_{a,h|x} - \sum_{h \le t-1} v_{a,h|x} \right\} \tag{6.15}$$

$$\mathcal{S}_{a,t|x} = \min \left\{ Q_{a,t|x}; K_a L_a + \sum_{h \le t - \frac{L_a}{W_{a|x}}} v_{a,h|x} - \sum_{h \le t-1} u_{a,h|x} \right\}. \tag{6.16}$$

In Eq. (6.15), the second part shows the remaining traffic at downstream point that is ready to exit link a at time t. Note that, due to the free-flow travel time $\frac{L_a}{V_{a|x}}$ on link a, any traffic entering link a after time $\left(t - \frac{L_a}{V_{a|x}}\right)$ is still travelling on this link, and will exit link a after time t. Similarly, for the supply capacity $\mathcal{S}_{a,t|x}$ at the upstream point of link a, it shows the vacancies made by the outgoing flows (that accumulate any $v_{a,h|x}$ for all $h \le t - \frac{L_a}{W_{a|x}}$) and the entering flows (that sum up any $u_{a,h|x}$ for all $h \le t-1$) of link a. These supply and demand capacity in each link are used in node model below to restrict the upstream and downstream flow at each node.

Node constraints

For any node in the network, the node constraints below guarantee the flow conservation and restriction of demand and supply.

$$u_{a,t|x} = \sum_{b \in \Upsilon_a^-} f_{ba,t|x} \qquad \forall a \in \mathbb{A}, t \in \mathbb{T}, x \in \mathcal{X} \tag{6.17}$$

$$v_{a,t|x} = \sum_{b \in \Upsilon_a^+} f_{ab,t|x} \qquad \forall a \in \mathbb{A}, t \in \mathbb{T}, x \in \mathcal{X} \tag{6.18}$$

$$u_{a,t|x} \le \mathcal{S}_{a,t|x} \qquad \forall a \in \mathbb{A}, t \in \mathbb{T}, x \in \mathcal{X} \tag{6.19}$$

$$v_{a,t|x} \leq \mathcal{D}_{a,t|x} \qquad \forall a \in \mathbb{A}, t \in \mathbb{T}, x \in \mathcal{X} \tag{6.20}$$

$$f_{ab,t|x} \geq 0 \qquad \forall a, b \in \mathbb{A}, t \in \mathbb{T} : b \in \Upsilon_a^+, x \in \mathcal{X}. \tag{6.21}$$

For source nodes, the upstream flow is given and identical to the demand at time t, i.e.,

$$u_{r,t|x} = \sum_{s \in \mathbb{A}_S} D_{t|x}^{(rs)} \qquad \forall r \in \mathbb{A}_R, t \in \mathbb{T}, x \in \mathcal{X}. \tag{6.22}$$

The path-based constraints

To trace the path flows on any links (needed for the computation of $T_{p,t|x}$), let $f_{a,p,t|x}$ represent the flow on path p arriving the upstream of link a at time t, and recall that $f_{p,t|x}$ is an amount of traffic departing from source r at time t and choosing path $p \in \mathbb{P}^{(rs)}$. Related to the link flow variables, the definite constraints follow:

$$u_{a,t|x} = \sum_{p \in \mathbb{P}: a \in p} f_{a,p,t|x} \qquad \forall a \in \mathbb{A}, t \in \mathbb{T} \tag{6.23}$$

$$v_{a,t|x} = \sum_{b \in \Upsilon_a^+} \sum_{p \in \mathbb{P}: a,b \in p} f_{b,p,t|x} \qquad \forall a \in \mathbb{A}, t \in \mathbb{T} \tag{6.24}$$

$$f_{ab,t|x} = \sum_{p \in \mathbb{P}: a,b \in p} f_{b,p,t|x} \qquad \forall a, b \in \mathbb{A}, t \in \mathbb{T} \tag{6.25}$$

$$f_{p,t|x} = f_{r,p,t|x} \qquad \forall p \in \mathbb{P}, t \in \mathbb{T}. \tag{6.26}$$

With the detail of traffic in a path, the constraint Eq. (6.22) is replaced by the following equation,

$$\sum_{p \in \mathbb{P}^{(rs)}} f_{p,t|x} = D_{t|x}^{(rs)}. \tag{6.27}$$

Similar to the demand capacity on link a due to the free-flow travel time $L_a / V_{a|x}$ (see Eq. (6.15)), the additional constraint of free-moving path flows is included, i.e.,

$$\sum_{h \leq t} f_{b,p,h|x} - \sum_{h \leq t - \frac{L_a}{V_a}} f_{a,p,h|x} \leq 0 \qquad \forall a \in \mathbb{A}_I, p \in \mathbb{P}, t \in \mathbb{T} : b \in \Upsilon_a^+, b \in p. \tag{6.28}$$

The First-In-First-Out (FIFO) constraints

The original LTM model, described in Sections 6.2.1 and 6.2.1, satisfies FIFO principle on any links (Yperman, 2007). However, by using the path-based constraints, it is required to represent FIFO among path flows in explicit constraints. The principle of FIFO describes an order of movement where a flow departs earlier at the origin would arrive sooner at the destination. Let $f_{a,p,t,h|x}$ denote an amount of traffic flow that enters link a on path p at time t and goes out of this link at time h in scenario x. The FIFO condition on link a, based on Carey (1992), follows

$$\sum_{p\in\mathbb{P}} f_{a,p,t^+,h|x} \sum_{p\in\mathbb{P}} f_{a,p,t,h^+|x} \leq 0 \qquad \forall a \in \mathbb{A}, \quad \forall h^+ > h, \quad \forall t^+ > t. \tag{6.29}$$

By Eq. (6.29), the later arriving flow $f_{a,p,t^+,h}$ (at time $t^+ > t$) could exit link a at time h only when there is no flow f_{a,p,t,h^+} that arrives sooner at time t and exits link a after time h. The relation between $f_{a,p,t,h}$ and $f_{a,p,t}$ reads,

$$f_{a,p,t|x} = \sum_{h\geq t+L_a/V_{a|x}} f_{a,p,t,h|x} \qquad \forall a \in \mathbb{A}, p \in \mathbb{P}, t \in \mathbb{T} \tag{6.30}$$

$$f_{a,p,t,h|x} = 0 \qquad \forall a \in \mathbb{A}, p \in \mathbb{P}, t \in \mathbb{T}, h < t + L_a/V_{a|x}. \tag{6.31}$$

The information-based constraints

Travellers receive real-time information y_t which could happen in any realisable scenarios $x \in \mathcal{X}_{y_t}$. Therefore, the traffic split is considered the same in these scenarios, i.e.,

$$f_{ab,t|x} = f_{ab,y_t} \tag{6.32}$$

for all $x \in \mathcal{X}, t \in \mathbb{T}, a \in \mathbb{A}, b \in \Upsilon_a^+$. This equation shows that, for any $x_1, x_2 \in \mathcal{X}_{y_t}$: $f_{ab,t|x_1} = f_{ab,t|x_2} = f_{ab,y_t}$, where f_{ab,y_t} is the common variable for traffic from link a to link b given the same y_t.

Furthermore, according to Eq. (6.4), we infer that the split of path choices is also the same for all indistinguishable scenario $x \in \mathcal{X}_{y_t}$, that is follows,

$$f_{p,t|x} = f_{p,y_t} \tag{6.33}$$

for all $x \in \mathcal{X}_{y_t}, t \in \mathbb{T}, p \in \mathbb{P}$. This constraint Eq. (6.33) implicitly requires the setting of $D_{t|x}^{(rs)}$ to be identical for any $x \in \mathcal{X}_{y_t}$. For convenience, we denote this demand value as $D_{y_t}^{(rs)}$. As a result, we can rewrite Eq. (6.27) as

$$D_{y_t}^{(rs)} - \sum_{p \in \mathbb{P}^{(rs)}} f_{p,y_t} = 0 \qquad \forall r \in \mathbb{A}_R, s \in \mathbb{A}_S, y_t \in \mathbb{Y}. \tag{6.34}$$

The above constraints, i.e., Eqs. (6.11)–(6.34), formulate the discrete-time link-based network loading model for a general single O-D road network. In the next section, we present the overall DTA model for ISUE-DTA problem.

6.2.2 The overall model for information-based stochastic user equilibrium of path choices

We have previously defined path-based variables in the LTM-based model, however, they are inseparable due to the propagation of traffic over time-space domain, and thus we cannot identify the impact of a path flow $f_{p,t|x}$ on the link flow $f_{ab,t|x}$, i.e., it is unable to determine the partial derivation $\dfrac{\partial f_{ab,t|x}}{\partial f_{p,t|x}}$ for any link and path flows. Hence, it is difficult to analyse these models under the Karush-Kuhn-Tucker (KKT) conditions for the optimal solution. For this reason, we describe the concept of a trip flow below (which is also similar to trip flow in Section 4.1.4).

Trips and trip flow variables During a trip from source to destination, each vehicle travels through several links at different times. To differentiate trips in time and space, we define a *trip* φ as a collection of all entering times and the associated arrival links for an individual (or a group of travellers) during their movement from a source to a destination, i.e.,

$$\varphi = \{(h, a) : h \in \mathbb{T}, a \in \mathbb{A}\}. \tag{6.35}$$

Each element $(h, a) \in \varphi$ means that the flow following trip φ enters link a at time h. For traffic belonging to demand $D_{y_t}^{(rs)}$ and travelling on path p, let t_a denote the arrival time to any link $a \in p$, then the trip $\varphi = \{(t_a, a) : \forall a \in p \text{ where } t_r = t, t_s > t\}$. The amount of flow following trip φ in the scenario x is called $f_{\varphi|x}$. The most important property of trip flow is *separability*, that any individual flow (or infinitesimal flow) has to belong to one and only one trip flow. A set of all (possible) trips is called Φ. Furthermore, for convenience, let Φ_p^t denote the set of trips belonging to the demand departing source r at time t and following path p to destination s ($r, s \in p$), and let $\Phi_{a,t}$ denote the set of trips entering link a at time t. Their mathematical definitions are follows,

$$\Phi_{a,t} = \{\varphi \mid \forall \varphi \in \Phi : (t, a) \in \varphi\}$$
$$\Phi_p^t = \{\varphi \mid \forall \varphi \in \Phi, \forall (t, r), (h, a) \in \varphi : r \in \mathbb{A}_R, a \in \mathbb{A}, h > t, a \text{ and } r \in p\}.$$

A set of all trip flows is called $\boldsymbol{f} = \{f_{\varphi|x} : \forall \varphi \in \Phi, \forall x \in \mathcal{X}\}$. As a result, the flow variables in previous sections are constructed from trip flows shown below,

$$f_{p,t,h|x} = \sum_{\varphi \in \Phi_p^t \bigcap \Phi_{s,h}} f_{\varphi|x} \qquad \forall p \in \mathbb{P}; s \in p \cap \mathbb{A}_S; t, h \in \mathbb{T}; x \in \mathcal{X} \tag{6.36}$$

$$f_{p,t|x} = \sum_{\varphi \in \Phi_p^t} f_{\varphi|x} \qquad \forall p \in \mathbb{P}, t \in \mathbb{T}, x \in \mathcal{X} \tag{6.37}$$

$$f_{ab,t|x} = \sum_{h<t} \sum_{\varphi \in \Phi_{b,t} \bigcap \Phi_{a,h}} f_{\varphi|x} \qquad \forall a, b \in \mathbb{A}, s \in \mathbb{A}_S, t \in \mathbb{T}, x \in \mathcal{X} \tag{6.38}$$

$$u_{a,t|x} = \sum_{\varphi \in \Phi_{a,t}} f_{\varphi|x} \qquad \forall a \in \mathbb{A}, s \in \mathbb{A}_S, t \in \mathbb{T}, x \in \mathcal{X} \tag{6.39}$$

$$v_{a,t|x} = \sum_{b \in \Upsilon_a^+} \sum_{h<t} \sum_{\varphi \in \Phi_{b,t} \bigcap \Phi_{a,h}} f_{\varphi|x} \qquad \forall a \in \mathbb{A}, s \in \mathbb{A}_S, t \in \mathbb{T}, x \in \mathcal{X}. \tag{6.40}$$

The ISUE-DTA problem

The general side constrained DTA model is defined as the optimization problem $\mathcal{P}_0$ below:

$$[\mathcal{P}_0] \qquad \min_{\boldsymbol{f}} F(\boldsymbol{f}) = \sum_{x \in \mathcal{X}} \sum_{p \in \mathbb{P}} \sum_{t \in \mathbb{T}} \rho_x \int_0^{f_{p,t|x}} T_{p,t|x}(u)\, \mathrm{d}u \tag{6.41}$$

s.t. Path-based model: Eqs. (6.11)–(6.33)

$$D_{y_t}^{(rs)} - \sum_{p \in \mathbb{P}} f_{p,y_t} = 0 \qquad \forall r \in \mathbb{A}_R, s \in \mathbb{A}_S, y_t \in \mathbb{Y}$$

$$f_{\varphi|x} \geq 0 \qquad \forall f_{\varphi|x} \in \boldsymbol{f}$$

$$\sum_{\varphi \in \Phi_p^t} f_{\varphi|x} = f_{p,t|x} \qquad \forall p \in \mathbb{P}, t \in \mathbb{T}, x \in \mathcal{X}$$

where all the constraints Eqs. (6.11)–(6.33) can be written in a standardized form

$$c_i(\boldsymbol{f}) \leq 0, \ \ \forall i \in \mathbb{I},$$

with $\mathbb{I}$ is the set of constraint indexes, and $\boldsymbol{f}$ is the set of separable trip flow variables $\boldsymbol{f} = \{f_{\varphi|x} : \forall \varphi \in \Phi, x \in \mathcal{X}\}$ (see Appendix C.1). The study in this section is generalized to a class of non-decreasing functions, stated in the following assumption.

Assumption (Non-decreasing constraint functions)**.**

A1. All constraints $c_i(\boldsymbol{f})$ are non-decreasing functions with respect to any $f_{\varphi|x} \in \boldsymbol{f}$, i.e.,

$$\frac{\partial c_i(\boldsymbol{f})}{\partial f_{\varphi|x}} \geq 0 \qquad \forall f_{\varphi|x} \in \boldsymbol{f}, \forall i \in \mathbb{I}.$$

The assumption A1 implies that, the constraint function $c_i(\boldsymbol{f})$ increases with respect to the increment of flow $f_{\varphi|x}$. However, due to the condition $c_i(\boldsymbol{f}) \leq 0$, the flow is only able to increase until $c_i(\boldsymbol{f}) = 0$. Because of this property, they are also called *capacitated* constraints. Note that, all constraints used in LTM-based DTA models in Section 6.2.1 satisfy this assumption (see in Appendix C.1). With this, the definition of unsaturated route follows.

Definition 8 (The unsaturated route)**.** *A path $p \in \mathbb{P}^{(rs)}$ is said to be unsaturated in a scenario $x \in \mathcal{X}$ for demand at time t if*

$$\exists \varphi \in \Phi_p^t \qquad \forall i \in \mathbb{I} : \frac{\partial c_i(\boldsymbol{f})}{\partial f_{\varphi|x}} > 0 \Rightarrow c_i(\boldsymbol{f}) < 0. \tag{6.42}$$

Equivalently, path p is unsaturated in a scenario x for demand at time t if there exists a trip flow $f_{\varphi|x}$ such that for all constraints $c_i(\boldsymbol{f})$, either $\dfrac{\partial c_i(\boldsymbol{f})}{\partial f_{\varphi|x}} = 0$ or $c_i(\boldsymbol{f}) < 0$.

The condition Eq. (6.42) shows that for a given $\boldsymbol{f}$, if $c_i(\boldsymbol{f})$ is strictly increasing with respect to flow f_φ, then in an unsaturated route, more flow could be added to

this route without violating $c_i(\boldsymbol{f}) \leq 0$. Conversely, a route p is *saturated* for demand at time t in scenario x if there exists a constraint $c_i(\boldsymbol{f}) = 0$ while the above antecedent is hold, i.e.,

$$\forall \varphi \in \Phi_p^t \qquad \exists i \in \mathbb{I} : c_i(\boldsymbol{f}) = 0 \text{ and } \frac{\partial c_i(\boldsymbol{f})}{\partial f_{\varphi|x}} > 0. \tag{6.43}$$

Based on the concept of unsaturated route, the information-based user equilibrium in a stochastic dynamic traffic network is presented in the next section.

Information-based stochastic user equilibrium condition

The Wardrop's first principle of route choice (Wardrop, 1952) is widely used in the study of user equilibrium solutions, i.e., the journey times in all traversed routes are equal and less than those that would be experienced by a single vehicle on any unused route. This, however, is not always held in a capacitated network where the amount of flows passing through a road is restricted by the maximum flow capacity or the capacity of downstream flow is dynamically constrained by the spill-back phenomenon (Larsson et al., 1999; Correa et al., 2004). In this type of network, we extend the definition of static capacitated UE, introduced in Correa et al. (2004), for a dynamic stochastic network as follows.

Definition 9 (Information-based stochastic user equilibrium - ISUE)**.** *A flow $\boldsymbol{f}$ represents an information-based stochastic user equilibrium if no O-D pair has an unsaturated route with strictly smaller cost than any used path for that pair where the route cost is averaged over all realisable scenarios at the departure time. Mathematically, for all $p \in \mathbb{P}^{(rs)}$ and $y_t \in \mathbb{Y}$,*

$$f_{p,y_t} > 0 \Rightarrow T_{p,y_t} \leq \min \big\{ T_{p',y_t} : \forall p' \in \mathbb{P}^{(rs)}, \exists x \in \mathcal{X}_{y_t} : \\ \textit{it is unsaturated in the scenariox for demand } D_{t|x}^{(rs)} \big\}. \tag{6.44}$$

The ISUE implies that all unsaturated paths have the same cost, which is not smaller than any cost of saturated paths. Therefore, in an *incapacitated* network

where every route is unsaturated, the travel costs T_{p,y_t} in all used paths are equilibrium.

Proposition 9. *Suppose problem $\mathcal{P}_0$, satisfying the assumptions A1, has an optimal solution $\boldsymbol{f}^*$. Then, $\boldsymbol{f}^*$ is an ISUE solution with the equilibrium cost $T^*_{p,y_t} = T_{p,y_t}(f^*_{p,y_t})$.*

Proof. See C.2. □

From Proposition 9 and Definition 9, we infer that for a given ISUE solution, if it exists a scenario $x \in \mathcal{X}$ so that path p is unsaturated for demand at time t then the average travel cost on this path is equal to the equilibrium cost.

Remark 6. *The proof of Proposition 9 does not require any specific definition of $T_{p,t|x}(u)$, that means $\mathcal{P}_0$ can describe a more general UE-DTA problem with any increasing-continuous travel cost functions, including the average travel time function shown in Eq.* (6.8).

In the literature, Larsson et al. (1999) shows that the capacitated UE solutions also satisfy the Wardrop equilibrium in terms of generalized route travel costs. Similarly, this cost is defined and proved below in the context of stochastic dynamic networks and ISUE.

Let $\mathcal{L}(\boldsymbol{f}, \boldsymbol{\pi}, \boldsymbol{\mu})$ define the Lagrangian function, i.e.,

$$\mathcal{L}(\boldsymbol{f}, \boldsymbol{\pi}, \boldsymbol{\mu}) = F(\boldsymbol{f}) + \sum_{r,s,y_t} \pi_{rs,y_t}\big(D^{(rs)}_{y_t} - \sum_{p,y_t} f_{p,y_t}\big) + \sum_{i\in\mathbb{I}} \mu_i c_i(\boldsymbol{f}) \tag{6.45}$$

where $\boldsymbol{\pi}$ and $\boldsymbol{\mu}$ are multipliers associated to the constraints in the problem $\mathcal{P}_0$.

Definition 10 (Generalized route travel cost)**.** *For a given solution $\boldsymbol{f}^*$ and multiplier $\boldsymbol{\mu}^*$, let's denote*

$$\mathcal{F}_{p,y_t} = T_{p,y_t} + \frac{1}{P_{y_t}} \min_{\varphi\in\Phi^t_p, x\in\mathcal{X}} \sum_{i\in\mathbb{I}} \mu^*_i \frac{\partial c_i(\boldsymbol{f}^*)}{\partial f_{\varphi|x}}. \tag{6.46}$$

$\mathcal{F}_{p,y_t}$ is called the generalized route travel cost for f_{p,y_t} regarding to demand $D^{(rs)}_{y_t}$.

Proposition 10 (The generalized Wardrop equilibrium (GWE) for DTA)**.** *The solutions to the problem $\mathcal{P}_0$ are Wardrop equilibrium flows for the generalized route travel cost, that*

means

$$f^*_{p,y_t} > 0 \Rightarrow \mathcal{F}_{p,y_t} = \frac{\pi^*_{rs,y_t}}{P_{y_t}} \qquad \forall p \in \mathbb{P}^{(rs)} \tag{6.47}$$

$$f^*_{p,y_t} = 0 \Rightarrow \mathcal{F}_{p,y_t} \geq \frac{\pi^*_{rs,y_t}}{P_{y_t}} \qquad \forall p \in \mathbb{P}^{(rs)}. \tag{6.48}$$

Proof. See Appendix C.3. □

Note that, GWE does not require the assumption A1, meaning that ISUE is GWE but not vice versa. In the next section, we study in detail both UE and SO objective functions and shows the relation between them.

6.3 The relation between ISUE and ISSO

In this section, we expand the expression of ISUE objective as the function of route travel time in the problem $\mathcal{P}_0$ and derive the relation between ISUE and ISSO objectives. Furthermore, we show that by using the marginal travel cost function $\tilde{T}_{p,t|x}(u)$ instead of the average travel cost function $T_{p,t|x}(u)$, $\mathcal{P}_0$ becomes ISSO-DTA problem.

6.3.1 Evaluation of ISUE objective

In this section, we first present the calculation of travel cost in a given scenario $x \in \mathcal{X}$. The path flow $f_{p,t|x}$ could be divided into a set of flows $\boldsymbol{f}_{p,t|x} = \{f_{p,t,h|x} : h \in \mathbb{T}, h \geq t\}$, where $f_{p,t,h|x}$ is a partial flow of $f_{p,t|x}$, arriving the destination at time h ($h > t$). Let's denote the accumulation of $f_{p,t,h|x}$ for any $h \leq k$ as $g_{p,t,k|x} = \sum_{h=t}^{k} f_{p,t,h|x}$ $\forall k \in \mathbb{N}, k \geq t$. Proposition 11 shows the evaluation of ISUE objective for a given path flow traffic.

Proposition 11. *ISUE objective function*

$$\begin{aligned} F(\boldsymbol{f}) &= \sum_{x\in\mathcal{X}} \sum_{p\in\mathbb{P}} \sum_{t\in\mathbb{T}} \rho_x \int_0^{f_{p,t|x}} T_{p,t|x}(u)\, \mathrm{d}u \\ &= \underbrace{\sum_{p,t,x} \sum_{h>t} \rho_x (h-t) f_{p,t,h|x}}_{F_T(\boldsymbol{f})} - \underbrace{\sum_{p,t,x} \sum_{h>t} \rho_x g_{p,t,h|x} \big(\ln f_{p,t|x} - \ln g_{p,t,h|x} \big)}_{E(\boldsymbol{f})} \end{aligned} \tag{6.49}$$

Proof. See C.4. □

In Eq. (6.49), $F_T(\boldsymbol{f})$ is the total system travel time defined below in Eq. (6.50), and $E(\boldsymbol{f})$ is the difference between ISUE and ISSO objective function.

6.3.2 The ISSO objective as the marginal equilibrium

The system optimal objective is defined by minimizing the total travel time $F_T(\boldsymbol{f})$, which is defined below:

$$F_T(\boldsymbol{f}) = \sum_{p,t,x} \sum_{h>t} \rho_x (h-t) f_{p,t,h|x}. \tag{6.50}$$

In Proposition 12, we show that, Eq. (6.41) could be used to define both ISUE and ISSO objectives by using different definitions of travel cost, i.e., marginal travel time Eq. (6.6) for SO and the average travel time Eq. (6.8) for UE.

Proposition 12. *The function Eq.* (6.41) *in problem* $\mathcal{P}_0$ *represents the ISSO objective by considering the marginal travel time function, i.e.,*

$$\tilde{F}(\boldsymbol{f}) = \sum_{p,t,x} \int_0^{f_{p,t|x}} \rho_x \tilde{T}_{p,t|x}(u) \mathrm{d}u = F_T(\boldsymbol{f}). \tag{6.51}$$

Proof. In the previous section, we have proved that, problem $\mathcal{P}_0$ provides ISUE solutions. Similarly, by applying Eq. (6.6) into this problem, we get:

$$\begin{aligned}
& \int_{g_{i-1}}^{g_i} \tilde{T}_{p,t|x}(u)\, \mathrm{d}u && = i(g_i - g_{i-1}) = i f_{p,t,t+i|x} && \forall i \geq 1 \\
\Rightarrow \quad & \int_0^{f_{p,t|x}} \tilde{T}_{p,t|x}(u)\, \mathrm{d}u && = \sum_{i \geq 1} \int_{g_{i-1}}^{g_i} \tilde{T}_{p,t|x}(u)\, \mathrm{d}u = \sum_{i \in \mathbb{T}} i f_{p,t,t+i|x} \\
\Rightarrow \quad & \tilde{F}(\boldsymbol{f}) && = \sum_{p,t,x} \int_0^{f_{p,t}} \rho_x \tilde{T}_{p,t|x}(u) \mathrm{d}u = \sum_{p,t,x} \sum_{h \geq t} \rho_x (h-t) f_{p,t,h} = F_T(\boldsymbol{f})
\end{aligned}$$

Therefore, solving problem $\mathcal{P}_0$ with the definition of travel cost in Eq. (6.6) gives ISSO solutions. □

Applying the ISUE definition to ISSO solutions, the marginal travel cost of any used paths are not smaller than the marginal cost of any unsaturated paths. In the case that all traffic exits the network at the end of time horizon, the ISSO objective is equivalent to the maximum aggregation of throughput at the destination, proved in Proposition 13.

Proposition 13. *If all traffic exits the network at the end of time horizon then*

$$F_T(\boldsymbol{f}) = C - H(\boldsymbol{f}) \tag{6.52}$$

where

$$C = \sum_{x \in \mathcal{X}} \sum_{t \in \mathbb{T}} \delta \rho_x (T + 1 - t) D_{t|x}^{(rs)} \tag{6.53}$$

$$H(\boldsymbol{f}) = \sum_{x \in \mathcal{X}} \sum_{t \in \mathbb{T}} \delta \rho_x (T + 1 - t) u_{s,t|x} \tag{6.54}$$

and $u_{s,t|x}$ is the flow entering the destination s at time t in scenario x.

Proof. See Appendix C.5. □

The Proposition 13 shows an efficient way to achieve ISSO solutions by only considering the traffic flow entering the destination. Note that, these variables are also supported by the linked-based traffic model described in Section 6.2.1. Moreover, using the original ISSO objective Eq. (6.50) requires the implicit assumption (or constraint) that all demands finish at destinations, otherwise, the minimization of $F_T(\boldsymbol{f})$ would result in a zero total travel time. For these reasons, we use $H(\boldsymbol{f})$ in the ISSO-DTA model in this work.

Proposition 14. *Given the same information evolution, let $\boldsymbol{f}^S$ denote the ISSO solution, and $\boldsymbol{f}^S_{w.o.i}$ denote the system optimal solution without considering the information evolution in the routing decision. Similarly, $\boldsymbol{f}^U$, $\boldsymbol{f}^U_{w.o.i}$ for ISUE with and without using information, respectively. By comparing these solutions under $H(\boldsymbol{f})$, we have the following order:*

$$H(\boldsymbol{f}^S) \geq H(\boldsymbol{f}^S_{w.o.i}) \geq H(\boldsymbol{f}^U_{w.o.i}) \tag{6.55}$$

$$H(\boldsymbol{f}^S) \geq H(\boldsymbol{f}^U). \tag{6.56}$$

Note that, the larger value of $H(\boldsymbol{f})$ means the smaller value of the total travel time (according to Eq. (6.52)*).*

Proof. The $\boldsymbol{f}^S_{w.o.i}$ is solved under the same problem defined for the ISSO solution $\boldsymbol{f}^S$ except for the constraints Eqs. (6.32) and (6.33) where $y_t = (\mathcal{X}, t)$ at any time t. Due to $\mathcal{X}_{y_t} \subseteq \mathcal{X}$, the feasible domain of $f^S_{w.o.i}$ is more restricted than f^S, therefore $H(\boldsymbol{f}^S) \geq H(\boldsymbol{f}^S_{w.o.i})$.

For $f^U_{w.o.i}$, the feasible solution domain is more restricted than $f^S_{w.o.i}$ due to the explicit FIFO constraint Eq. (6.29) even though the information-based constraints Eqs. (6.32) and (6.33) are the same. Furthermore, the objective function directly optimizes $H(\boldsymbol{f})$ in $\boldsymbol{f}^S_{w.o.i}$, therefore $H(\boldsymbol{f}^S_{w.o.i}) \geq H(\boldsymbol{f}^U_{w.o.i})$. Similarly, we have $H(\boldsymbol{f}^S) \geq H(\boldsymbol{f}^U)$ in the case of utilising information updates for route choice. □

The Proposition 14 proves mathematically the impact of information under two situations, i.e., with and without using information, in two typical route choices, i.e., SO and UE. It is well-known in the literature that in some networks, the UE solution could be as good as SO solution, meaning that $\boldsymbol{f}^U$ could be as good as $\boldsymbol{f}^S$ (for example the parallel-two-route network in Hall (1996)) and possibly outperform $\boldsymbol{f}^S_{w.o.i}$ or $\boldsymbol{f}^U_{w.o.i}$. Oppositely via numerical result in Section 6.5, we show an example that $\boldsymbol{f}^U_{w.o.i}$ is better than $\boldsymbol{f}^U$ in terms of $H(\boldsymbol{f})$ (or total system travel time). All of these suggest that the use of updated information from individual perspective does not always improve the system performance.

6.3.3 The solution existence and model complexity

For the existence of solutions, we refer to the discussion in Section 4.2.3. Similarly, we also propose the incremental method to solve the ISUE solution via a series of ISSO solutions. Therefore, the ISUE solution is achievable if there is an ISSO solution for each ISSO-DTA sub-problem.

Table 6.1 summarises the proposed DTA models and includes the complexity for each model. The ISUE-DTA problem, called $\mathcal{P}_{UE}$, requires the path-flow variables $f_{p,t,h|x}$ with double time indexes. Due to this extra variables and constraints, the complexity of $\mathcal{P}_{UE}$ in terms of the number of constraints and/or variables is proportional to $O(|\mathbb{A}||\mathbb{P}|T^2|\mathcal{X}|)$. In contrast, the ISSO-DTA problem, called $\mathcal{P}_{SO}$, only

TABLE 6.1: Summary of DTA models (Notation: $\surd$ for Used, and 'X' for Unused).

Objective or constraints	UE-DTA ($\mathcal{P}_{UE}$)	SO-DTA ($\mathcal{P}_{SO}$)	Type	# constraints or # variables
UE (Eq. (6.49))	$\surd$	X	Non-linear	-
SO (Eq. (6.54))	X	$\surd$	Linear	-
Link-based LTM model				
Eqs. (6.11)–(6.22)	X	$\surd$	Linear	$O(\lvert\mathbb{A}\rvert T\lvert\mathcal{X}\rvert)$
Path-based LTM model				
Eqs. (6.11)–(6.31)	$\surd$	X	Non-linear	$O(\lvert\mathbb{A}\rvert\lvert\mathbb{P}\rvert T^2\lvert\mathcal{X}\rvert)$
Information-based constraints				
Eqs. (6.32) and (6.33)	$\surd$	$\surd$	Linear	$O\left((\lvert\mathbb{A}\rvert+\lvert\mathbb{P}\rvert)T\lvert\mathcal{X}\rvert\right)$

requires path-based decision variables at source, therefore its complexity is significantly smaller than $\mathcal{P}_{UE}$ and proportional to $O\left((|\mathbb{A}|+|\mathbb{P}|)T|\mathcal{X}|\right)$ which avoids the double time indexes as in $\mathcal{P}_{UE}$.

Furthermore, the problem $\mathcal{P}_{SO}$ is linear in comparison with the non-linearity of $\mathcal{P}_{UE}$, therefore, it is much more efficient to solve $\mathcal{P}_{SO}$ than $\mathcal{P}_{UE}$. In the next section, we utilise the relation between ISUE and ISSO objectives to develop a so-called incremental solution method to compute the ISUE solutions based on solving a series of related ISSO-DTA problems instead of directly solving $\mathcal{P}_{UE}$.

6.4 Solution method

In the previous sections, we develop DTA models which are represented in the discrete time. In this section, we first show that the difference between ISSO and ISUE objective functions, i.e., $E(\boldsymbol{f})$, reduces according to the length of a time step. Based on this observation, we develop the ISM method to solve these models effectively.

In fact, the time domain in the proposed models are discretised from the continuous time horizon $[0,\mathcal{T}]$ with the time interval δ, i.e., $\mathbb{T}=\{1,2,3,...,T\}$ where $T=\dfrac{\mathcal{T}}{\delta}$. In continuous time, let $d_x^{(rs)}(\tau)$ denote a time-dependent demand function, showing the arrival rate at source r to destination s at time $\tau\in[0,\mathcal{T}]$ in scenario x. The relation between $D_{t|x}^{(rs)}$ (in discrete time domain) and $d_x^{(rs)}(\tau)$ (in continuous time domain) reads,

$$D_{t|x}^{(rs)}=\int_{\delta(t-1)}^{\delta t} d_x^{(rs)}(\tau)\mathrm{d}\tau. \tag{6.57}$$

Therefore, the same total traffic demand D in both the time domains can be expressed as

$$D = \sum_{t \in \mathbb{T}} D_{t|x}^{(rs)} = \int_0^{\mathcal{T}} d_x^{(rs)}(\tau) \mathrm{d}\tau. \tag{6.58}$$

Note that, the t index is for discrete time, and τ index is for continuous time.

Similarly, in a capacitated network, the link flow capacity $Q_{a,t|x}$ is also discretized from the continuous-time link capacity $q_{a|x}(\tau)$ where

$$Q_{a,t|x} = \int_{\delta(t-1)}^{\delta t} q_{a|x}(\tau) \mathrm{d}\tau \tag{6.59}$$

for all $a \in \mathbb{A}, t \in \mathbb{T}, x \in \mathcal{X}$. Given functions $d_x^{(rs)}(\tau)$ and $q_{a|x}(\tau)$, the parameters $D_{t|x}^{(rs)}$ and $Q_{a,t|x}$ are function of the time interval δ. We rewrite in this section the relation between ISUE and ISSO objective in Eq. (6.49) with the appearance of the time interval δ as follows,

$$F(\boldsymbol{f}) = \underbrace{\sum_{p,t,x} \sum_{h>t} \delta \rho_x (h-t) f_{p,t,h|x}}_{F_T(\boldsymbol{f})} - \underbrace{\sum_{p,t,x} \sum_{h>t} \delta \rho_x g_{p,t,h|x} \left(\ln f_{p,t|x} - \ln g_{p,t,h|x} \right)}_{E(\boldsymbol{f})}. \tag{6.60}$$

6.4.1 Linear approximation of ISUE objective

In the previous section, we show the computation of ISUE objective that includes the linear formulation of ISSO function $F_T(\boldsymbol{f})$ and the non-linear function $E(\boldsymbol{f})$. In Lemma 4, we show that the difference between ISUE and ISSO objective, i.e., $E(\boldsymbol{f})$, approaches zero at the infinitesimal system time step.

Lemma 4. *Let $N_{p,t|x}$ denote the number of non-zero elements in the set $\boldsymbol{f}_{p,t|x}$. If $N_{p,t|x}$ is bounded by a constant N (independent from δ, p and t and x), then*

(i) $0 \leq E(\boldsymbol{f}) \leq \dfrac{\delta N}{e} \sum_{x,t} \rho_x D_{t|x}^{(rs)}$.

(ii) $\lim_{\delta \to 0} E(\boldsymbol{f}) = 0$.

Proof. See Appendix C.6. □

Remark 7. *From Eqs.* (6.57) *and* (6.59)*, both the discrete demand* $D_{t|x}^{(rs)}$ *and the link flow capacity* $Q_{a,t|x}$ *are approximately proportional to the value of time interval* δ*. The value of* $N_{p,t|x}$ *is normally proportional to* $D_{t|x}^{(rs)}$ *and inversely proportional to* $Q_{a,t|x}$*, then the impact of* δ *to* $N_{p,t}$ *is destructive, that means* N *is likely independent from* δ *(and* p, t *or* x*). Similarly, Carey (2009) observed the time-independence of* $N_{p,t|x}$*, where each time halving the length of the time intervals, the number of time-space links in the fans (i.e.,* $N_{p,t|x}$ *in our study) remains much the same in the study of UE solutions. From this, by applying Lemma 4, we have* $\lim_{\delta \to 0} E(\boldsymbol{f}^*) = 0$*, where* $\boldsymbol{f}^*$ *is an ISUE solution.*

Note that, $F_T(\boldsymbol{f})$ *is linear while* $E(\boldsymbol{f})$ *is non-linear. By reducing the time step and ignoring the term* $E(\boldsymbol{f})$ *in the approximation of* $F(\boldsymbol{f})$*, we are able to gain more detail of traffic propagation in the network, and at the same time lessen the computational complexity via solving a linear formulation. Nevertheless, the FIFO principle implicitly requires the term* $E(\boldsymbol{f})$ *in the computation of path flow variables and still has to be maintained, which will guarantee the requirement of the FIFO principle in the obtained ISUE solutions.*

6.4.2 The ISM framework

Based on the insights from the previous section, we propose herein an incremental approach, called the incremental solution method (ISM), to obtain the ISUE solution for the ISUE-DTA problem.

The incremental procedure: By using a sufficiently small time step and based on the Lemma 4 and Remark 4, the $E(\boldsymbol{f})$ term in the ISUE objective function (6.49) can be omitted which results in a linear system optimal objective. The incremental procedure also guarantees the FIFO relation by imposing the FIFO requirement at any links in each scenario. The detail method is described below.

Step 0 Choose the time step δ, and set $t^* = 1$.

Step 1 Given the solution for previous demands ($\{D_{h|x}^{(rs)} : \forall x \in \mathcal{X}, \forall h \in \mathbb{T}, h < t^*\}$): $\hat{\boldsymbol{f}} = \{\hat{f}_{ab,t|x}, \forall a, b \in \mathbb{A}, t \in \mathbb{T}, x \in \mathcal{X}\}$, find an ISSO solution for demand at time t^* by solving the following linear problem:

$$[\mathcal{P}_1] \qquad \max_{f} H(\boldsymbol{f})$$

$$\text{s.t. :} \quad \text{Link-based model: Eqs. (6.11)–(6.22) and (6.32)}$$

$$f_{ab,t|x} \geq \hat{f}_{ab,t|x} \qquad a,b \in \mathbb{A}, t \in \mathbb{T}, x \in \mathcal{X} \tag{6.61}$$

$$\sum_{h \leq t - \frac{L_a}{V_{a|x}}} u_{a,h|x} - \sum_{h \leq t} v_{a,h|x} \geq \sum_{h \leq t - \frac{L_a}{V_{a|x}}} \hat{u}_{a,h|x} - \sum_{h \leq t} \hat{v}_{a,h|x} \quad a \in \mathbb{A}, t \in \mathbb{T}, x \in \mathcal{X} \tag{6.62}$$

$$\sum_{t \in \mathbb{T}} f_{ab,t|x} = \sum_{t \in \mathbb{T}} f_{ab,t|x'} \qquad a,b \in \mathbb{A}, x, x' \in \mathcal{X}_{y_{t^*}} \tag{6.63}$$

where $\hat{u}_{a,h|x}, \hat{v}_{a,h|x}$ are parameters computed from $\hat{f}_{ab,t|x}$ by Eqs. (6.17) and (6.18).

Step 2 If there is no more demand after t^*, then the last solution is the ISUE solution. Otherwise, update $t^* = t^* + 1$, and jump to step 1.

Note that, the problem $\mathcal{P}_1$ is the *link-based linear* model with the new constraints Eqs. (6.61)–(6.63) replacing all path-based constraints in the original problem $\mathcal{P}_0$. In particular Eq. (6.29) in $\mathcal{P}_0$ is replaced with Eqs. (6.61) and (6.62) in the ISM framework to maintain the FIFO principle, while Eq. (6.63) is equivalent to the information-based constraint Eq. (6.33) for demand at time t^* $(t^* > t)$. Note that the total traffic demands departing up to time t^* and choosing path p in $\mathcal{P}_1$ is written as

$$\sum_{t \leq t^*} f_{p,t|x} = \sum_{t \in \mathbb{T}} f_{ab,t|x}$$

for any $a, b \in p$ and $x \in \mathcal{X}$. As a result, the complexity of $\mathcal{P}_1$ reduces to $O\left(|\mathbb{A}|T|\mathcal{X}|\right)$ which is considerably less compared to that of the models $\mathcal{P}_{SO}$ and $\mathcal{P}_{UE}$ in Table 6.1.

Study of convergence: Given a solution $\boldsymbol{f}^*$ generated from this framework, we could evaluate $E(\boldsymbol{f}^*)$ in Eq. (6.49) to assess the goodness of the resulting ISUE solution. Based on Lemma 4, the smaller value of $E(\boldsymbol{f}^*)$, the better the solution $\boldsymbol{f}^*$ approximates the true ISUE value. We show in the numerical results (Section 6.5) the solution convergence by reducing the time step δ. Furthermore, the second example for large network demonstrates the efficiency of using higher time-scale solution for the computation at the smaller time scale.

6.5 Numerical results

The numerical results are studied in two different network sizes: the small-sized simple network and the large-sized Fort-Worth network. In the simple network, we demonstrate the impact of information updates to user equilibrium route choices. For the large network, the performance of our method is measured in terms of modelling complexity (i.e., a number of constraints and variables) and computational time. All results are computed in the desktop computer (Intel Core i5 3.2GHz, 8GB RAM) by the open-source solver COIN-OR CBC (Forrest, 2012) for LP problems.

6.5.1 The simple network

The network topology and its parameters are shown in Fig. 6.1 and Table 6.2. The traffic demand departs from node R at rate 2 veh/s during the first 120 seconds (Fig. 6.2). In this example, we generate two scenarios x_0 ($\rho_0 = 0.2$) and x_1 ($\rho_1 = 0.8$) to represent the stochastic link capacity on link (2,4) shown in Fig. 6.2. With the probability $\rho_1 = 0.8$, link (2,4) drops its flow capacity from time $t = 30$ to 120. For other links, they operate normally without any disruption in both scenarios. According to this setting, the evolution of information update is shown below,

- Before time $t = 30$ (s), travellers consider two possibilities x_0 and x_1 in their decisions of route choices.
- At and after time $t = 30$ (s), travellers realise either x_0 or x_1, therefore, their decisions are specially made for the realised scenarios at presence.

The dynamic and stochastic network settings impact on route choices in one of the three paths: P0 (1,3,4), P1 (1,2,3,4), and P2 (1,2,4). In the non-congested network state, paths P0 and P2 are shortest with 120-second travel time, while path P1 takes 150 seconds. However, the total flow capacity on both P0 and P2 (1.33 veh/s) is smaller than the traffic demand (2 veh/s) entering the network, therefore it causes travel delay in these paths. If the delay is large enough, a portion of traffic will traverse P1. Furthermore, the uncertainty of link (2,4) also alters the route choice by travellers to avoid the long delay on the link (2,4) in the scenario x_1.

In this simple example, the impact of information is studied in three cases:

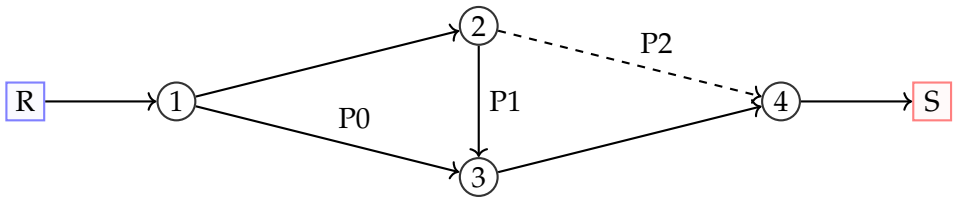


FIGURE 6.1: The simple network topology with source R and sink S.

TABLE 6.2: Configuration for each link in Fig. 6.1.

Links	**L** (m)	**V** (m/s)	**W** (m/s)	**K** (veh/m)	**Q** (veh/s)	Free-flow travel time (s)
(1,2), (3,4)	1000	16.67	8.33	0.24	1.33	60
(1,3), (2,4)	1000	16.67	8.33	0.12	0.67	60
(2,3)	500	16.67	8.33	0.12	0.67	30

- Full information: Every travellers know the full information (at any point in time) of traffic demand and network capacity before entering the network. Note that in this case, the solution is in fact the standard UE solution for each scenario.

- Real-time information: The information of network state and demand *up to* the current time is available (i.e., up-to-date) to all travellers.

- No real-time information: There is no information update as time passes, however, travellers are aware of the distribution of all scenarios based on their past experience.

For the last two cases, the optimal solutions in terms of path flow splits and travel cost are shown in Fig. 6.3. In the right figure, without updating the real-time information, we observe the same decision of splitting traffic over the three paths in

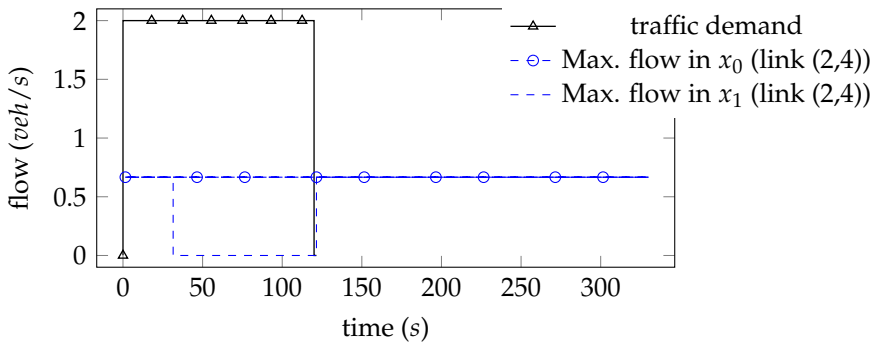


FIGURE 6.2: The dynamic setting for network in Fig. 6.1.

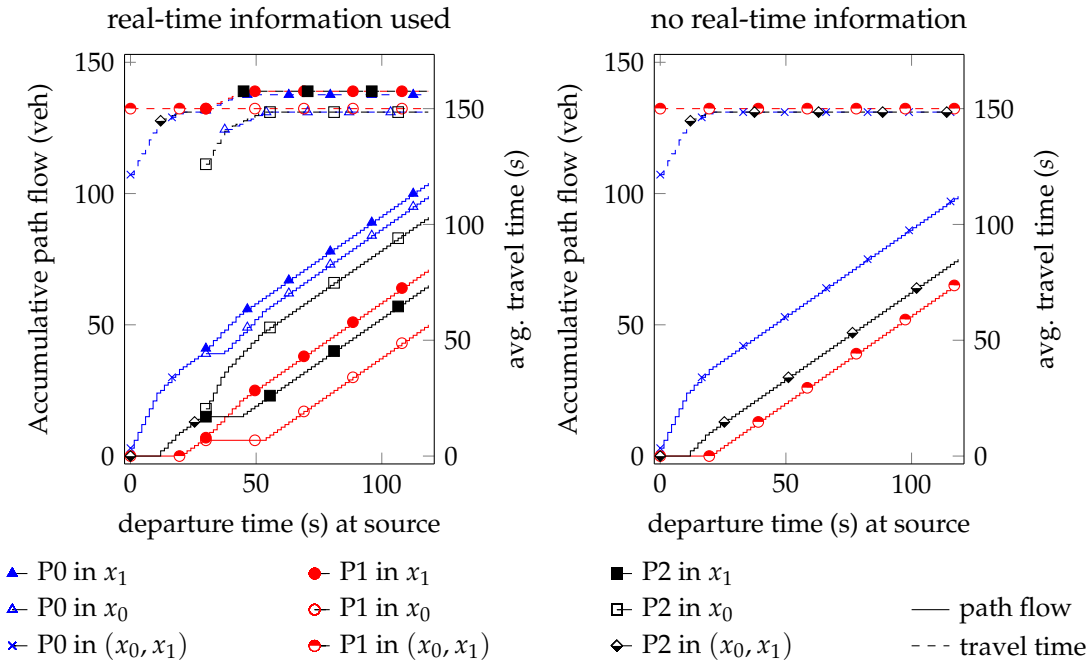


FIGURE 6.3: Travel time on each path in the UE solution for the simple network

both x_0 and x_1. After time $t = 20$, the UE travel cost is 150 seconds which is also the free-flow travel time on P1.

More adaptively, the information updates enable travellers from time $t = 30$ to adapt better according to the revealed scenarios as shown in the left figure of Fig. 6.3. There are two different flow splits after this time. In particular, there is more traffic on path P2 and less on path P1 in scenario x_1 than in x_0 to better utilize the link (2,4) which possesses some uncertainty. Hence, depending on the received information, the ISUE travel costs are stabilized in two different points instead of one as in the case of no information update.

The impact of information update on different types of route choices is shown in Table 6.3. Although the real-time information is useful to have a more adaptive route choice, the first row in Table 6.3 shows an increase in average travel time compared to the no-information case. The reason is that, without a real-time update, it is too risky to choose path P2 at the beginning, therefore more traffic selects paths P0 and P1, that improves the system performance.

In Table 6.3, by considering the average of travel time on three paths, the solution with the real-time update is close to the full information solution while the no-information solution is considerably similar to the SO solution over paths P1 and

TABLE 6.3: The impact of information on the system performance.

	Full information	Real-time update	No information	System optimal
Overall avg. travel time (s)	149.9	150.5	149.6	143.8
Avg. cost (s) on P0 for x_0, x_1	140.2, 148.5	144.8, 148.3	146.5, 146.5	121.3, 133.7
Avg. cost (s) on P1 for x_0, x_1	150.0, 154.1	150.3, 153.8	151.4, 151.4	150.0, 153.5
Avg. cost (s) on P2 for x_0, x_1	140.4, 151.8	142.0, 155.2	120.0, 160.5	121.2, 162.6
UE MSRE (s) for x_0, x_1	0.0, 0.0	7.2, 0.5	9.1, 0.4	16.4, 13.1
SO MSRE (s) for x_0, x_1	16.4, 13.1	18.0, 13.0	12.4, 13.1	0.0, 0.0

P2. Let's consider the full information and SO solutions as the baselines, and define the distance of any solutions to these baseline solutions as the mean square root error (MSRE) of travel cost in each scenario x, formulated below:

$$MSRE_x = \sqrt{\frac{\sum_{t\in\mathbb{T}} D_{t|x}(T_{t|x} - T^*_{t|x})^2}{\sum_{t\in\mathbb{T}} D_{t|x}}} \tag{6.64}$$

where $D_{t|x}$ is the total traffic demand at time t in scenario x, i.e., $D_{t|x} = D^{(rs)}_{t|x}$ in the single-O-D network. The quantity $T_{t|x}$ is the average travel time for demand at time t in scenario x, and $T^*_{t|x}$ is for the baseline solution. The last two rows in Table 6.3 shows that real-time update of information could help to improve route choice to the ideal (full information) UE choices while no information pushes the ISUE solution to the system optimal point in this example.

As discussed in the previous section that we able to obtain UE solutions by reducing the time step δ, and the value of $E(\boldsymbol{f})$ approaches zero as $\delta \to 0$. In Fig. 6.4, we show the results of $E(\boldsymbol{f})$ and $H(\boldsymbol{f})$ in different time scales from 1.5 to 30 seconds. It indicates that while $H(\boldsymbol{f})$ slowly reduces with respect to δ, the value of $E(\boldsymbol{f})$ drops quickly toward zero and is much smaller than $H(\boldsymbol{f})$, i.e., at $\delta = 1.5(s)$, $\frac{E(\boldsymbol{f})}{H(\boldsymbol{f})} = 0.0014$.

In the simple network, we study ISUE solutions in different configurations of information updates and show the solution convergence in terms of $E(\boldsymbol{f})$. In the next example, our method is applied to the large network to evaluate its modelling

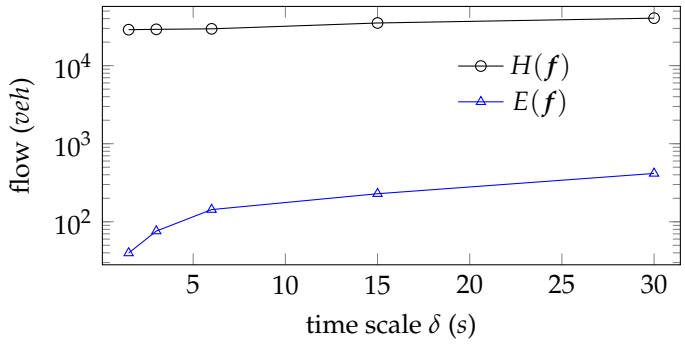


FIGURE 6.4: The evaluation of $E(f)$ in the simple network.

complexity and computational performance.

6.5.2 Large network

For a real-sized problem, we perform our method on the Fort-Worth network (Mahmassani, 2001), which includes 182 nodes, 441 links shown in Fig. 6.5. We consider two scenarios where the capacity drop happens on randomly chosen links in one of the scenarios. The total amount of traffic demand is 500 vehicles, moving from the source R to the destination S. In this example, only the last 250 vehicles are received real-time information that enables them to determine the current network state belonging to a particular scenario.

Table 6.4 shows the complexity and computational time by solving the problem in different time scales from 15 seconds to 60 seconds. Halving the time unit doubles the number of constraints $\mathcal{C}$ and variables $\mathcal{V}$, however, the model preparation time $T_{prepare}$ and the solution time T_{sovle} increase considerably. Note that, the time $T_{prepare}$ is essential only for the first time of solving a problem, that takes into account the model generation for a given set of parameters. Without or minor changes of parameters, this time is negligible for the later use. On the other hand, the solving time $\mathcal{T}_{solve}$ is the time to obtain the optimal solution by solving the generated model everytime a solution is sought.

For solving this problem at 15-second time unit, it takes more than 24 hours to be completed without any initial solutions used in the warmed-start process. To improve $\mathcal{T}_{solve}$, we reuse the solution at 30-second time unit as the initial point before solving the problem. The warmed-start process improves significantly the solution time as shown in Table 6.4. This demonstrates the significant practical applicability

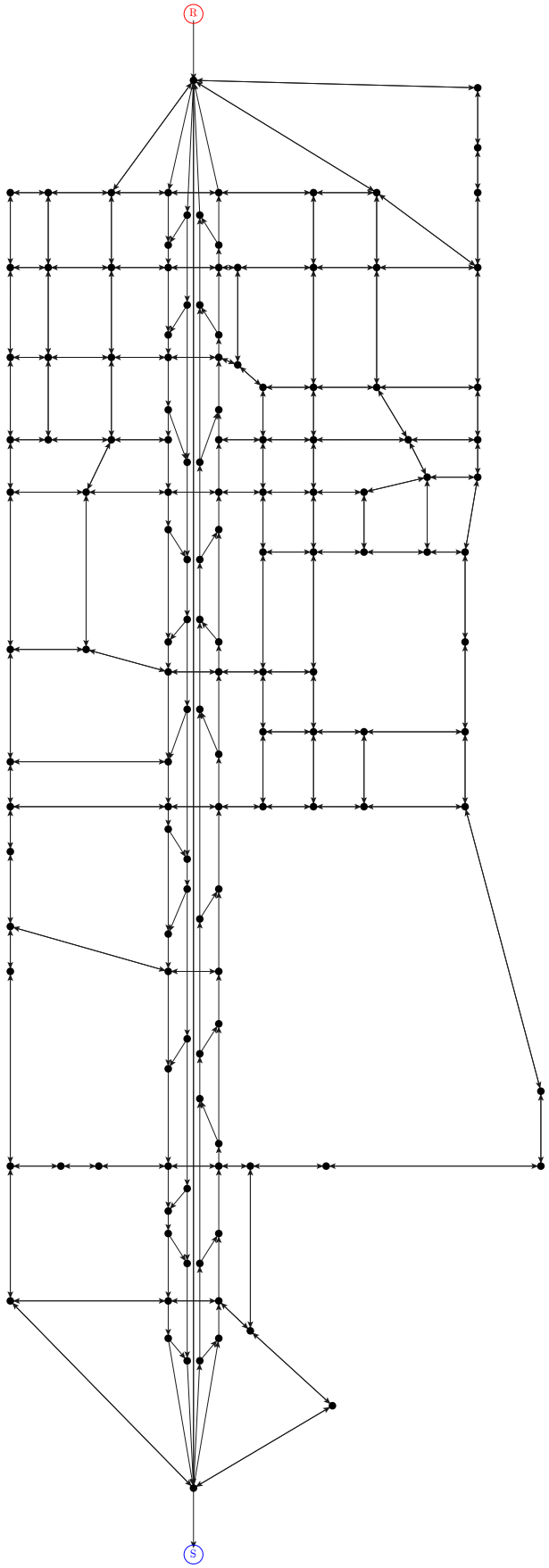


FIGURE 6.5: Fort-Worth network.

TABLE 6.4: Results for Fort-Worth network.

Time unit δ (s)	$\mathcal{C}$	$\mathcal{V}$	$\mathcal{T}_{prepare}(s)$	$\mathcal{T}_{solve}(s)$	$H(\boldsymbol{f})$
60	73K	61K	60	17	276.14K
30	109K	116K	350	8900	171.55K
15 (warmed start)	218K	232K	2000	800	164.55K

of our method that enables the use of higher time scale solutions for a more detailed and accurate solution at the smaller time scales in real-time operation.

6.6 Summary

This chapter develops a comprehensive linear mathematical framework to study the benefit of real-time information and the impact of resulting user adaptive route choice behaviours on network performance. The framework formulates the information-based stochastic user equilibrium (ISUE) dynamic traffic assignment (DTA) problem for a single origin-destination network. Using the framework, it proves the linkage between the ISUE and ISSO solutions underpinned by the FIFO principle. This important property then enables us to develop an incremental loading method to obtain the ISUE solutions efficiently by solving a sequence of linear programs. Moreover, the proposed method is more scalable that avoids a huge enumeration of paths in large-scale networks as done in path-based methods of the existing literature on this topic. The numerical examples show the impact of information to both route choices and network performance, and demonstrate the significant improvements in the obtained ISUE solution both in terms of accuracy and computational complexity.

Chapter 7

Conclusion

7.1 Summary

This thesis proposes macroscopic DTA models for the study of real-time information in the transportation network. It provides an insight understanding of the interaction among route choice, network state (including signal control) and information. The impact of real-time information has been studied from both system's and user's point of view. Especially for UE route choices, novel solution methods are developed for the proposed objective-based mathematical programming models which are in contrast with the existing constraint-based models in the literature.

DTA SO (Chapter 3) This chapter presents an optimisation framework to find SO solutions which can achieve better traffic throughput by optimally distributing the congestion over links inside the network. This framework formulates the Two regime Transmission Model (TTM), an extended version of the Link Transmission Model (LTM), as a set of side linear constraints for a SO problem. Via numerical results, they show the capability of this framework to achieve minimal spillbacks and less complexity in comparison with the CTM-based SO-DTA model.

DTA UE (chapter 4) This chapter first proposes a mathematical programming framework for the UE-DTA problem in a single O-D capacitated network. It then develops an incremental solution method (ISM) to efficiently obtain the UE solution by solving a sequence of linear problems. This ISM method is underpinned by an important theoretical result proven in this chapter that in the limit, the SO objective can be used to obtain the UE solution as the system time step goes to zero given the satisfaction of

the FIFO constraint. Via numerical results, the incremental method provides significant improvements in the UE solution both in terms of accuracy and computational complexity.

DTA SO Info (Chapter 5) This chapter develops a novel information model describing the evolution of knowledge in the form of the real-time information that accounts for the (stochastic) uncertainty in demand and network capacity. It shows that the proposed information model can be effectively integrated into the mathematical framework based on the cell transmission model (CTM) for the SO-DTA problem. Using the framework, we demonstrate the significant value of the real-time information via real-time strategy model that includes two types of route choice: path choice and policy choice. The former enables temporal adjustments, while the latter allows both temporal and spatial adaptations. To this end, it proposes and develops two linear programming models: PoSCTM for policy choice, and PaSCTM for path choice for DTA SO problems in the context of a stochastic time-dependent network. The numerical results show that, for system optimality, PoSCTM solutions are always better than (or at least equal to) that of PaSCTM in both objective values and computational performance. Besides, the policy-based approach can replace path-based approach to solve multi-destination DTA SO problems without changing the objective value but providing much lower computational cost.

DTA UE Info (Chapter 6) This chapter first proposes a novel mathematical programming framework for the information-based stochastic UE-DTA problem in a single O-D dynamic stochastic network. We then develop an incremental solution method (ISM) to efficiently obtain the ISUE solution by solving a sequence of linear ISSO-DTA problems. Our proposed method for this problem extends the ISM method in Chapter 4 for stochastic traffic flow. We prove that in the limit, the information-based SO objective can be used to obtain the ISUE solution as the system time step goes to zero given the satisfaction of the FIFO constraint. Via numerical results, our incremental method shows significant improvements in finding global optimal ISUE solution both in terms of accuracy and computational complexity. Even though our method requires a small system time unit for the convergence

of solutions, we demonstrate that the use of large-time-scale solution (lower computational time) can efficiently solve the small-time-scale problem as presented in our example using a realistic large network.

7.2 Future research

While most of the current models developed in this thesis only consider a single O-D pair, our mathematical programming framework is extensible for solving a more general network, that includes multiple commodities, different types of travel cost, etc. It paves a way to develop sophisticated models to study different aspects in the combination of real-time information, user behaviour and system control. In the vision to future development in transportation, the following directions are valuable as the potential extensions of this work.

Future of transportation In recent years, we observe the social attraction of the new generation in the automotive industry, i.e., driver-less vehicles, electric vehicles and personal hybrid car-drones. In Australia, the New South Wales state government is currently testing the driver-less buses for public uses. The future vehicles are expected to provide not only self-driving but also multiple functions, from collecting traffic data, providing assistant information, to becoming an artificially intelligent entity (communicating with both human and other vehicles). They provide new transport options for all travellers, however, raise a question about the applicability of previous transport models with these new types of vehicles.

Similarly, there are signs of a new generation of traffic management. We currently see the application of new technologies in signal controls, variable message signs (VMS), surveillance cameras, etc. It is expected that they would help to adaptively manage the transport system in real-time. While DTA framework mainly studies the traffic flow in the network, it can be potentially extended to incorporate with various control strategies above. Furthermore, the traffic management evolves slowly, from fixed to dynamic rules, but for each link in the network. Thanks to the communication technology, it is possible now that the network operator would communicate directly and individually to each (or group of) vehicles. In this case,

each vehicle would receive a unique set of instructions during its journey which might differ from one vehicle to another. This futuristic vision certainly changes the way of making traffic policies as its dimensions include not only space, time but also vehicles.

Enhancements of information model The vehicle and GPS data, including the trajectory data either from floating cars or connected and automated vehicles, are increasingly available. It certainly improves the quality of information, i.e., providing a more accurate estimation of OD matrix and the joint distribution of demand and supply. Based on the models developed in this work, more realistic features could be added to reflect the process of information propagation, including the error probability and spatial coverage. The understanding of these properties currently is within the scope of communication technologies, and few works have partially studied its impact on transport applications.

In the past, the role of information was limited. Nowadays, it increasingly influences and contributes to the normal travellers' behaviour in many developed countries. In the future, the interaction among information providers will further increase its role in shaping mobility. There is a need for better information models not only to enhance system performance but also to investigate new mobility management options, e.g., a study of advance intersection control using a wireless communication system.

The choice models This thesis currently focuses on the choice of routes in the DTA framework. Therefore, the models and their solution methods are extensible to include the other dimensions of choices, for example, the departure time choice as seen in the literature. Additionally, the trend of sharing vehicles (e.g., bike or cars) and autonomous vehicles could be popular in the future. As a result, a new dimension of choices emerges such as a multi-modal transport. Studying its impact on the system performance is thus a promising future research direction that could be based on the work developed in this thesis.

In conclusion, this thesis develops novel analytical frameworks for the study of DTA problem with real-time information. Based on this work, a variety of extensions

could be performed as the future works in different aspects, i.e., traffic model, travel choice model and information model, that is also influenced by the development of future transportation modes and technologies.

Appendix A

Appendix for Chapter 4

A.1 Satisfaction of assumption A1 for any constraints in Section 4.1.2

We recall the definition of LTM-based variables based on trip flows below,

$$
\begin{aligned}
f_{ab,t} &= \sum_{\substack{\varphi \in \Phi_{b,t} \\ \exists (h,a) \in \varphi, h<t}} f_\varphi \\
u_{a,t} &= \sum_{\varphi \in \Phi_{a,t}} f_\varphi \\
v_{a,t} &= \sum_{b \in \Upsilon_a^+} f_{ab,t} = \sum_{b \in \Upsilon_a^+} \sum_{\substack{\varphi \in \Phi_{b,t} \\ \exists (h,a) \in \varphi, h<t}} f_\varphi \\
v_{a,h,t} &= \sum_{b \in \Upsilon_a^+} \sum_{\varphi \in \Phi_{b,t} \cap \Phi_{a,h}} f_\varphi .
\end{aligned}
$$

Upper-bound of flow capacity $c(\boldsymbol{f}) = u_{a,t} - Q_{a,t} \leq 0$. We easily see that this constraint satisfy the assumption A1, because

$$
\frac{\partial c(\boldsymbol{f})}{\partial f_\varphi} = \begin{cases} 1 & \text{if } \varphi \in \Phi_{a,t} \\ 0 & \text{otherwise.} \end{cases}
$$

We get similar result for this constraint at downstream flow, i.e., $c(f) = v_{a,t} - Q_{a,t} \leq 0$.

Constraint of downstream flow limited by upstream flow (forward direction)

$$c(\boldsymbol{f}) = \sum_{h \le t - \frac{L_a}{V_a}} u_{a,h} - \sum_{h \le t} v_{a,h} \le 0.$$

Because of the free-flow travel time L_a/V_a, we able to set $f_\varphi = 0$ for any $\varphi \in \Phi_a^t \cap \Phi_b^h$ where $b \in \Upsilon_a^+$, $h \le t - L_a/V_a$. For any $\varphi \in \Phi_{a,h} \cap \Phi_{b,k}$, where $b \in \Upsilon_a^+$:

- If $h \le t - L_a/V_a$ and $k \le t$: $\dfrac{\partial c(\boldsymbol{f})}{\partial f_\varphi} = \dfrac{\partial u_{a,h}}{\partial f_\varphi} - \dfrac{\partial v_{a,k}}{\partial f_\varphi} = 0.$
- If $h \le t - L_a/V_a$ and $k > t$: $\dfrac{\partial c(\boldsymbol{f})}{\partial f_\varphi} = 1.$
- If $h > t - L_a/V_a$ and $k \le t$: Since $f_\varphi = 0$, it becomes a constant or parameter in the model.
- If $h > t - L_a/V_a$ and $k > t$: $\dfrac{\partial c(\boldsymbol{f})}{\partial f_\varphi} = 0.$

Constraint of upstream flow limited by downstream flow (backward direction)

$$c(\boldsymbol{f}) = \sum_{h \le t} u_{a,h} - \sum_{h \le t - \frac{L_a}{W_a}} v_{a,h} - K_a L_a \le 0.$$

For any $\varphi \in \Phi_{a,h} \cap \Phi_{b,k}$,where $b \in \Upsilon_a^+$:

- If $h \le t$, and $k \le t - \dfrac{L_a}{W_a}$: $\dfrac{\partial c(\boldsymbol{f})}{\partial f_\varphi} = \dfrac{\partial u_{a,h}}{\partial f_\varphi} - \dfrac{\partial v_{a,k}}{\partial f_\varphi} = 0.$
- If $h \le t$, and $k > t - \dfrac{L_a}{W_a}$: $\dfrac{\partial c(\boldsymbol{f})}{\partial f_\varphi} = \dfrac{\partial u_{a,h}}{\partial f_\varphi} = 1.$
- If $h > t$, and $k \le t - \dfrac{L_a}{W_a}$: Since $k < h$, $f_\varphi = 0$ in this case, it is not considered as a variable in the model.
- If $h > t$, and $k > t - \dfrac{L_a}{W_a}$: $\dfrac{\partial c(\boldsymbol{f})}{\partial f_\varphi} = 0.$

FIFO constraint

$$\sum_{p \in \mathbb{P}} f_{a,p,t^+,h} \sum_{p \in \mathbb{P}} f_{a,p,t,h^+} \le 0 \qquad \forall a \in \mathbb{A}, \quad \forall h^+ > h, \quad \forall t^+ > t.$$

Since FIFO constraint is the product of two flow variables where each variable is the accumulation of trip flows f_φ, therefore, its derivation on each trip flow is always larger or equal to 0.

A.2 Proof of Proposition 2

Let $\mathcal{L}(\boldsymbol{f}, \boldsymbol{\pi}, \boldsymbol{\mu})$ denote the Lagrangian function with the definition in Eq. (4.38). It is known that, if $\boldsymbol{f}^*$ is the minimum solution of $\mathcal{P}_0$, then there exists multipliers $\boldsymbol{\mu}^*$ and $\boldsymbol{\pi}^*$ such that,

$$f_\varphi^* \frac{\partial \mathcal{L}(\boldsymbol{f}^*, \boldsymbol{\pi}^*, \boldsymbol{\mu}^*)}{\partial f_\varphi} = 0 \quad \Leftrightarrow \quad f_\varphi^* \Big(\frac{\partial F(\boldsymbol{f})}{\partial f_\varphi} + \sum_{i \in \mathbb{I}} \mu_i^* \frac{\partial c_i(\boldsymbol{f}^*)}{\partial f_\varphi} - \pi_{rs,t}^* \Big) = 0 \quad \forall f_\varphi \in \Phi_p^t \tag{A.1}$$

$$\frac{\partial \mathcal{L}(\boldsymbol{f}^*, \boldsymbol{\pi}^*, \boldsymbol{\mu}^*)}{\partial f_n} \geq 0 \quad \Leftrightarrow \quad \frac{\partial F(\boldsymbol{f})}{\partial f_\varphi} + \sum_{i \in \mathbb{I}} \mu_i^* \frac{\partial c_i(\boldsymbol{f}^*)}{\partial f_\varphi} - \pi_{rs,t}^* \geq 0 \quad \forall f_\varphi \in \Phi_p^t \tag{A.2}$$

$$\mu_i^* c_i(\boldsymbol{f}^*) = 0 \qquad \forall i \in \mathbb{I} \tag{A.3}$$

$$\mu_i^* \geq 0 \qquad \forall i \in \mathbb{I}. \tag{A.4}$$

These equations are called the Karush-Kuhn-Tucker (KKT) necessary condition for minimum solutions. Note that, $\frac{\partial F(\boldsymbol{f})}{\partial f_\varphi} = \frac{\partial F(\boldsymbol{f})}{\partial f_{p,t}} \frac{\partial f_{p,t}}{\partial f_\varphi} = T_{p,t} \; \forall f_\varphi \in \Phi_p^t$. From Eqs. (A.1) and (A.2), we get

$$f_{p,t}^* > 0 \Rightarrow T_{p,t}^* = \pi_{rs,t}^* - \frac{1}{f_{p,t}} \sum_{i \in \mathbb{I}} \mu_i^* \underbrace{\sum_{f_\varphi \in \Phi_p^t} f_\varphi \frac{\partial c_i(\boldsymbol{f}^*)}{\partial f_\varphi}}_{\tilde{c}_i(\boldsymbol{f}^*)}.$$

$$f_{p,t}^* = 0 \Rightarrow T_{p,t}^* \geq \pi_{rs,t}^* - \min_{f_\varphi \in \Phi_p^t} \sum_{i \in \mathbb{I}} \mu_i^* \frac{\partial c_i(\boldsymbol{f}^*)}{\partial f_\varphi}.$$

If path p is unsaturated, then from A1 and Eqs. (4.35) and (A.3), we infer that there exists $f_\varphi \in \Phi_p^t$ so that either $\mu_i^* = 0$ or $\frac{\partial c_i(\boldsymbol{f}^*)}{\partial f_\varphi} = 0 \; \forall i \in \mathbb{I}$. Therefore, from Eq. (A.2), $T_{p,t}^* \geq \pi_{rs,t}^*$ for any unsaturated path p. For any used route, the first equation above shows the opposition that $T_{p,t}^* \leq \pi_{rs,t}^*$ (from the assumption A1, $\tilde{c}_i(\boldsymbol{f}) \geq 0$). Hence, $\pi_{rs,t}^*$ is the route travel cost of any used and unsaturated paths. It is also the lower bound cost of any unused and unsaturated paths. As a result, $\boldsymbol{f}^*$ satisfies the CUE principle (Eq. (4.37)).

A.3 Proof of Proposition 3

In this proof, we also apply the KKT conditions presented in A.2. For the internal flow equilibrium Eqs. (4.43) and (4.44), they are easily derived from Eqs. (A.1) and (A.2), i.e.,

$$\begin{aligned} f_{\varphi}^{*} > 0 &\Rightarrow \mathcal{F}_{p,t}^{\varphi} = \pi_{rs,t}^{*} && \forall \varphi \in \Phi_{p}^{t} \\ f_{\varphi}^{*} = 0 &\Rightarrow \mathcal{F}_{p,t}^{\varphi} \geq \pi_{rs,t}^{*} && \forall \varphi \in \Phi_{p}^{t}. \end{aligned}$$

Furthermore, due to the identical equilibrium value for any positive trip flow, the Wardrop conditions Eqs. (4.41) and (4.42) are proved, that is,

- If $f_{p,t}^{*} = 0$: $\mathcal{F}_{p,t} = \min\limits_{\varphi \in \Phi_{p}^{t}} \mathcal{F}_{p,t}^{\varphi} \geq \pi_{rs,t}^{*}$.
- If $f_{p,t}^{*} > 0$: $\mathcal{F}_{p,t}^{\varphi} = \min\limits_{\varphi \in \Phi_{p}^{t}} \mathcal{F}_{p,t}^{\varphi} = \min \left\{ \min\limits_{\varphi \in \Phi_{p}^{t}: f_{\varphi} > 0} \mathcal{F}_{p,t}^{\varphi}, \min\limits_{\varphi \in \Phi_{p}^{t}: f_{\varphi} = 0} \mathcal{F}_{p,t}^{\varphi} \right\} = \pi_{rs,t}^{*}$. Note that, it exists at least a trip flow $f_{\varphi} > 0$ ($\exists \varphi \in \Phi_{p}^{t}$) in this case.

A.4 Proof of Proposition 6

From Eqs. (4.17) and (4.50), we infer that:

$$\lim_{\delta \to 0} D_{t}^{(rs)} = 0 \Rightarrow \lim_{\delta \to 0} f_{p,t} = 0.$$

Since the demand function $d^{(rs)}(t)$ is given, we also have the conservation of total demand:

$$\sum_{p \in \mathbb{P}} \sum_{t \in \mathbb{T}} f_{p,t} = \sum_{rs} \int_{0}^{\mathcal{T}} d^{(rs)}(u) \mathrm{d}u = D$$

where D is the constant of total demand. In general, we can prove that,

$$0 \leq g_{p,t,t+i} \ln f_{p,t} - g_{p,t,t+i} \ln g_{p,t,t+i} \leq \frac{f_{p,t}}{e}$$

for any $g_{p,t,t+i} \in [0, f_{p,t}]$. Applying this boundary to $E(\boldsymbol{f})$ in Eq. (4.53), we get:

$$0 \leq E(\boldsymbol{f}) \leq \frac{\delta}{e} \sum_{p,t} \sum_{N_{p,t}} f_{p,t} = \frac{\delta N}{e} \sum_{rs,t} D_t^{(rs)} \leq \frac{\delta N D}{e} \xrightarrow{\delta \to 0} 0.$$

For the limitation of $E(\boldsymbol{f})$, we achieve: $\lim_{\delta \to 0} E(\boldsymbol{f}) = 0$.

Appendix B

Appendix for Chapter 5

B.1 Proofs of Proposition 7

If the network is clear at the end time, minimising total system cost is equivalent to maximise F.

Proof. Consider traffic flow approaching a destination s at time t, we have:

$$\sum_{i\in\Gamma_s^-} y_{is,t|x} = \sum_{\forall t_0<t} y_{s,t_0,t|x}$$

$$n_{s,t|x} = \sum_{h\le t}\sum_{i\in\Gamma_s^-} y_{is,h|x} = \sum_{h\le t}\sum_{t_0<h} y_{s,t_0,h|x}$$

$$\Rightarrow \sum_{h\le t} n_{s,h|x} = \sum_{h\le t}\sum_{i\in\Gamma_s^-} (t+1-h) y_{is,h|x}$$

Because of the clear network at the ending time T, we get:

$$\sum_{t>t_0} y_{s,t_0,t|x} = \sum_{r\in\mathbf{C}_R} D_{t_0|x}^{(rs)} = \text{Constant} \qquad \forall t_0\in\mathbb{T}, s\in\mathbf{C}_S, x\in\mathcal{X}$$

$$\sum_{s\in\mathbf{C}_S} n_{s,T|x} = \sum_{t\in\mathbb{T}}\sum_{\forall rs} D_{t|x}^{(rs)} = \text{Constant} \qquad \forall t\in\mathbb{T}, x\in\mathcal{X}$$

The total system cost at a specific scenario x is transformed to:

$$\begin{aligned} C_x &= \sum_{s\in\mathbf{C}_S}\sum_{t_0\in\mathbb{T}}\sum_{t>t_0}(t-t_0) y_{s,t_0,t|x} = \sum_{s\in\mathbf{C}_S}\sum_{t_0\in\mathbb{T}}\sum_{t>t_0} t.y_{s,t_0,t|x} - \sum_{t_0\in\mathbb{T}} t_0 \sum_{\forall rs} D_{t_0|x}^{(rs)} \\ &= \sum_{s\in\mathbf{C}_S}\sum_{t\in\mathbb{T}} t \sum_{i\in\Gamma_s^-} y_{is,t|x} - \sum_{t\in\mathbb{T}}\sum_{\forall rs} t.D_{t|x}^{(rs)} \\ &= (T+1)\sum_{s\in\mathbf{C}_S}\sum_{t\in\mathbb{T}}\sum_{i\in\Gamma_s^-} y_{is,t|x} - \sum_{s\in\mathbf{C}_S}\sum_{t\in\mathbb{T}}\sum_{i\in\Gamma_s^-}(T+1-t) y_{is,t|x} - \sum_{t\in\mathbb{T}}\sum_{\forall rs} t.D_{t|x}^{(rs)} \end{aligned}$$

$$= (T+1)\sum_{s\in\mathbb{C}_S} n_{s,T|x} - \sum_{s\in\mathbb{C}_S}\sum_{t\in\mathbb{T}} n_{s,t|x} - \sum_{t\in\mathbb{T}}\sum_{\forall rs} tD_{t|x}^{(rs)}$$

$$= \underbrace{\sum_{t\in\mathbb{T}}\sum_{\forall rs}(T+1-t)D_{t|x}^{(rs)}}_{\text{Constant}} - \sum_{s\in\mathbb{C}_S}\sum_{t\in\mathbb{T}} n_{s,t|x}$$

Thus, $\min TS = \min \sum_{x\in\mathcal{X}} q_x C_x \Leftrightarrow \max F = \sum_{x\in\mathcal{X}}\sum_{s\in\mathbb{C}_S}\sum_{t\in\mathbb{T}} q_x n_{s,t|x}$. □

B.2 Proofs of Lemma 3

In a non-stochastic (or a single scenario) network, for any feasible PoSCTM solution, there exists an appropriate PaSCTM solution, and vice versa.

Proof. The relationship between policy-based and path-based models are presented in Eqs. (5.15), (5.16), (5.25) and (5.26). In the context of a single scenario, we ignore the index of scenario x to simplify the representation of variables. We observe that, any feasible PaSCTM solutions can be converted to feasible PoSCTM solutions which satisfy all policy-based constraints. In contrast, we get a feasible PoSCTM solution $\{n_{i,t}^{(-s)}, y_{ij,t}^{(-s)}\}$, and find an appropriate PaSCTM solution $\{n_{i,p,t}, y_{ij,p,t}\}$, by solving the system:

$$\sum_{\substack{p\in\mathbb{P}\\ i,s\in p}} n_{i,p,t} = n_{i,t}^{(-s)} \qquad \forall i\in\mathbb{C}, t\in\mathbb{T}, s\in\mathbb{C}_S \qquad \text{(B.1)}$$

$$\begin{aligned} n_{i,p,t+1} &= n_{i,p,t} + y_{bi,p,t} - y_{ie,p,t} \\ &= n_{i,p,1} + \sum_{h\le t} y_{bi,p,h} - \sum_{h\le t} y_{ie,p,h} \end{aligned} \qquad \forall b,i,e\in\mathbb{C}_I, t\in\mathbb{T}, p\in\mathbb{P} \qquad \text{(B.2)}$$

$$\begin{aligned} n_{s,p,t+1} &= n_{s,p,t} + y_{bs,p,t} \\ &= n_{s,p,1} + \sum_{h\le t} y_{bs,p,h} \end{aligned} \qquad s\in\mathbb{C}_S; t\in\mathbb{T}; b\in\Gamma_i^-; b,s\in p \qquad \text{(B.3)}$$

$$\begin{aligned} n_{r,p,t+1} &= n_{r,p,t} + f_{p,t} - y_{ri,p,t} \\ &= n_{r,p,1} + \sum_{h\le t} f_{p,h} - \sum_{h\le t} y_{ri,p,h} \end{aligned} \qquad r\in\mathbb{C}_R; t\in\mathbb{T}; i\in\Gamma_r^+; r,i\in p \qquad \text{(B.4)}$$

$$\sum_{p\in\mathbb{P}^{(rs)}} f_{p,t} = D_t^{(rs)} \qquad r\in\mathbb{C}_R; s\in\mathbb{C}_S; t\in\mathbb{T} \qquad \text{(B.5)}$$

$$\sum_{\substack{p\in\mathbb{P}\\ ij,s\in p}} y_{ij,p,t} = y_{ij,t}^{(-s)} \qquad \forall (i,j)\in\mathbb{A}, t\in\mathbb{T}, s\in\mathbb{C}_S \qquad \text{(B.6)}$$

$$n_{i,p,t} \geq y_{ie,p,t} \overset{Eq.(B.2)}{\Longleftrightarrow} \quad n_{i,p,1} + \sum_{h \leq t} y_{bi,p,h} \geq \sum_{h \leq t+1} y_{ie,p,h}$$

$$\forall b, i, e \in \mathbb{C}, t \in \mathbb{T}, p \in \mathbb{P} \qquad \text{(B.7)}$$

$$n_{r,p,t} \geq y_{ri,p,t} \overset{Eq.(B.4)}{\Longleftrightarrow} \quad n_{r,p,1} + \sum_{h \leq t} f_{p,h} \geq \sum_{h \leq t+1} y_{ri,p,h}$$

$$\forall r \in \mathbb{C}_R, i \in \mathbb{C}_I, t \in \mathbb{T}, p \in \mathbb{P} \qquad \text{(B.8)}$$

From the above system, the independent variables are only $n_{i,p,1}$ (only at time 1) and $y_{ij,p,t}$ for any $i \in \mathbb{C}, (i,j) \in \mathbb{A}, p \in \mathbb{P}, t \in \mathbb{T}$. We prove this Lemma by

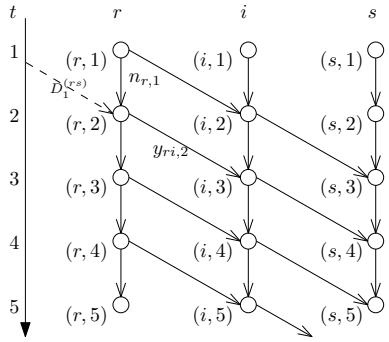


FIGURE B.1: Time-expanded cell network.

proposing a method of assigning policy flow to path flow. Let's consider the path flow sequence: $p_t = \{(ij, t_i) | \forall (i,j) \in p, t_j > t_i\}$, where (ij, t_i) represents a path flow y_{ij,p,t_i} from cell i to cell j at time t_i. Consistently, if cell j is the downstream neighbour of cell i, then $t_j > t_i$. In fact, this path flow p_t is a feasible path in time-expanded cell network, like in Fig. B.1. In this expansion, the flow $y_{ij,t}$ becomes flow from node (i,t) to node $(j, t+1)$ in the expanded network, and the number of vehicles on cell i ($n_{i,t}$) becomes flow from node $(i, t-1)$ to node (i,t). There is no arc at the same or upper time level. By building this new network, we allow no-queue at every nodes. The flow from node (i,t) to node $(j, t+1)$ is:

$$y^*_{ij,t} = \begin{cases} y_{ij,t} & \text{if} \quad i \neq j \\ n_{i,t} & \text{if} \quad i = j. \end{cases}$$

Similarly, the policy flow in this network is:

$$y^{*(-s)}_{ij,t} = \begin{cases} y^{(-s)}_{ij,t} & \text{if} \quad i \neq j \\ n^{(-s)}_{i,t} & \text{if} \quad i = j. \end{cases}$$

We observe that, it is safe to assign: $y_{ij,p,t_i} = f_{p,t_i} = \min_{(ij,t_i)\in p_t} y^{*(-s)}_{ij,t_i} \quad \forall (ij,t_i) \in p_t$. This assigned flow can be put aside by this transformation:

$$y^{(-s)}_{ij,t_i} \quad = \quad y^{(-s)}_{ij,t_i} - f_{p,t_i} \qquad \forall (ij,t_i) \in p_t$$

By doing this, Eqs. (B.7) and (B.8) are not violated. Furthermore, if at destination s, $y^{(-s)}_{is,t} = 0 \quad \forall t \in \mathbb{T}$, then we simply do not consider link (i,s) in any paths going through (i,s), which help us to shrink down the network size. Having transformed, the problem becomes smaller, because at least one policy-based arc flow is clear. Repeat these steps above until we assign all path flows. □

Appendix C

Appendix for Chapter 6

C.1 Increasing constraint functions in the DTA model

Definitions of LTM-based variables based on trip flows:

$$f_{ab,t|x} = \sum_{\substack{\varphi \in \Phi_{b,t} \\ \exists (h,a) \in \varphi, h<t}} f_{\varphi|x}$$

$$u_{a,t|x} = \sum_{\varphi \in \Phi_{a,t}} f_{\varphi|x}$$

$$v_{a,t|x} = \sum_{b \in \Upsilon_a^+} f_{ab,t|x} = \sum_{b \in \Upsilon_a^+} \sum_{\substack{\varphi \in \Phi_{b,t} \\ \exists (h,a) \in \varphi, h<t}} f_{\varphi|x}$$

$$v_{a,h,t|x} = \sum_{b \in \Upsilon_a^+} \sum_{\varphi \in \Phi_{b,t} \cap \Phi_{a,h}} f_{\varphi|x}$$

Upper-bound of flow capacity $c(\boldsymbol{f}) = u_{a,t|x} - Q_{a,t|x} \leq 0$. We easily see that this constraint satisfy the assumption A1, because

$$\frac{\partial c(\boldsymbol{f})}{\partial f_{\varphi|x}} = \begin{cases} 1 & \text{if } \varphi \in \Phi_{a,t} \\ 0 & \text{otherwise.} \end{cases}$$

We get similar result for this constraint at downstream flow, i.e., $c(\boldsymbol{f}) = v_{a,t|x} - Q_{a,t|x} \leq 0$.

Constraint of downstream flow limited by upstream flow (forward direction)

$$c(\boldsymbol{f}) = \sum_{h \leq t - \frac{L_a}{V_{a|x}}} u_{a,h|x} - \sum_{h \leq t} v_{a,h|x} \leq 0.$$

Because of the free-flow travel time $L_a/V_{a|x}$, we able to set $f_{\varphi|x} = 0$ for any $\varphi \in \Phi_a^t \cap \Phi_b^h$ where $b \in \Upsilon_a^+$, $h \leq t - L_a/V_{a|x}$. For any $\varphi \in \Phi_{a,h} \cap \Phi_{b,k}$, where $b \in \Upsilon_a^+$:

- If $h \leq t - L_a/V_{a|x}$ and $k \leq t$: $\dfrac{\partial c(\boldsymbol{f})}{\partial f_{\varphi|x}} = \dfrac{\partial u_{a,h|x}}{\partial f_{\varphi|x}} - \dfrac{\partial v_{a,k|x}}{\partial f_{\varphi|x}} = 0.$
- If $h \leq t - L_a/V_{a_x}$ and $k > t$: $\dfrac{\partial c(\boldsymbol{f})}{\partial f_{\varphi|x}} = 1.$
- If $h > t - L_a/V_{a|x}$ and $k \leq t$: Since $f_{\varphi|x} = 0$, it becomes a constant or parameter in the model.
- If $h > t - L_a/V_{a|x}$ and $k > t$: $\dfrac{\partial c(\boldsymbol{f})}{\partial f_{\varphi|x}} = 0.$

Constraint of upstream flow limited by downstream flow (backward direction)

$$c(\boldsymbol{f}) = \sum_{h \leq t} u_{a,h|x} - \sum_{h \leq t - \frac{L_a}{W_{a|x}}} v_{a,h|x} - K_a L_a \leq 0.$$

For any $\varphi \in \Phi_{a,h} \cap \Phi_{b,k}$,where $b \in \Upsilon_a^+$:

- If $h \leq t$, and $k \leq t - \dfrac{L_a}{W_{a|x}}$: $\dfrac{\partial c(\boldsymbol{f})}{\partial f_{\varphi|x}} = \dfrac{\partial u_{a,h|x}}{\partial f_{\varphi|x}} - \dfrac{\partial v_{a,k|x}}{\partial f_{\varphi|x}} = 0.$
- If $h \leq t$, and $k > t - \dfrac{L_a}{W_{a|x}}$: $\dfrac{\partial c(\boldsymbol{f})}{\partial f_{\varphi|x}} = \dfrac{\partial u_{a,h|x}}{\partial f_{\varphi|x}} = 1.$
- If $h > t$, and $k \leq t - \dfrac{L_a}{W_{a|x}}$: Since $k < h$, $f_{\varphi|x} = 0$ in this case, it is not considered as a variable in the model.
- If $h > t$, and $k > t - \dfrac{L_a}{W_{a|x}}$: $\dfrac{\partial c(\boldsymbol{f})}{\partial f_{\varphi|x}} = 0.$

FIFO constraint

$$v_{a,h^+,t|x}\, v_{a,h,t^+|x} \leq 0.$$

Since FIFO constraint is the product of two flow variables where each variable is the accumulation of trip flows $f_{\varphi|x}$, therefore, its derivation on each trip flow is always larger or equal to 0.

C.2 Proof of Proposition 9

Let $\mathcal{L}(\boldsymbol{f}, \boldsymbol{\pi}, \boldsymbol{\mu})$ define the Lagrangian function, i.e.,

$$\mathcal{L}(\boldsymbol{f}, \boldsymbol{\pi}, \boldsymbol{\mu}) = F(\boldsymbol{f}) + \sum_{y_k, rs} \pi_{y_k, rs}\big(D_{y_k}^{(rs)} - \sum_{p, y_k} f_{p, y_k}\big) + \sum_{i \in \mathbb{I}} \mu_i c_i(\boldsymbol{f}) \tag{C.1}$$

If $\boldsymbol{f}^*$ is the minimum solution of $\mathcal{P}_0$, then there exists multipliers $\boldsymbol{\mu}^*$ and $\boldsymbol{\pi}^*$ such that,

$$f^*_{\varphi|x} \frac{\partial \mathcal{L}(\boldsymbol{f}^*, \boldsymbol{\pi}^*, \boldsymbol{\mu}^*)}{\partial f_{\varphi|x}} = 0 \quad \Leftrightarrow \quad f^*_{\varphi|x}\big(\frac{\partial F(\boldsymbol{f})}{\partial f_{\varphi|x}} + \sum_{i \in \mathbb{I}} \mu_i^* \frac{\partial c_i(\boldsymbol{f}^*)}{\partial f_{\varphi|x}} - \pi^*_{y_k, rs}\big) = 0 \quad \forall \varphi \in \Phi \tag{C.2}$$

$$\frac{\partial \mathcal{L}(\boldsymbol{f}^*, \boldsymbol{\pi}^*, \boldsymbol{\mu}^*)}{\partial f_{\varphi|x}} \geq 0 \quad \Leftrightarrow \quad \frac{\partial F(\boldsymbol{f})}{\partial f_{\varphi|x}} + \sum_{i \in \mathbb{I}} \mu_i^* \frac{\partial c_i(\boldsymbol{f}^*)}{\partial f_{\varphi|x}} - \pi^*_{y_t, rs} \geq 0 \quad \forall \varphi \in \Phi \tag{C.3}$$

$$\mu_i^* c_i(\boldsymbol{f}^*) = 0 \qquad \forall i \in \mathbb{I} \tag{C.4}$$

$$\mu_i^* \geq 0 \qquad \forall i \in \mathbb{I} \tag{C.5}$$

These equations are called the Karush-Kuhn-Tucker (KKT) necessary condition for minimum solutions. Due to $\dfrac{\partial f_{p, y_t}}{\partial f_{\varphi|x}} = 1$ for all $\varphi \in \Phi_p^t$ (and equal to zero otherwise), the partial derivation of objective is equal to

$$\frac{\partial F(\boldsymbol{f})}{\partial f_{\varphi|x}} = T_{p, y_t} P_{y_t} \tag{C.6}$$

for all $\varphi \in \Phi_p^t$ and $r, s \in p$. From Eqs. (C.2) and (C.3), we get

$$f^*_{p, y_t} > 0 \Rightarrow P_{y_t} T_{p, y_t} = \pi^*_{y_t, rs} - \min_{\varphi \in \Phi_p^t, x \in \mathcal{X}} \sum_{i \in \mathbb{I}} \mu_i^* \frac{\partial c_i(\boldsymbol{f}^*)}{\partial f_{\varphi|x}} \tag{C.7}$$

$$f^*_{p, y_t} = 0 \Rightarrow P_{y_t} T_{p, y_t} \geq \pi^*_{y_t, rs} - \min_{\varphi \in \Phi_p^t, x \in \mathcal{X}} \sum_{i \in \mathbb{I}} \mu_i^* \frac{\partial c_i(\boldsymbol{f}^*)}{\partial f_{\varphi|x}}. \tag{C.8}$$

If path p is unsaturated, then from A1 and Eqs. (6.42) and (C.4), we infer that there exists $\varphi \in \Phi_p^t$ so that either $\mu_i^* = 0$ or $\dfrac{\partial c_i(\boldsymbol{f}^*)}{\partial f_{\varphi|x}} = 0 \; \forall i \in \mathbb{I}$. Therefore, from Eq. (C.3),

$T_{p,y_t} \geq \dfrac{\pi^*_{rs,y_t}}{P_{y_t}}$ for any unsaturated path p. For any used route, the first equation above shows the opposition that $T_{p,y_t}(f^*_{p,y_t}) \leq \pi^*_{rs,y_t}$ (from the assumption A1, $\tilde{c}_i(\boldsymbol{f}) \geq 0$). Hence, $\dfrac{\pi^*_{rs,y_t}}{P_{y_t}}$ is the route travel cost of any used and unsaturated paths. It is also the lower bound cost of any unused and unsaturated paths. As a result, $\boldsymbol{f}^*$ satisfies the ISUE principle (Eq. (6.44)).

C.3 Proof of Proposition 10

In this proof, we also apply the KKT conditions presented in C.2. Due to the identical equilibrium value for any positive trip flow shown in Eqs. (C.7) and (C.8), the Wardrop conditions Eqs. (6.47) and (6.48) are proved.

C.4 Proof of Proposition 11

According to the definition of $f_{p,t,h|x}$ and $g_{p,t,k|x}$, we have:

$$f_{p,t|x} = \sum_{h \geq t} f_{p,t,h|x}$$
$$g_{p,t,k|x} = \sum_{h=t}^{k} f_{p,t,h|x}.$$

We assume that, each flow has to take at least one time unit from source to destination, i.e., $f_{p,t,h|x} = 0$ for all $h \leq t$. For a flow $f_{p,t,h|x}$ ($h > t$), it takes exactly $(h-t)$ time units to arrive the destination. Initially, we have $g_{p,t,t|x} = 0$. Let $m = \max_{h \geq t}\{i : g_{p,t,h|x} = 0\}$ denote the time up to which none of the traffic departing at time t has arrived at the destination. For any positive path flow $f_{p,t|x} > 0$, the average travel time function $T_{p,t|x}(u)$ for an amount u of flow is computed by

$$\begin{aligned} T_{p,t|x}(u) &= \frac{\delta(h-t+1)(u - g_{p,t,h|x}) + \sum_{j=t+1}^{h} \delta(j-t) f_{p,t,j|x}}{u} \\ &= \frac{\delta(h-t+1)(u - g_{p,t,h|x}) + \sum_{j=t+1}^{h} \delta(j-t)(g_{p,t,j|x} - g_{p,t,j-1|x})}{u} \end{aligned}$$

for any $u \in [g_{p,t,h|x}, g_{p,t,h+1|x}]$ and $i \geq m$.

If $i \geq m$, the integration of $T_{p,t|x}(u)$ in the range $[g_{p,t,h|x}, g_{p,t,h+1|x}]$ becomes:

$$\int_{g_{p,t,h|x}}^{g_{p,t,h+1|x}} T_{p,t|x}(u)\mathrm{d}u = \int_{g_{p,t,h|x}}^{g_{p,t,h+1|x}} \frac{\delta(h-t+1)(u-g_{p,t,h|x}) + \sum_{j=t+1}^{h}\delta(j-t)f_{p,t,j|x}}{u}\mathrm{d}u$$

$$= \begin{cases} \delta(h-t+1)f_{p,t,h+1|x} - \sum_{j=t+1}^{h}\delta g_{p,t,j|x}\ln\frac{g_{p,t,h+1|x}}{g_{p,t,h|x}} & \text{if } h > m \\ \delta(h-t+1)f_{p,t,h+1|x} & \text{if } i = m. \end{cases}$$

From that, for the whole range $(0, f_{p,t|x})$, after some simple transformations, we get:

$$\int_0^{f_{p,t|x}} T_{p,t|x}(u)\mathrm{d}u = \sum_{h\geq m}\int_{g_{p,t,h|x}}^{g_{p,t,h+1|x}} T_{p,t|x}(u)\mathrm{d}u$$

$$= \delta\sum_{h>t}(h-t)f_{p,t,h|x} - \delta\sum_{h\geq t} g_{p,t,h|x}\big(\ln f_{p,t|x} - \ln g_{p,t,h|x}\big).$$

The computation of ISUE objective function follows:

$$F(\boldsymbol{f}) = \sum_{x\in\mathcal{X}}\sum_{p\in\mathbb{P}}\sum_{t\in\mathbb{T}}\rho_x\int_0^{f_{p,t|x}} T_{p,t|x}(u)\ \mathrm{d}u$$

$$= \sum_{p,t,x}\sum_{h>t}\delta\rho_x(h-t)f_{p,t,h|x} - \sum_{p,t,x}\sum_{h\geq t}\delta\rho_x g_{p,t,h|x}\big(\ln f_{p,t|x} - \ln g_{p,t,h|x}\big).$$

C.5 Proofs of Proposition 13

The total system cost at a specific scenario x is transformed to:

$$F_x = \sum_{p\in\mathbb{P}^{(rs)}}\sum_{t\in\mathbb{T}}\sum_{h>t}(h-t)f_{p,t,h|x} = \sum_{p\in\mathbb{P}}\sum_{t\in\mathbb{T}}\sum_{t>t_0} h.f_{p,t,h|x} - \sum_{t\in\mathbb{T}} t\sum_{\forall rs} D_{t|x}^{(rs)}$$

$$= \sum_{s\in\mathbb{A}_S}\sum_{t\in\mathbb{T}} t.u_{s,t|x} - \sum_{t\in\mathbb{T}}\sum_{\forall rs} tD_{t|x}^{(rs)}$$

$$= -\sum_{s\in\mathbb{A}_S}\sum_{t\in\mathbb{T}}(T+1-t)u_{s,t|x} + \sum_{s\in\mathbb{A}_S}\sum_{t\in\mathbb{T}}(T+1)u_{s,t|x} - \sum_{t\in\mathbb{T}}\sum_{\forall rs} tD_{t|x}^{(rs)}$$

$$= \underbrace{\sum_{t\in\mathbb{T}}\sum_{\forall rs}(T+1-t)D_{t|x}^{(rs)}}_{\text{Constant}} - \sum_{s\in\mathbb{A}_S}\sum_{t\in\mathbb{T}}(T+1-t)u_{s,t|x}$$

Therefore, $F_T(\boldsymbol{f}) = C - H(\boldsymbol{f})$.

C.6 Proof of Lemma 4

Based on Eqs. (6.27) and (6.57), we infer that:

$$\lim_{\delta \to 0} D_{k|x}^{(rs)} = 0 \Rightarrow \lim_{\delta \to 0} f_{p|x}^{k} = 0$$

Since the demand function $d^{(rs)}(t)$ is given, the total demand is conserved:

$$\sum_{p \in \mathbb{P}} \sum_{t \in \mathbb{T}} f_{p,t|x} = \sum_{rs} \int_{0}^{\mathcal{T}} d_{x}^{(rs)}(u) \mathrm{d}u = D_x$$

where D_x is the constant of total demand in scenario x. In general, we can prove that, for any $g_{p,t,h|x} \in [0, f_{p,t|x}]$:

$$0 \leq g_{p,t,h|x} \ln f_{p,t|x} - g_{p,t,h|x} \ln g_{p,t,h|x} \leq \frac{f_{p,t|x}}{e}$$

Applying this boundary to $E(\boldsymbol{f})$, we get:

$$0 \leq E(\boldsymbol{f}) \leq \frac{\delta}{e} \sum_{p,t,x} \sum_{N_{p,t|x}} \rho_x f_{p,t|x} = \frac{\delta N}{e} \sum_{x} \rho_x D_x \xrightarrow{\delta \to 0} 0$$

For the limitation of $E(\boldsymbol{f})$, we achieve: $\lim_{\delta \to 0} E(\boldsymbol{f}) = 0$.

Bibliography

Ahmed, Afzal, Dong Ngoduy, and David Watling (2016). "Prediction of traveller information and route choice based on real-time estimated traffic state". In: *Transportmetrica B: Transport Dynamics* 4.1, pp. 23–47. DOI: 10.1080/21680566.2015.1052110.

Allsop, Richard E (1974). "Some possibilities for using traffic control to influence trip distribution and route choice". In: *Transportation and Traffic Theory, Proceedings*. Vol. 6.

Altman, Eitan and Laura Wynter (2004). "Equilibrium, games, and pricing in transportation and telecommunication networks". In: *Networks and Spatial Economics* 4.1, pp. 7–21.

Amaral, Rodrigo Rezende and El-Houssaine Aghezzaf (2015). "City Logistics and Traffic Management: Modelling the Inner and Outer Urban Transport Flows in a Two-tiered System". In: *Transportation Research Procedia* 6, pp. 297–312.

Arnott, Richard, André De Palma, and Robin Lindsey (1999). "Information and time-of-usage decisions in the bottleneck model with stochastic capacity and demand". In: *European Economic Review* 43.3, pp. 525–548.

Aubin, Jean-Pierre, Alexandre M Bayen, and Patrick Saint-Pierre (2011). *Viability Theory: New Directions, Systems & Control: Foundations & Applications*.

Balakrishna, Ramachandran, Haris Koutsopoulos, Moshe Ben-Akiva, Bruno Fernandez Ruiz, and Manish Mehta (2005). "Simulation-based evaluation of advanced traveler information systems". In: *Transportation Research Record: Journal of the Transportation Research Board* 1910, pp. 90–98.

Balakrishna, Ramachandran, Moshe Ben-Akiva, Jon Bottom, and Song Gao (2013). "Information Impacts on Traveler Behavior and Network Performance: State of Knowledge and Future Directions". In: *Advances in Dynamic Network Modeling in Complex Transportation Systems*. Springer, pp. 193–224.

Balijepalli, NC, D Ngoduy, and DP Watling (2013). "The two-regime transmission model for network loading in dynamic traffic assignment problems". In: *Transportmetrica A: Transport Science* ahead-of-print, pp. 1–22.

Ban, Xuegang Jeff, Jong-Shi Pang, Henry X Liu, and Rui Ma (2012). "Modeling and solving continuous-time instantaneous dynamic user equilibria: A differential complementarity systems approach". In: *Transportation Research Part B: Methodological* 46.3, pp. 389–408.

Bar-Gera, Hillel (2005). "Continuous and discrete trajectory models for dynamic traffic assignment". In: *Networks and Spatial Economics* 5.1, pp. 41–70.

Barceló, Jaume and Lídia Montero (2015). "A Robust Framework for the Estimation of Dynamic OD Trip Matrices for Reliable Traffic Management". In: *Transportation Research Procedia* 10, pp. 134–144.

Beard, Christopher and Athanasios Ziliaskopoulos (2006). "System optimal signal optimization formulation". In: *Transportation Research Record: Journal of the Transportation Research Board* 1978, pp. 102–112.

Beckmann, Martin, CB McGuire, and Christopher B Winsten (1956). "Studies in the Economics of Transportation". In: *Yale University Press, New Haven, Connecticut*.

Biegler, Lorenz T and Victor M Zavala (2009). "Large-scale nonlinear programming using IPOPT: An integrating framework for enterprise-wide dynamic optimization". In: *Computers & Chemical Engineering* 33.3, pp. 575–582.

Bierlaire, Michel and Frank Crittin (2004). "An efficient algorithm for real-time estimation and prediction of dynamic OD tables". In: *Operations Research* 52.1, pp. 116–127.

Bifulco, Gennaro N, Giulio E Cantarella, Fulvio Simonelli, and Pietro Velonà (2016). "Advanced traveller information systems under recurrent traffic conditions: Network equilibrium and stability". In: *Transportation Research Part B: Methodological* 92, pp. 73–87.

BITRE (2014). "Impact of road trauma and measures to improve outcomes (Report 140)". In: Bureau of Infrastructure, Transport and Regional Economics (BITRE), Canberra ACT. ISBN: 978-1-925216-10-3.

— (2017). "Costs and benefits of emerging road transport technologies (Report 146)". In: Bureau of Infrastructure, Transport and Regional Economics (BITRE), Canberra ACT. ISBN: 978-1-925531-37-4.

Bliemer, Michiel CJ and Piet HL Bovy (2003). "Quasi-variational inequality formulation of the multiclass dynamic traffic assignment problem". In: *Transportation Research Part B: Methodological* 37.6, pp. 501–519.

Brackstone, Mark and Mike McDonald (1999). "Car-following: a historical review". In: *Transportation Research Part F: Traffic Psychology and Behaviour* 2.4, pp. 181–196.

Braess, Dietrich, Anna Nagurney, and Tina Wakolbinger (2005). "On a paradox of traffic planning". In: *Transportation science* 39.4, pp. 446–450.

Bureau of Public Roads (1964). "Traffic Assignment Manual". In: *US Department of Commerce, Urban Planning Division*.

Carey, M (2009). "A framework for user equilibrium dynamic traffic assignment". In: *Journal of the Operational Research Society* 60.3, pp. 395–410.

Carey, Malachy (1992). "Nonconvexity of the dynamic traffic assignment problem". In: *Transportation Research Part B: Methodological* 26.2, pp. 127–133.

Carey, Malachy and YE Ge (2012a). "Comparison of methods for path flow reassignment for dynamic user equilibrium". In: *Networks and Spatial Economics* 12.3, pp. 337–376.

Carey, Malachy and David Watling (2012b). "Dynamic traffic assignment approximating the kinematic wave model: System optimum, marginal costs, externalities and tolls". In: *Transportation Research Part B: Methodological* 46.5, pp. 634–648.

Carey, Malachy, Hillel Bar-Gera, David Watling, and Chandra Balijepalli (2014). "Implementing first-in–first-out in the cell transmission model for networks". In: *Transportation Research Part B: Methodological* 65, pp. 105–118.

Chen, Owen and Moshe Ben-Akiva (1998). "Game-theoretic formulations of interaction between dynamic traffic control and dynamic traffic assignment". In: *Transportation Research Record: Journal of the Transportation Research Board* 1617, pp. 179–188.

Chorus, Caspar G, Eric JE Molin, and Bert Van Wee (2006). "Use and effects of Advanced Traveller Information Services (ATIS): a review of the literature". In: *Transport Reviews* 26.2, pp. 127–149.

Claudel, Christian G and Alexandre M Bayen (2010a). "Lax–Hopf based incorporation of internal boundary conditions into Hamilton–Jacobi equation. Part I: Theory". In: *IEEE Transactions on Automatic Control* 55.5, pp. 1142–1157.

— (2010b). "Lax–hopf based incorporation of internal boundary conditions into hamilton-jacobi equation. part ii: Computational methods". In: *IEEE Transactions on Automatic Control* 55.5, pp. 1158–1174.

Connors, Richard D and Agachai Sumalee (2009). "A network equilibrium model with travellers' perception of stochastic travel times". In: *Transportation Research Part B: Methodological* 43.6, pp. 614–624.

Correa, José R, Andreas S Schulz, and Nicolás E Stier-Moses (2004). "Selfish routing in capacitated networks". In: *Mathematics of Operations Research* 29.4, pp. 961–976.

Daganzo, Carlos F (1994). "The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory". In: *Transportation Research Part B: Methodological* 28.4, pp. 269–287.

— (1995). "The cell transmission model, part II: network traffic". In: *Transportation Research Part B: Methodological* 29.2, pp. 79–93.

— (2005). "A variational formulation of kinematic waves: basic theory and complex boundary conditions". In: *Transportation Research Part B: Methodological* 39.2, pp. 187–196.

De Palma, André, Robin Lindsey, and Nathalie Picard (2012). "Risk aversion, the value of information, and traffic equilibrium". In: *Transportation Science* 46.1, pp. 1–26.

Diakaki, Christina, Markos Papageorgiou, Ioannis Papamichail, and Ioannis Nikolos (2015). "Overview and analysis of Vehicle Automation and Communication Systems from a motorway traffic management perspective". In: *Transportation Research Part A: Policy and Practice* 75, pp. 147–165.

Doan, Kien and Satish V Ukkusuri (2015). "Dynamic system optimal model for multi-OD traffic networks with an advanced spatial queuing model". In: *Transportation Research Part C: Emerging Technologies* 51, pp. 41–65.

Dong, Wei et al. (2013). "Shortest Paths in Stochastic Time-Dependent Networks with Link Travel Time Correlation". In: *Transportation Research Record: Journal of the Transportation Research Board* 2338.1, pp. 58–66.

Evans, Gwynne, J Blackledge, and Peter Yardley (2012). *Numerical methods for partial differential equations*. Springer Science & Business Media.

Flötteröd, Gunnar and Jannis Rohde (2011). "Operational macroscopic modeling of complex urban road intersections". In: *Transportation Research Part B: Methodological* 45.6, pp. 903–922.

Forrest, John (2012). "Cbc (coin-or branch and cut) open-source mixed integer programming solver". In: *URL: https://projects.coin-or.org/Cbc*.

Friesz, Terry L (2010). *Dynamic optimization and differential games*. Vol. 135. Springer Science & Business Media.

Friesz, Terry L, Javier Luque, Roger L Tobin, and Byung-Wook Wie (1989). "Dynamic network traffic assignment considered as a continuous time optimal control problem". In: *Operations Research* 37.6, pp. 893–901.

Friesz, Terry L, David Bernstein, Tony E Smith, Roger L Tobin, and BW Wie (1993). "A variational inequality formulation of the dynamic network user equilibrium problem". In: *Operations Research* 41.1, pp. 179–191.

Friesz, Terry L, Taeil Kim, Changhyun Kwon, and Matthew A Rigdon (2011). "Approximate network loading and dual-time-scale dynamic user equilibrium". In: *Transportation Research Part B: Methodological* 45.1, pp. 176–207.

Gao, Song (2012). "Modeling strategic route choice and real-time information impacts in stochastic and time-dependent networks". In: *Intelligent Transportation Systems, IEEE Transactions on* 13.3, pp. 1298–1311.

Gao, Song and Ismail Chabini (2006). "Optimal routing policy problems in stochastic time-dependent networks". In: *Transportation Research Part B: Methodological* 40.2, pp. 93–122.

Gao, Song and He Huang (2012). "Real-time traveler information for optimal adaptive routing in stochastic time-dependent networks". In: *Transportation Research Part C: Emerging Technologies* 21.1, pp. 196–213.

Gartner, Nathan H and Chronis Stamatiadis (1998). "Integration of dynamic traffic assignment with real-time traffic adaptive control system". In: *Transportation Research Record: Journal of the Transportation Research Board* 1644.1, pp. 150–156.

Gawron, Christian (1999). "Simulation-based traffic assignment: Computing user equilibria in large street networks". PhD thesis.

Gentile, Guido (2016). "Solving a Dynamic User Equilibrium model based on splitting rates with Gradient Projection algorithms". In: *Transportation Research Part B: Methodological* 92, pp. 120–147.

Gentile, Guido, Lorenzo Meschini, and Natale Papola (2005). "Macroscopic arc performance models with capacity constraints for within-day dynamic traffic assignment". In: *Transportation Research Part B: Methodological* 39.4, pp. 319–338.

— (2007). "Spillback congestion in dynamic traffic assignment: a macroscopic flow model with time-varying bottlenecks". In: *Transportation Research Part B: Methodological* 41.10, pp. 1114–1138.

Gentile, Guido et al. (2010). "The General Link Transmission Model for dynamic network loading and a comparison with the DUE algorithm". In: *New developments in transport planning: advances in Dynamic Traffic Assignment* 178, p. 153.

Haddad, Jack, Mohsen Ramezani, and Nikolas Geroliminis (2013). "Cooperative traffic control of a mixed network with two urban regions and a freeway". In: *Transportation Research Part B: Methodological* 54, pp. 17–36.

Hall, Randolph W (1996). "Route choice and advanced traveler information systems on a capacitated and dynamic network". In: *Transportation Research Part C: Emerging Technologies* 4.5, pp. 289–306.

Hall, Randolph William (1983). "Traveler route choice: travel time implications of improved information and adaptive decisions". In: *Transportation Research Part A: General* 17.3, pp. 201–214.

Hamdouch, Younes, Patrice Marcotte, and Sang Nguyen (2004). "A strategic model for dynamic traffic assignment". In: *Networks and Spatial Economics* 4.3, pp. 291–315.

Hamdouch, Younes, WY Szeto, and Y Jiang (2014). "A new schedule-based transit assignment model with travel strategies and supply uncertainties". In: *Transportation Research Part B: Methodological* 67, pp. 35–67.

Han, Ke, Vikash V Gayah, Benedetto Piccoli, Terry L Friesz, and Tao Yao (2014). "On the continuum approximation of the on-and-off signal control on dynamic traffic networks". In: *Transportation Research Part B: Methodological* 61, pp. 73–97.

Han, Ke, Yuqi Sun, Hongcheng Liu, Terry L Friesz, and Tao Yao (2015a). "A bi-level model of dynamic traffic signal control with continuum approximation". In: *Transportation Research Part C: Emerging Technologies* 55, pp. 409–431.

Han, Ke, WY Szeto, and Terry L Friesz (2015b). "Formulation, existence, and computation of boundedly rational dynamic user equilibrium with fixed or endogenous user tolerance". In: *Transportation Research Part B: Methodological* 79, pp. 16–49.

Han, Ke, Benedetto Piccoli, and WY Szeto (2016). "Continuous-time link-based kinematic wave model: formulation, solution existence, and well-posedness". In: *Transportmetrica B: Transport Dynamics* 4.3, pp. 187–222.

Han, Sangjin (2003). "Dynamic traffic modelling and dynamic stochastic user equilibrium assignment for general road networks". In: *Transportation Research Part B: Methodological* 37.3, pp. 225–249.

He, Fang, Di Wu, Yafeng Yin, and Yongpei Guan (2013). "Optimal deployment of public charging stations for plug-in hybrid electric vehicles". In: *Transportation Research Part B: Methodological* 47, pp. 87–101.

Hoang, N. H., H. L. Vu, M. Smith, and D. Ngoduy (2016). "Convex signal control model in a single-destination dynamic traffic assignment". In: *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, pp. 790–794. DOI: 10.1109/ITSC.2016.7795645.

Hoang, N. H., Hai Le Vu, Panda Manoj, and K. Hong Lo (2017a). "A linear framework for dynamic user equilibrium traffic assignment in a single origin-destination capacitated network". In: *Transportation Research Part B: Methodological*. DOI: 10.1016/j.trb.2017.11.013.

Hoang, N. H., Hai Le Vu, Panda Manoj, and Dong Ngoduy (2017b). "An information model for real-time adaptive routing in the dynamic system optimal assignment problem". *Under-review to Networks and Spatial Economics*.

Hoang, N. H. and Hai Le Vu (2017c). "An informed user equilibrium dynamic traffic assignment problem in a single origin-destination stochastic network". *Under-review to Transportation Research Part B: Methodological*.

Huang, Hai-Jun and William HK Lam (2002). "Modeling and solving the dynamic user equilibrium route and departure time choice problem in network with queues". In: *Transportation Research Part B: Methodological* 36.3, pp. 253–273.

Huang, Hai-Jun, Tian-Liang Liu, and Hai Yang (2008). "Modeling the evolutions of day-to-day route choice and year-to-year ATIS adoption with stochastic user equilibrium". In: *Journal of Advanced Transportation* 42.2, pp. 111–127.

Ibarra-Rojas, OJ, F Delgado, R Giesen, and JC Muñoz (2015). "Planning, operation, and control of bus transport systems: A literature review". In: *Transportation Research Part B: Methodological* 77, pp. 38–75.

Jahn, Olaf, Rolf H Möhring, Andreas S Schulz, and Nicolás E Stier-Moses (2005). "System-optimal routing of traffic flows with user constraints in networks with congestion". In: *Operations research* 53.4, pp. 600–616.

Jayakrishnan, R, Hani S Mahmassani, and Ta-Yin Hu (1994). "An evaluation tool for advanced traffic information and management systems in urban networks". In: *Transportation Research Part C: Emerging Technologies* 2.3, pp. 129–147.

Jiang, Yanqun et al. (2011). "A dynamic traffic assignment model for a continuum transportation system". In: *Transportation Research Part B: Methodological* 45.2, pp. 343–363.

Jin, Wen-Long (2015). "Continuous formulations and analytical properties of the link transmission model". In: *Transportation Research Part B: Methodological* 74, pp. 88–103.

Joueiai, Mahtab, Ludovic Leclercq, Hans Van Lint, and Serge P Hoogendoorn (2015). "Multiscale Traffic Flow Model Based on the Mesoscopic Lighthill–Whitham and Richards Models". In: *Transportation Research Record: Journal of the Transportation Research Board* 2491, pp. 98–106.

Kamal, Md Abdus Samad, Jun-ichi Imura, Tomohisa Hayakawa, Akira Ohata, and Kazuyuki Aihara (2015). "Traffic signal control of a road network using MILP in the MPC framework". In: *International journal of intelligent transportation systems research* 13.2, pp. 107–118.

Karmarkar, Narendra (1984). "A new polynomial-time algorithm for linear programming". In: *Proceedings of the sixteenth annual ACM symposium on Theory of computing*. ACM, pp. 302–311.

Kim, Seongmoon, Mark E Lewis, and Chelsea C White III (2005). "Optimal vehicle routing with real-time traffic information". In: *Intelligent Transportation Systems, IEEE Transactions on* 6.2, pp. 178–188.

Küchler, Christian and Stefan Vigerske (2010). "Numerical evaluation of approximation methods in stochastic programming". In: *Optimization* 59.3, pp. 401–415.

Lam, William HK and Hai-Jun Huang (1995). "Dynamic user optimal traffic assignment model for many to one travel demand". In: *Transportation Research Part B: Methodological* 29.4, pp. 243–259.

Lam, William HK, Hu Shao, and Agachai Sumalee (2008). "Modeling impacts of adverse weather conditions on a road network with uncertainties in demand and supply". In: *Transportation research part B: methodological* 42.10, pp. 890–910.

Larsson, Torbjörn and Michael Patriksson (1999). "Side constrained traffic equilibrium models-analysis, computation and applications". In: *Transportation Research Part B: Methodological* 33.4, pp. 233–264.

Le, Tung, Hai L. Vu, Yoni Nazarathy, Quoc Bao Vo, and Serge Hoogendoorn (2013). "Linear-quadratic model predictive control for urban traffic networks". In: *Transportation Research Part C: Emerging Technologies* 36.Supplement C, pp. 498–512. ISSN: 0968–090X. DOI: https://doi.org/10.1016/j.trc.2013.06.021. URL: http://www.sciencedirect.com/science/article/pii/S0968090X1300154X.

Le, Tung et al. (2017). "Utility optimization framework for a distributed traffic control of urban road networks". In: *Transportation Research Part B: Methodological* 105, pp. 539–558.

Li, Pengfei, Pitu Mirchandani, and Xuesong Zhou (2015). "Solving simultaneous route guidance and traffic signal optimization problem using space-phase-time hypernetwork". In: *Transportation Research Part B: Methodological* 81, pp. 103–130.

Li, Yue, S Travis Waller, and Thanasis Ziliaskopoulos (2003). "A decomposition scheme for system optimal dynamic traffic assignment models". In: *Networks and Spatial Economics* 3.4, pp. 441–455.

Lighthill, Michael J and Gerald Beresford Whitham (1955). "On kinematic waves. II. A theory of traffic flow on long crowded roads". In: *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*. Vol. 229. 1178. The Royal Society, pp. 317–345.

Lin, Wei-Hua and Hong K Lo (2000). "Are the objective and solutions of dynamic user-equilibrium models always consistent?" In: *Transportation Research Part A: Policy and Practice* 34.2, pp. 137–144.

Lin, Wei-Hua and Chenghong Wang (2004). "An enhanced 0-1 mixed-integer LP formulation for traffic signal control". In: *IEEE Transactions on Intelligent transportation systems* 5.4, pp. 238–245.

Lindsey, Robin, Terry Daniel, Eyran Gisches, and Amnon Rapoport (2014). "Pre-trip information and route-choice decisions with stochastic travel conditions: Theory". In: *Transportation Research Part B: Methodological* 67, pp. 187–207.

Liu, Tian-Liang, Hai-Jun Huang, Hai Yang, and Xiaolei Guo (2009). "Equilibria and inefficiency in traffic networks with stochastic capacity and information provision". In: *Transportation and Traffic Theory 2009: Golden Jubilee*. Springer, pp. 263–281.

Lo, Hong, Bin Ran, and Bruce Hongola (1996). "Multiclass dynamic traffic assignment model: formulation and computational experiences". In: *Transportation Research Record: Journal of the Transportation Research Board* 1537, pp. 74–82.

Lo, Hong K (1999). "A novel traffic signal control formulation". In: *Transportation Research Part A: Policy and Practice* 33.6, pp. 433–448.

Lo, Hong K and WY Szeto (2002). "A cell-based variational inequality formulation of the dynamic user optimal assignment problem". In: *Transportation Research Part B: Methodological* 36.5, pp. 421–443.

Lo, Hong K and Wai Yuen Szeto (2004). "Modeling advanced traveler information services: static versus dynamic paradigms". In: *Transportation Research Part B: Methodological* 38.6, pp. 495–515.

Ma, Rui, Xuegang Jeff Ban, and Jong-Shi Pang (2014). "Continuous-time dynamic system optimum for single-destination traffic networks with queue spillbacks". In: *Transportation Research Part B: Methodological* 68, pp. 98–122.

Ma, Rui, Xuegang Jeff Ban, and WY Szeto (2017). "Emission modeling and pricing on single-destination dynamic traffic networks". In: *Transportation Research Part B: Methodological* 100, pp. 255–283.

Mackie, Peter, Tom Worsley, and Jonas Eliasson (2014). "Transport appraisal revisited". In: *Research in Transportation Economics* 47, pp. 3–18.

Mahmassani, Hani S (2001). "Dynamic network traffic assignment and simulation methodology for advanced system management applications". In: *Networks and spatial economics* 1.3-4, pp. 267–292.

Mahmassani, Hani S et al. (2012). *Implementation and evaluation of weather responsive traffic estimation and prediction system*. Tech. rep.

Mahut, Michael, Michael Florian, and Nicolas Tremblay (2003). "Space-time queues and dynamic traffic assignment: A model, algorithm and applications". In: *Transportation Research Board 82nd Annual Meeting*.

Marcotte, Patrice and Sang Nguyen (1998). "Hyperpath formulations of traffic assignment problems". In: *Equilibrium and advanced transportation modelling*. Springer, pp. 175–200.

Merchant, Deepak K and George L Nemhauser (1978). "A model and an algorithm for the dynamic traffic assignment problems". In: *Transportation science* 12.3, pp. 183–199.

Mohan, Ranju and Gitakrishnan Ramadurai (2013). "State-of-the art of macroscopic traffic flow modelling". In: *International Journal of Advances in Engineering Sciences and Applied Mathematics* 5.2-3, pp. 158–176.

Nagel, Kai and Michael Schreckenberg (1992). "A cellular automaton model for freeway traffic". In: *Journal de physique I* 2.12, pp. 2221–2229.

Newell, Gordon F (1993). "A simplified theory of kinematic waves in highway traffic (part I, II, III)". In: *Transportation Research Part B: Methodological* 27.4, pp. 281–313.

Ng, ManWo and S Travis Waller (2010). "Reliable evacuation planning via demand inflation and supply deflation". In: *Transportation Research Part E: Logistics and Transportation Review* 46.6, pp. 1086–1094.

Ngoduy, D (2010). "Multiclass first-order modelling of traffic networks using discontinuous flow-density relationships". In: *Transportmetrica* 6.2, pp. 121–141.

Ngoduy, D, S Hoogendoorn, and J van Lint (2005). "Modeling traffic flow operation in multilane and multiclass urban networks". In: *Transportation Research Record: Journal of the Transportation Research Board* 1923, pp. 73–81.

Ngoduy, Dong, N. H. Hoang, Hai Le Vu, and D Watling (2016). "Optimal queue placement in dynamic system optimum solutions for single origin-destination traffic networks". In: *Transportation Research Part B: Methodological*. DOI: 10.1016/j.trb.2015.11.011.

Nguyen, Sang and Clermont Dupuis (1984). "An efficient method for computing traffic equilibria in networks with asymmetric transportation costs". In: *Transportation Science* 18.2, pp. 185–202.

Nie, Xiaojian and H Michael Zhang (2005). "A comparative study of some macroscopic link models used in dynamic traffic assignment". In: *Networks and Spatial Economics* 5.1, pp. 89–115.

Nuzzolo, Agostino, Umberto Crisalli, Antonio Comi, and Luca Rosati (2016). "A mesoscopic transit assignment model including real-time predictive information on crowding". In: *Journal of Intelligent Transportation Systems* 20.4, pp. 316–333.

Osorio, Carolina, Gunnar Flötteröd, and Michel Bierlaire (2011). "Dynamic network loading: a stochastic differentiable model that derives link state distributions". In: *Transportation Research Part B: Methodological* 45.9, pp. 1410–1423.

Osorio, Carolina and Gunnar Flötteröd (2014). "Capturing dependency among link boundaries in a stochastic dynamic network loading model". In: *Transportation Science* 49.2, pp. 420–431.

Paz, Alexander and Srinivas Peeta (2009a). "Behavior-consistent real-time traffic routing under information provision". In: *Transportation Research Part C: Emerging Technologies* 17.6, pp. 642–661.

— (2009b). "Information-based network control strategies consistent with estimated driver behavior". In: *Transportation Research Part B: Methodological* 43.1, pp. 73–96.

Peeta, Srinivas and Athanasios K Ziliaskopoulos (2001). "Foundations of dynamic traffic assignment: The past, the present and the future". In: *Networks and Spatial Economics* 1.3-4, pp. 233–265.

Peeta, Srinivas and Chao Zhou (2002). "A hybrid deployable dynamic traffic assignment framework for robust online route guidance". In: *Networks and Spatial Economics* 2.3, pp. 269–294.

Perakis, Georgia and Guillaume Roels (2006). "An analytical model for traffic delays and the dynamic user equilibrium problem". In: *Operations Research* 54.6, pp. 1151–1171.

Polychronopoulos, George H and John N Tsitsiklis (1996). "Stochastic shortest path problems with recourse". In: *Networks* 27.2, pp. 133–143.

Punzo, Vincenzo and Fulvio Simonelli (2005). "Analysis and comparison of microscopic traffic flow models with real traffic microscopic data". In: *Transportation Research Record: Journal of the Transportation Research Board* 1934, pp. 53–63.

Rahman, Mizanur, Mashrur Chowdhury, Yuanchang Xie, and Yiming He (2013). "Review of microscopic lane-changing models and future research opportunities". In: *IEEE transactions on intelligent transportation systems* 14.4, pp. 1942–1956.

Ran, Bin, Peter J Jin, David Boyce, Tony Z Qiu, and Yang Cheng (2012). "Perspectives on future transportation research: Impact of intelligent transportation system technologies on next-generation transportation modeling". In: *Journal of Intelligent Transportation Systems* 16.4, pp. 226–242.

Richards, Paul I (1956). "Shock waves on the highway". In: *Operations research* 4.1, pp. 42–51.

Rinaldi, Marco and Chris MJ Tampère (2015). "An extended coordinate descent method for distributed anticipatory network traffic control". In: *Transportation Research Part B: Methodological* 80, pp. 107–131.

Rudloff, Christian et al. (2015). "Influence of Weather on Transport Demand: Case Study from the Vienna, Austria, Region". In: *Transportation Research Record: Journal of the Transportation Research Board* 2482, pp. 110–116.

Shaked, Moshe and George Shanthikumar (2007). *Stochastic orders*. Springer Science & Business Media.

Shen, Wei, Yu Nie, and H Zhang (2007). "Dynamic network simplex method for designing emergency evacuation plans". In: *Transportation Research Record: Journal of the Transportation Research Board* 2022, pp. 83–93.

Shen, Wei and H Zhang (2008). "What do different traffic flow models mean for system-optimal dynamic traffic assignment in a many-to-one network?" In: *Transportation Research Record: Journal of the Transportation Research Board* 2088, pp. 157–166.

Shen, Wei and HM Zhang (2014). "System optimal dynamic traffic assignment: Properties and solution procedures in the case of a many-to-one network". In: *Transportation Research Part B: Methodological* 65, pp. 1–17.

Sims, Arthur G and Kenneth W Dobinson (1980). "The Sydney coordinated adaptive traffic (SCAT) system philosophy and benefits". In: *Vehicular Technology, IEEE Transactions on* 29.2, pp. 130–137.

Smits, Erik-Sander, Michiel CJ Bliemer, Adam J Pel, and Bart van Arem (2015). "A family of macroscopic node models". In: *Transportation Research Part B: Methodological* 74, pp. 20–39.

Spiess, Heinz and Michael Florian (1989). "Optimal strategies: a new assignment model for transit networks". In: *Transportation Research Part B: Methodological* 23.2, pp. 83–102.

Sun, Chao, Lin Cheng, Senlai Zhu, Fei Han, and Zhaoming Chu (2017). "Multi-criteria user equilibrium model considering travel time, travel time reliability and distance". In: *Transportation Research Part D: Transport and Environment*.

Sun, Dazhi, Rahim F Benekohal, and S Travis Waller (2006). "Bi-level Programming Formulation and Heuristic Solution Approach for Dynamic Traffic Signal Optimization". In: *Computer-Aided Civil and Infrastructure Engineering* 21.5, pp. 321–333.

Szeto, WY, Yu Jiang, and Agachai Sumalee (2011a). "A Cell-Based Model for Multi-class Doubly Stochastic Dynamic Traffic Assignment". In: *Computer-Aided Civil and Infrastructure Engineering* 26.8, pp. 595–611.

Szeto, WY, Muthu Solayappan, and Yu Jiang (2011b). "Reliability-Based Transit Assignment for Congested Stochastic Transit Networks". In: *Computer-Aided Civil and Infrastructure Engineering* 26.4, pp. 311–326.

Szeto, WY and SC Wong (2012). "Dynamic traffic assignment: model classifications and recent advances in travel choice principles". In: *Central European Journal of Engineering* 2.1, pp. 1–18.

Tampère, Chris MJ, Ruben Corthout, Dirk Cattrysse, and Lambertus H Immers (2011). "A generic class of first order node models for dynamic macroscopic simulation of traffic flows". In: *Transportation Research Part B: Methodological* 45.1, pp. 289–309.

Taylor, Nicholas B (2003). "The CONTRAM dynamic traffic assignment model". In: *Networks and Spatial Economics* 3.3, pp. 297–322.

Timotheou, Stelios, Christos G Panayiotou, and Marios M Polycarpou (2015). "Distributed Traffic Signal Control Using the Cell Transmission Model via the Alternating Direction Method of Multipliers". In: *Intelligent Transportation Systems, IEEE Transactions on* 16.2, pp. 919–933.

Toledo, Tomer, Oded Cats, Wilco Burghout, and Haris N Koutsopoulos (2010). "Mesoscopic simulation for transit operations". In: *Transportation Research Part C: Emerging Technologies* 18.6, pp. 896–908.

Ukkusuri, Satish, Kien Doan, and HM Aziz (2013). "A bi-level formulation for the combined dynamic equilibrium based traffic signal control". In: *Procedia-Social and Behavioral Sciences* 80, pp. 729–752.

Ukkusuri, Satish V and Gopal R Patil (2007). "Exploring user behavior in online network equilibrium problems". In: *Transportation Research Record: Journal of the Transportation Research Board* 2029.1, pp. 31–38.

Ukkusuri, Satish V and S Travis Waller (2008). "Linear programming models for the user and system optimal dynamic network design problem: formulations, comparisons and extensions". In: *Networks and Spatial Economics* 8.4, pp. 383–406.

Ukkusuri, Satish V, Gitakrishnan Ramadurai, and Gopal Patil (2010). "A robust transportation signal control problem accounting for traffic dynamics". In: *Computers & Operations Research* 37.5, pp. 869–879.

Ukkusuri, Satish V, Lanshan Han, and Kien Doan (2012). "Dynamic user equilibrium with a path based cell transmission model for general traffic networks". In: *Transportation Research Part B: Methodological* 46.10, pp. 1657–1684.

Unnikrishnan, Avinash and Steven Travis Waller (2009). "User equilibrium with recourse". In: *Networks and Spatial Economics* 9.4, pp. 575–593.

Van Vliet, Dirck (1976). "Road assignment-II: The GLTS model". In: *Transportation Research* 10.3, pp. 145–149.

Vickrey, William S (1969). "Congestion theory and transport investment". In: *The American Economic Review*, pp. 251–260.

Wallace, Charles E, Kenneth G Courage, and Mohammed A Hadi (1991). *TRANSYT-7F Users Manual*. Federal Highway Administration.

Waller, S Travis, David Fajardo, Melissa Duell, and Vinayak Dixit (2013). "Linear Programming Formulation for Strategic Dynamic Traffic Assignment". In: *Networks and Spatial Economics* 13.4, pp. 427–443.

Wang, Chen, Bertrand David, René Chalon, and Chuantao Yin (2015). "Dynamic road lane management study: A Smart City application". In: *Transportation Research Part E: Logistics and Transportation Review*.

Wang, David ZW and Bo Du (2016). "Continuum modelling of spatial and dynamic equilibrium in a travel corridor with heterogeneous commuters—A partial differential complementarity system approach". In: *Transportation Research Part B: Methodological* 85, pp. 1–18.

Wardrop, John Glen (1952). "ROAD PAPER. SOME THEORETICAL ASPECTS OF ROAD TRAFFIC RESEARCH." In: *ICE Proceedings: Engineering Divisions*. Vol. 1. 3. Thomas Telford, pp. 325–362.

Watling, David (2006). "User equilibrium traffic network assignment with stochastic travel times and late arrival penalty". In: *European journal of operational research* 175.3, pp. 1539–1556.

Wie, Byung-Wook, Terry L Friesz, and Roger L Tobin (1990). "Dynamic user optimal traffic assignment on congested multidestination networks". In: *Transportation Research Part B: Methodological* 24.6, pp. 431–442.

Wie, Byung-Wook and Roger L Tobin (1998). "Dynamic congestion pricing models for general traffic networks". In: *Transportation Research Part B: Methodological* 32.5, pp. 313–327.

Wie, Byung-Wook, Roger L Tobin, and Malachy Carey (2002). "The existence, uniqueness and computation of an arc-based dynamic network user equilibrium formulation". In: *Transportation Research Part B: Methodological* 36.10, pp. 897–918.

Willumsen, Luis G et al. (2011). *Modelling transport*. Wiley.

Wright, Stephen E and James J Rohal (2014). "Solving the continuous nonlinear resource allocation problem with an interior point method". In: *Operations Research Letters* 42.6, pp. 404–408.

Xiao, Lin and Hong K Lo (2014). "Adaptive vehicle navigation with en route stochastic traffic information". In: *IEEE Transactions on Intelligent Transportation Systems* 15.5, pp. 1900–1912.

Xie, Chi and Zugang Liu (2014). "On the stochastic network equilibrium with heterogeneous choice inertia". In: *Transportation Research Part B: Methodological*.

Yang, Baiyu and Elise Miller-Hooks (2004a). "Adaptive routing considering delays due to signal operations". In: *Transportation Research Part B: Methodological* 38.5, pp. 385–413.

Yang, Hai (1998). "Multiple equilibrium behaviors and advanced traveler information systems with endogenous market penetration". In: *Transportation Research Part B: Methodological* 32.3, pp. 205–218.

Yang, Hai and Hai-Jun Huang (2004b). "The multi-class, multi-criteria traffic network equilibrium and systems optimum problem". In: *Transportation Research Part B: Methodological* 38.1, pp. 1–15.

Yperman, Isaak (2007). "The link transmission model for dynamic network loading (Doctoral disseration)". In: *KU Leuven, Belgium*.

Yperman, Isaak, Steven Logghe, Chris MJ Tampere, and Ben Immers (2006). "Multicommodity link transmission model for dynamic network loading". In: *Transportation Research Board 85th Annual Meeting*. 06-1062.

Zhang, HM, Yu Nie, and Zhen Qian (2013). "Modelling network flow with and without link interactions: the cases of point queue, spatial queue and cell transmission model". In: *Transportmetrica B: Transport Dynamics* 1.1, pp. 33–51.

Zhong, RX, A Sumalee, TL Friesz, and William HK Lam (2011). "Dynamic user equilibrium with side constraints for a traffic network: theoretical development and numerical solution algorithm". In: *Transportation Research Part B: Methodological* 45.7, pp. 1035–1061.

Ziliaskopoulos, Athanasios K (2000). "A linear programming model for the single destination system optimum dynamic traffic assignment problem". In: *Transportation science* 34.1, pp. 37–49.

Zito, Pietro, Gianfranco Amato, Salvatore Amoroso, and Maria Berrittella (2011). "The effect of Advanced Traveller Information Systems on public transport demand and its uncertainty". In: *Transportmetrica* 7.1, pp. 31–43.

February 15, 2018