

Organizers



Department of Civil Engineering &  
Institute of Transport Studies,  
The University of Hong Kong



The Hong Kong University of  
Science and Technology



Hong Kong Society for  
Transportation Studies

Proceedings of  
**The 7th International Symposium on  
Dynamic Traffic Assignment**

6-8 June 2018  
Hong Kong

# SMART TRANSPORTATION



Edited by: W.Y. Szeto, Hong K. Lo and S.C. Wong

# Message from the Co-Chairs of the Conference Organizing Committee



**W.Y. Szeto**  
Co-chair



**Hong K. Lo**  
Co-chair



**S.C. Wong**  
Co-chair

It is our distinct pleasure to welcome you to the 7th International Symposium on Dynamic Traffic Assignment, co-organized by the Department of Civil Engineering and Institute of Transport Studies of The University of Hong Kong, the Department of Civil and Environmental Engineering of The Hong Kong University of Science and Technology, and the Hong Kong Society for Transportation Studies.

This Symposium aims to foster excellence in dynamic traffic assignment (DTA) research and practice, and provide a forum for exchanging innovative ideas and challenges on DTA and related transportation science research problems. The theme of this Symposium is “smart transportation”. Nowadays, in order to provide smooth traffic flow, high degree of safety and sustainable urban environment, smart transportation systems become popular. The systems include the integrated application of advanced management strategies and emerging technology. Such evolution can enhance the current transportation systems and come out positive in benefit-cost analyses. The systems are considered to be essential especially in addressing the challenges of complex matrix within smart cities. The DTA theory together with the latest technology, such as autonomous and electric vehicles, APPs, and GPS, are important ingredients towards the development of a smart city. Hong Kong is one of the cities moving towards this direction. This Symposium attracts international scholars to share their latest finding on DTA with local practitioners and academics to encourage the development of Hong Kong to be a leading smart city with more competitive advantages to adjacent cities, and eventually to make a positive impact on Hong Kong society.

This year, we are honored that Professor Terry Friesz, Harold and Inge Marcus Chaired Professor of Industrial Engineering of the Pennsylvania State University, has kindly accepted our invitation to be the keynote speaker. We would like to thank the Croucher Foundation for their financial sponsorship. We would also like to thank the Organizing Committee members, the helpers, and especially our Symposium Secretary, Miss Ruby Kwok, for their assistance in putting together this Symposium.

Finally, on behalf of the Conference Organizing Committee, we hope that you will find this Symposium a most worthwhile and enjoyable time.

Handwritten signature of W.Y. Szeto in black ink.

W.Y. Szeto

Handwritten signature of Hong K. Lo in black ink.

Hong K. Lo

Handwritten signature of S.C. Wong in black ink.

S.C. Wong

Co-Chairs  
Conference Organizing Committee  
The 7th International Symposium on Dynamic Traffic Assignment

## Co-Chairs

- Dr. W.Y. Szeto, The University of Hong Kong
- Professor Hong K. Lo, The Hong Kong University of Science and Technology
- Professor S.C. Wong, The University of Hong Kong

## Members

- Professor Anthony Chen, The Hong Kong Polytechnic University
- Dr. Andy Chow, City University of Hong Kong
- Dr. Sylvia Y. He, The Chinese University of Hong Kong
- Dr. Yong-Hong Kuo, The University of Hong Kong
- Professor William H.K. Lam, The Hong Kong Polytechnic University
- Professor Janny M.Y. Leung, The Chinese University of Hong Kong
- Professor Siu Ming Lo, City University of Hong Kong
- Professor Becky P.Y. Loo, The University of Hong Kong
- Professor Agachai Sumalee, The Hong Kong Polytechnic University
- Professor Donggen Wang, Hong Kong Baptist University
- Professor Alan W.G. Wong, Technological and Higher Education Institute of Hong Kong
- Professor Hai Yang, The Hong Kong University of Science and Technology

# International Scientific Advisory Committee

- Michiel Bliemer, The University of Sydney, Australia
- Giulio Cantarella, University of Salerno, Italy
- Terry Friesz, Pennsylvania State University, USA
- Benjamin Heydecker, University College London, UK
- Mark Hickman, University of Queensland, Australia
- Masao Kuwahara, Tohoku University, Japan
- Hong K. Lo, The Hong Kong University of Science and Technology, Hong Kong
- Srinivas Peeta, Purdue University, USA
- Mike Smith, University of York, UK
- Iryo Takamasa, Kobe University, Japan
- Chris Tampere, Katholieke Universiteit Leuven, Belgium
- Satish Ukkusuri, Purdue University, USA
- David Watling, University of Leeds, UK

We gratefully acknowledge the support of Croucher Foundation.



Croucher Foundation  
裘槎基金會

# Schedule Overview

| Date/Time   | 6 Jun (Wed)                                                                      | 7 Jun (Thur)                                               |                                       | 8 Jun (Fri)                                                             |
|-------------|----------------------------------------------------------------------------------|------------------------------------------------------------|---------------------------------------|-------------------------------------------------------------------------|
| 08:15-09:00 | <b>Registration</b>                                                              |                                                            |                                       |                                                                         |
| Venue       | WGW Theatre                                                                      | WGW Theatre                                                | Rm P5-03                              | WGW Theatre                                                             |
| 09:00-10:40 | <b>A1</b><br><b>Opening,<br/>Keynote<br/>Presentation &amp;<br/>Photo Taking</b> | <b>B1</b><br>Traffic and<br>Demand<br>Management I         | <b>B2</b><br>Day-To-Day<br>Dynamics I | <b>C1</b><br>Within-Day<br>Equilibrium II                               |
| 10:40-11:00 | <b>Coffee Break</b>                                                              |                                                            |                                       |                                                                         |
| 11:00-12:40 | <b>A2</b><br>Within-Day<br>Equilibrium I                                         | <b>B3</b><br>Behavior Modeling                             | <b>B4</b><br>Sharing Economy          | <b>C2</b><br>Dynamic Traffic<br>Control II                              |
| 12:40-14:00 | <b>Lunch Break</b>                                                               |                                                            |                                       |                                                                         |
| 14:00-15:40 | <b>A3</b><br>Network Loading<br>and Simulation I                                 | <b>B5</b><br>Travel Demand<br>Management and<br>Estimation | <b>B6</b><br>Public Transport I       | <b>C3</b><br>Day-To-Day<br>Dynamics II                                  |
| 15:40-16:00 | <b>Coffee Break</b>                                                              |                                                            |                                       |                                                                         |
| 16:00-17:40 | <b>A4</b><br>Dynamic Traffic<br>Control I                                        | <b>B7</b><br>Network Loading<br>and Simulation II          | <b>B8</b><br>Public Transport II      | <b>C4</b><br>Traffic and<br>Demand<br>Management II &<br><b>Closing</b> |
| 18:00-21:30 | <b>Conference Dinner</b>                                                         |                                                            |                                       |                                                                         |

# Detailed Presentation Schedule

## Day 1: 6 June 2018 (Wednesday), Wang Gungwu Theatre, Graduate House

| Time                                                                                | Event                                                                                                                                                                                                                                       |
|-------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Session A1: Opening &amp; Keynote Presentation</b><br><i>Chair: W.Y. Szeto</i>   |                                                                                                                                                                                                                                             |
| 09:00-10:40                                                                         | <b>Welcome Speech</b><br>W.Y. Szeto                                                                                                                                                                                                         |
|                                                                                     | <b>The Near-term Future of Dynamic Traffic Assignment</b><br>Terry L. Friesz                                                                                                                                                                |
| <b>Session A2 Within-Day Equilibrium</b><br><i>Chair: Xuegang Ban</i>               |                                                                                                                                                                                                                                             |
| 11:00-12:40                                                                         | <b>Dynamic User Equilibrium with Recourse and Variable Message Signs</b><br>Kenneth A. Perrine and Michael W. Levin                                                                                                                         |
|                                                                                     | <b>Modeling Link-Based Dynamic User Equilibria with Differential Complementarity Systems</b><br>Xuegang Ban                                                                                                                                 |
|                                                                                     | <b>A Scenario-Based Approach for the Reliability-Based User Equilibrium Problem in Dynamic Stochastic Networks Considering Travel Time Correlations and Heterogeneous Users</b><br>Ali Zockaie, Hani S. Mahmassani and Fatemeh Fakhrmoosavi |
|                                                                                     | <b>An Inner Approximation Algorithm for the Time Varying Network User Equilibrium Problem</b><br>Athanasios Ziliaskopoulos                                                                                                                  |
| <b>Session A3 Network Loading and Simulation</b><br><i>Chair: Stephen D. Boyles</i> |                                                                                                                                                                                                                                             |
| 14:00-15:40                                                                         | <b>A Policy Based Chronological Link Transmission Model for Stochastic Time Dependent Traffic Networks</b><br>Hemant Gehlot and Satish Ukkusuri                                                                                             |
|                                                                                     | <b>Handling Non-Unique Flows in Macroscopic First-Order Intersection Models – An Equilibrium Theory</b><br>Yen-Hsiang Chen and Chris M.J. Tampère                                                                                           |
|                                                                                     | <b>Riemann Solvers at a Junction for Traffic Flow on a Road Network by Using a Conserved Higher-Order Model</b><br>D.L. Qiao, P. Zhang, K. Zhang and H. Liu                                                                                 |
|                                                                                     | <b>Understanding the Trade Offs between DTA Models Realism and Robustness: The Impact of Spillback Modeling</b><br>Stephen D. Boyles and Natalia Ruiz Juri                                                                                  |
| <b>Session A4 Dynamic Traffic Control I</b><br><i>Chair: Feng Zhu</i>               |                                                                                                                                                                                                                                             |
| 16:00-17:40                                                                         | <b>Controlling Pedestrian Flows Using a Dynamic Traffic Management System</b><br>Nicholas Molyneaux, Riccardo Scarinci and Michel Bierlaire                                                                                                 |
|                                                                                     | <b>Architectures of Urban Traffic Control with Consideration of Drivers' Route Choices</b><br>Andy H.F. Chow                                                                                                                                |
|                                                                                     | <b>Dynamic Traffic Signal Optimization Considering Network Equilibrium Flows</b><br>Wei Huang and Hong K. Lo                                                                                                                                |
|                                                                                     | <b>Safety, Stability and Smoothness in Control of Connected Autonomous Vehicles</b><br>Benjamin Heydecker and Dirk Yihua Chen                                                                                                               |

# Detailed Presentation Schedule

## Day 2: 7 June 2018 (Thursday), Wang Gungwu Theatre, Graduate House

| Time                                                                                       | Event                                                                                                                                                                                                                                     |
|--------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Session B1 Traffic and Demand Management I</b><br><i>Chair: Takamasa Iryo</i>           |                                                                                                                                                                                                                                           |
| 09:00-10:40                                                                                | <b>Traffic Network Partitioning for Hierarchical Macroscopic Fundamental Diagram Applications Based on Fusion of GPS Probe and Loop Detector Data</b><br>Kang An, Xianbiao Hu and Xiaohong Chen                                           |
|                                                                                            | <b>A Cell-Based Dynamic Congestion Pricing Considering Travel Distance and Congestion Level</b><br>Qixiu Cheng and Zhiyuan Liu                                                                                                            |
|                                                                                            | <b>An Online Interactive Experiment to Evaluate Tradable Mobility Credit Strategies</b><br>Ye Tian and Yi-Chang Chiu                                                                                                                      |
|                                                                                            | <b>An Optimal Control Framework for Multi-Region Macroscopic Fundamental Diagram Systems Considering Route Choice and Departure Time Choice</b><br>Renxin Zhong, Yunping Huang, Jianhui Xiong, Nan Zheng, William Lam and Agachai Sumalee |
| <b>Session B3 Behavior Modeling</b><br><i>Chair: Andy H.F. Chow</i>                        |                                                                                                                                                                                                                                           |
| 11:00-12:40                                                                                | <b>Modeling the Users Departure Time Choice Adaption for a Tolling System</b><br>Benedikt Bracher and Klaus Bogenberger                                                                                                                   |
|                                                                                            | <b>Empirical Analysis of Detours on Freeways in UK, France and Germany</b><br>Markus Auer, Hubert Rehborn and Klaus Bogenberger                                                                                                           |
|                                                                                            | <b>Assess the Effectiveness of En-Route Information Provision in No-Notice Evacuation Using Agent-Based Simulation</b><br>Yi Wang, Miltos Kyriakidis and Vinh N. Dang                                                                     |
|                                                                                            | <b>User's Behavior and Dynamic Network Loading</b><br>Sergio F.A. Batista and Ludovic Leclercq                                                                                                                                            |
| <b>Session B5 Travel Demand Management and Estimation</b><br><i>Chair: Carolina Osorio</i> |                                                                                                                                                                                                                                           |
| 14:00-15:40                                                                                | <b>A Fuzzy Set-Based Approach to Estimate the Dynamic Demands in Urban Road Networks with Imprecise Data</b><br>Bei Wang and Yu Jiang                                                                                                     |
|                                                                                            | <b>An Efficient Algorithm for High-Dimensional Dynamic Origin-Destination Demand Estimation of Inefficient Large-Scale Stochastic Simulation-Based Traffic Models</b><br>Carolina Osorio                                                  |
|                                                                                            | <b>Hourly Origin-Destination Demand Estimation Using Time Coefficients from Traffic Counts Based on Semi-Dynamic Traffic Assignment</b><br>Motohiro Fujita and Shinji Yamada                                                              |
|                                                                                            | <b>A Linear Programming Formulation for Incentive-Based Travel Demand Management</b><br>Shantanu Chakraborty, David Rey and S. Travis Waller                                                                                              |
| <b>Session B7 Network Loading and Simulation II</b><br><i>Chair: Ludovic Leclercq</i>      |                                                                                                                                                                                                                                           |
| 16:00-17:40                                                                                | <b>A Probabilistic Analytical Traffic-Theoretic Network Loading Model for Large-Scale Network Optimization</b><br>Jing Lu and Carolina Osorio                                                                                             |
|                                                                                            | <b>Interfacing a Dynamic Discrete Choice Based Travel Demand Model to an Agent Based Traffic Simulation Model</b><br>Mohammad Saleem, Oskar B. Västberg and Anders Karlström                                                              |
|                                                                                            | <b>A Probabilistic Sensitivity Analysis Guided Cross Entropy Method for Efficient Calibration of Traffic Models</b><br>Xiuxia Xie, Siwei Hu, Kaiyan Fu and Renxin Zhong                                                                   |
|                                                                                            | <b>Efficient Calibration of an Activity-Based Multimodal Transport Microsimulation Model</b><br>Ashraf Uz Zaman Patwary, Wei Huang and Hong K. Lo                                                                                         |

## Day 2: 7 June 2018 (Thursday), Rm P5-03, Graduate House

| Time                                                               | Event                                                                                                                                                                                 |
|--------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Session B2 Day-To-Day Dynamics I</b><br><i>Chair: Feng Xiao</i> |                                                                                                                                                                                       |
| 09:00-10:40                                                        | <b>Day-To-Day Dynamics under Advanced Traffic Information</b><br>Hongbo Ye, Feng Xiao and Hai Yang                                                                                    |
|                                                                    | <b>Day-To-Day Multimodal Dynamic Traffic Assignment: Impacts of the Learning Process in Case of Non-Unique Solutions</b><br>Mostafa Ameli, Jean Patrick Lebacque and Ludovic Leclercq |
|                                                                    | <b>Effects of Traffic Information on Drivers' Day-To-Day Route Choices</b><br>Genaro Peque Jr., Toshihiko Miyagi and Fumitaka Kurauchi                                                |
|                                                                    | <b>Effect of Information Collection Behaviour on Stability of a Simplified Departure Time Choice Problem</b><br>Takamasa Iryo                                                         |
| <b>Session B4 Sharing Economy</b><br><i>Chair: Jiancheng Long</i>  |                                                                                                                                                                                       |
| 11:00-12:40                                                        | <b>A Dynamic Green Bike Repositioning Problem</b><br>Chin Sum Shui and W.Y. Szeto                                                                                                     |
|                                                                    | <b>The Fleet Deployment Problem for Leisure-Oriented Electric Vehicle Sharing Systems with Uncertain Demand</b><br>Chung-Cheng Lu, Hui-Ju Chen and Hsiao-Tung Wang                    |
|                                                                    | <b>Modeling and Managing Morning Commute with Park and Ride-Sharing</b><br>Zhihui Huang, Jiancheng Long and Haoxiang Liu                                                              |
|                                                                    | <b>On the Morning Commute Problem in a Y-Shaped Network with Individual and Household Travelers</b><br>Dongdong He, Yang Liu, Qiuyan Zhong and David Z.W. Wang                        |
| <b>Session B6 Public Transport I</b><br><i>Chair: Mark Hickman</i> |                                                                                                                                                                                       |
| 14:00-15:40                                                        | <b>Passenger-Oriented Train Schedule Generation Model for a Railway System</b><br>Jiemin Xie, Shuguang Zhan, S.C. Wong and S.M. Lo                                                    |
|                                                                    | <b>Mixed Bus Fleet Routing and Scheduling Considering Driving Range and Locations of Charging Stations</b><br>Lu Li, Hong K. Lo and Feng Xiao                                         |
|                                                                    | <b>Real-Time Integrated Re-Scheduling for Tramway Operations</b><br>Kam-Fung Cheung, Yong-Hong Kuo, David S.W. Lai and Janny M.Y. Leung                                               |
|                                                                    | <b>Train Rescheduling Strategies in a Complete Blockage</b><br>Shuguang Zhan, S.C. Wong, Qiyuan Peng and S.M. Lo                                                                      |
| <b>Session B8 Public Transport II</b><br><i>Chair: Yu Jiang</i>    |                                                                                                                                                                                       |
| 16:00-17:40                                                        | <b>Simulating the Behaviors of Taxi-Cruising and Taxi-Dispatching Markets in a Cell-Based Network</b><br>R.C.P. Wong, W.H. Yang, W.Y. Szeto, and R.T.H. Wong                          |
|                                                                    | <b>Scheduling Synchronization with Time-Dependent Data</b><br>Yu Jiang, Kelvin Lee, Avishai Ceder and Otto Nielsen                                                                    |
|                                                                    | <b>Commuter's Departure Time Choice with Fare-Reward Scheme in a Multi-To-Single Mass Transit System</b><br>Yili Tang, Hai Yang and Chenzhe Liu                                       |
|                                                                    | <b>Bus Schedule Optimization for Multimodal Emergency Evacuation</b><br>Xia Yang, Xiaozheng (Sean) He and Xuegang (Jeff) Ban                                                          |

# Detailed Presentation Schedule

## Day 3: 8 June 2018 (Friday), Wang Gungwu Theatre, Graduate House

| Time                                                                                               | Event                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
|----------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Session C1 Within-Day Equilibrium II</b><br><i>Chair: Yang Liu</i>                              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
| 09:00-10:40                                                                                        | <p><b>Dynamic Traffic Assignment in a Corridor Network Optimum vs. Equilibrium</b><br/>Haoran Fu, Takashi Akamatsu and Kentaro Wada</p> <p><b>A Link-Based Probit Model for Dynamic Traffic Assignment</b><br/>Genaro Peque Jr., Junji Urata and Takamasa Iryo</p> <p><b>A Dynamic System Optimal Model for Hybrid Network of Reservoirs and Connecting Links</b><br/>Qian Ge, Ke Han, Haoran Fu and Daisuke Fukuda</p> <p><b>Route Choice Equilibration Algorithm for Stochastic Dynamic Traffic Assignment with Full Route Set</b><br/>Jeroen Verstraete, Willem Himpe and Chris M.J. Tampère</p>                                                                                           |
| <b>Session C2 Dynamic Traffic Control II</b><br><i>Chair: Gunnar Flötteröd</i>                     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
| 11:00-12:40                                                                                        | <p><b>Modeling and Simulation of Air Pollution from the Dispersion of Vehicle Exhaust in a City: Continuum Modeling Approach</b><br/>Liangze Yang, Tingting Li, S.C. Wong, Chi-Wang Shu and Mengping Zhang</p> <p><b>Autonomous Intersection Management for Connected and Automated Vehicles Using Multi-Agent Reinforcement Learning</b><br/>Yuanyuan Wu, Shunzhi Qian and Feng Zhu</p> <p><b>Traffic Control Strategies for Urban Network: Centralised vs Decentralised Methods</b><br/>Roberta Di Pace, Andy H.F. Chow, Stefano de Luca and Rui Sha</p> <p><b>Dynamic Traffic Signal Control, Dynamic Queueing, and Dynamic Route Choice</b><br/>Gunnar Flötteröd and Michael J. Smith</p> |
| <b>Session C3 Day-to-day Dynamics II</b><br><i>Chair: David Z.W. Wang</i>                          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
| 14:00-15:40                                                                                        | <p><b>One Is Not Like the Other: Deterministic vs Stochastic Day-To-Day DTA Models: Part 1</b><br/>Katharina Parry, Ahmad Mahmoodjanlou and Martin Hazelton</p> <p><b>One Is Not Like the Other: Deterministic vs Stochastic Day-To-Day DTA Models: Part 2</b><br/>Ahmad Mahmoodjanlou, Martin Hazelton and Katharina Parry</p> <p><b>Day-To-Day Flow Dynamics for Stochastic User Equilibrium and a General Lyapunov Function</b><br/>Feng Xiao, Minyu Shen, Zhengtian Xu, Ruijie Li, Hai Yang and Yafeng Yin</p> <p><b>Day-To-Day Network Dynamics over Different Scales: From Micro to Macro</b><br/>Weimeng Li and Ning Jia</p>                                                           |
| <b>Session C4 Traffic and Demand Management II &amp; Closing</b><br><i>Chair: Michael W. Levin</i> |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
| 16:00-17:40                                                                                        | <p><b>Robust Multi-Period Tradable Credit Scheme to Promote Zero-Emissions Vehicles under Travel Demand Uncertainty</b><br/>Mohammad Miralinaghi and Srinivas Peeta</p> <p><b>Real-Time Dynamic Traffic Control Based on Traffic State Estimation</b><br/>Afzal Ahmed, Ahsan Naqvi, Dong Ngoduy and David Watling</p> <p><b>An Algorithm for the User Equilibrium with Agent-Based Dynamic Transit Assignment</b><br/>Mark Hickman</p> <p><b>An Optimization Model for a Station-Based Bike Sharing System Rebalancing Problem</b><br/>Yufeng Zhang, Alireza Khani</p> <p><b>Closing Speech and DTA2020</b><br/>Hong K. Lo</p>                                                                |

**Session A2**  
**Within-Day Equilibrium I**

# DYNAMIC USER EQUILIBRIUM WITH RECOURSE AND VARIABLE MESSAGE SIGNS

Kenneth A. PERRINE <sup>a</sup> and Michael W. LEVIN <sup>b</sup>

<sup>a</sup> *Center for Transportation Research,  
The University of Texas at Austin, USA*

*Email: [kperrine@utexas.edu](mailto:kperrine@utexas.edu)*

<sup>b</sup> *Department of Civil, Environmental, and Geo-Engineering,  
University of Minnesota, USA*

*Email: [mlevin@umn.edu](mailto:mlevin@umn.edu)*

Variable message signs (VMS) are used in cities to alert travelers to traffic congestion and incidents. Information can help travelers understand the incident state of the network. Once that is known through VMS or other dissemination mechanism, an appropriate action is to consider an alternative path. This is facilitated via an extension to the online shortest path (OSP) hyperpath problem that allows for the modeling of VMS. The dynamic user equilibrium with recourse (DUER) DTA methodology models each traveler's perception of the incident state via VMS as a stochastic process. DUER is tested on a model of a city's central business district. Results include an overall decrease in total system travel time, as well as several changes in route choice that specifically visit VMS sites. This can be used in any locale to evaluate proposed VMS locations, to analyze travel behavior, and to characterize effects of malfunctioning or cyber attacked VMS.

**Keywords:** Online Shortest Path, Dynamic User Equilibrium, Hyperpath, Variable Message Sign

## 1. INTRODUCTION

Many cities have implemented variable message signs (VMS), which can be used to alert travelers to traffic congestion or incidents. The purpose of this research is to more accurately model changes in travelers' route choices in response to VMS messages about incidents. Drivers who have experienced delays due to recurrent incidents will anticipate the possibility of lane closures before departing. Rather than wait until learning about the incident state before rerouting, drivers may proactively change their route to minimize the expected travel time. Drivers can significantly decrease their expected travel time by taking a slightly longer route to pass by a VMS.

A hyperpath is a routing policy or strategy that is responsive to information about travel times. Travelers following a hyperpath are likely to take different routes in normal conditions versus after learning that an incident is active. For example, a traveler might prefer to take a freeway corridor in normal conditions. However, if the traveler learns through a VMS that road work has resulted in lane closures on that freeway, the traveler may take an arterial corridor instead.

Although previous work on VMS has not considered hyperpath routing strategies, the problem of finding the optimal hyperpath is known as the online shortest path (OSP) problem, and was studied by Waller and Ziliaskopoulos (2002). Our work differs from Waller and Ziliaskopoulos (2002) in that we assume information may be acquired through VMS located far from the incident itself. We also assume that travelers consider the congestion caused by other travelers in their routing decisions. This is typically known as the user equilibrium assumption. Here, since travelers follow a hyperpath rather than a fixed path, we assume that traveler routing decisions form a dynamic user equilibrium with recourse (DUER). Static UER was previously studied by Unnikrishnan and Waller (2009), and Gao (2012) extended their work to dynamic flow models. As with OSP, the DUER model and results change significantly due to information provisioned by VMS located far from incidents.

Incidents are essentially "fixed" from a vehicle's perspective. However, each vehicle's perception of the incident state via VMS is a stochastic process. While traveling through the network, the incident state perception may update by receiving information through VMS or through visually observing incidents.

Table 1: Scenarios for evaluating the methodology

| Scenario | Description                                                           |
|----------|-----------------------------------------------------------------------|
| A        | 1 VMS with high adherence 0.8 and 1 incident of high probability      |
| B        | 1 VMS with high adherence 0.8 and 10 incidents of low probability     |
| C        | 6 VMS with low adherence 0.1 and 10 incidents of low probability      |
| D        | 6 VMS with moderate adherence 0.5 and 10 incidents of low probability |
| E        | 6 VMS with high adherence 0.8 and 10 incidents of low probability     |

## 2. FORMULATION SUMMARY

Consider a traffic network  $G = (N, A)$  with set of nodes  $N$  and set of links  $A$ . Let  $S$  be the set of possible incidents, which represent changes in travel times due to temporary events including construction, collisions, or weather. Each incident  $\sigma_{\text{net}} \in S$  delays travel on one or more links.

The true incident state of the network is  $\sigma_{\text{net}} \in S$ , which includes the possibility of no incident ( $\sigma_{\text{net}} = \diamond$ ), which are expected to occur during one travel period, such as the morning peak. Unless a VMS is placed at traveler origins, travelers do not know the incident state before departing. Incident occurs with probability  $p_\sigma$ . Each vehicle may perceive incident state as each travels past a VMS. Let  $S_n$  be the set of incidents that the VMS at link  $a \in A$  provides information about. If  $\sigma_{\text{net}} \in S_n$ , then the VMS at  $a$  will communicate that the incident state is  $\sigma_{\text{net}}$  with probability  $\rho_{\sigma_{\text{net}}}^a$ . This can be interpreted as the probability of driver adherence;  $1 - \rho_{\sigma_{\text{net}}}^a$  is the probability that the driver ignores the message.

The online shortest path with information (OSP-I) algorithm defined here is an all-to-one algorithm, and will find the optimal hyperpath from every origin to a single destination. Each vehicle will initially choose routes assuming that perceived incident  $\sigma_{\text{per}} = \diamond$ , and may change its route as  $\sigma_{\text{per}}$  changes. The OSP-I algorithm is further developed using Markov decision process theory, and employs the user equilibrium assumption. The formulation is a non-discounted, infinite horizon Markov decision process (MDP) with a termination state that can be solved using standard solution methods. We refer the reader to Bertsekas (2012) for analytical results and solution algorithms for MDPs.

We determined the optimal hyperpaths for individual vehicles, equivalent to the shortest path problem for traffic assignment. We now seek to study the effects of VMS on the network by finding the equilibrium hyperpaths. Intuitively, the optimal hyperpath depends on the choices made by other vehicles. It is likely that some vehicles will choose a hyperpath that specifies diversion, and other vehicles will choose to remain on the incident corridor. The problem of finding the equilibrium hyperpaths is known as user equilibrium with recourse (UER). Our work differs from Unnikrishnan and Waller (2009) and Gao (2012) in that drivers can learn of incidents from VMS in addition to observing nearby traffic conditions.

We present a variational inequality to find UER to formally define the problem. However, our objective is not to solve a static traffic assignment problem. Rather, we develop an algorithm for solving DUER, using a mesoscopic traffic flow model including congestion propagation and queue spillback.

We use the method of successive averages (MSA) to find a DUER assignment. While not provable, MSA empirically works well for finding a dynamic user equilibrium (Levin et al., 2015), and we expect to show that it appears to converge for the DUER problem also.

For the experiment design, we use an Austin, Texas, USA model as shown in Figure 1 with AM peak demand applied for 5 hours, and set the simulation duration for 10 hours. The model consists of several scenarios where the VMS placement and possible incident locations are varied. The covered scenarios are identified in Table 1.

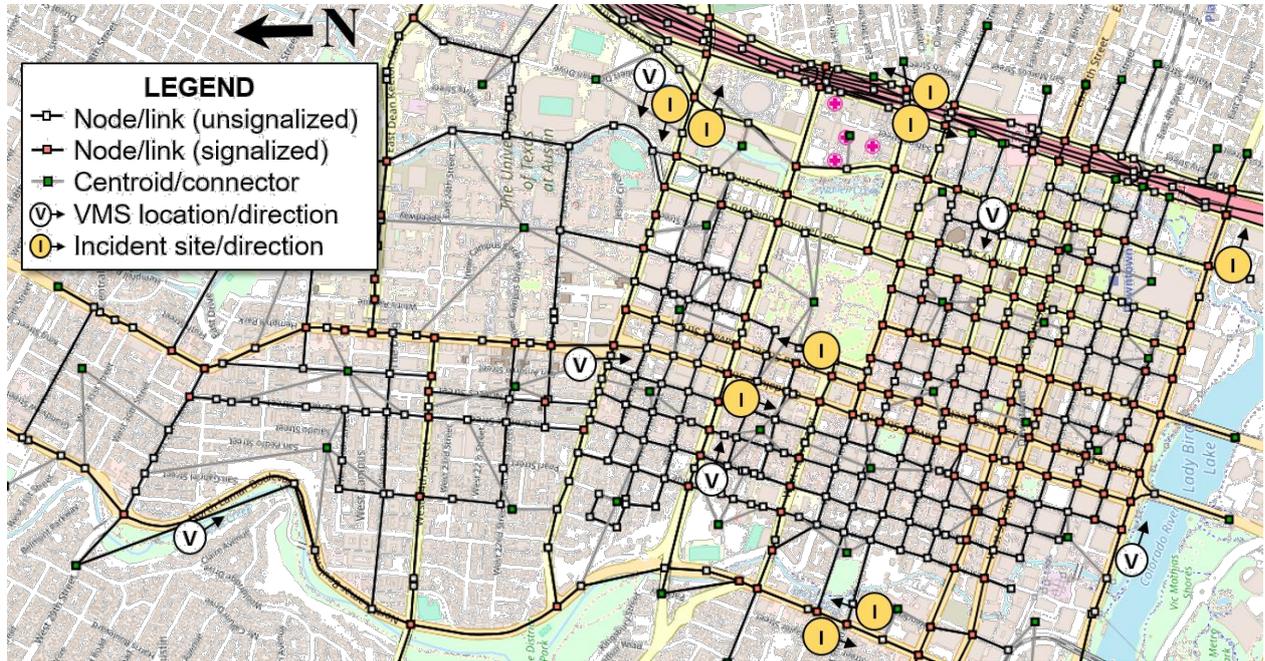


Figure 1. Network model of downtown Austin, Texas, USA with preliminary VMS and incident locations Map imagery: OpenStreetMap contributors.

Scenarios **A**, **B**, and **C/D/E** include a base case where VMS is completely ignored ( $\rho_{\sigma_{net}}^a = 0$ ), and cases where VMS is adhered by drivers as specified. Total system travel time is compared to measure impact. Because of the discrete nature of running a hyperpath assignment and simulation for each incident, the multiple-incident cases are designed to account for the most problematic roadway segments within the downtown area.

Further exploration is made in Scenarios **A** and **B** to investigate the hyperpath choice and corresponding comparative time savings in following VMS. Scenario **C** is designed to reflect a poorly-designed VMS system where drivers don't heed advice because of erroneous information, cybersecurity attack (Olofsson (2014)) or system malfunction. Scenarios **D** and **E** implement comprehensive use cases and show sensitivity when adherence rate  $\rho_{\sigma_{net}}^a$  is varied.

### 3. EXPECTED RESULTS AND FUTURE WORK

Because of the benefit of obtaining information and altering route choice, it is expected that the impact of VMS significantly decreases the total system travel time from the case where no incident occurs. As part of the Scenario **A** and **B** hyperpath explorations, driver route choice is investigated. For the comprehensive use cases Scenarios **D** and **E**, it is appropriate to work with different  $\rho_{\sigma_{net}}^a$  values because VMS adherence varies dramatically from region to region (Chatterjee and McDonald (2004)), and establishing an exact probability for a given region is difficult (Peeta et al. (2000)).

Scenarios can also be designed to look at the effects of messages that indicate ways to achieve faster travel times, such as an opening of a managed lane to all traffic—an “anti-incident”. This would be modeled by increasing roadway capacity in the incident location, versus decreasing it. As with lane-blocking incidents,

it is also expected that some drivers will alter their route choice to gather information on the possibility of a positive event.

This research is a first attempt at analyzing the impact of VMS (and other information sources such as messages conveyed to connected vehicles) by employing DUER to achieve travel time equilibrium given the use of hyperpaths. Each vehicles' discovery of information is modeled stochastically and paired with an MSA scheme for vehicle assignment. It is intended that this methodology will be useful within any locale for evaluating the placement of VMS signs to yield a greater impact, analyze travel behavior effects on the transportation system, and to characterize the effect of disrupted signs, either through malfunction, incorrect information, or cyber attack.

## REFERENCES

- Bertsekas, D.P. (2012) *Dynamic programming and optimal control, Volume 2*. Athena Scientific Belmont, MA.
- Chatterjee, K. and McDonald, M. (2004) Effectiveness of using variable message signs to disseminate dynamic traffic information: Evidence from field trials in European cities. *Transport Reviews*, 24(5), pp. 559–585.
- Gao, S. (2012) Modeling strategic route choice and real-time information impacts in stochastic and time-dependent networks. *IEEE Transactions on Intelligent Transportation Systems*, 13(3), pp. 1298–1311.
- Levin, M.W., Pool, M., Owens, T., Juri, N.R. and Waller, S.T. (2015) Improving the convergence of simulation-based dynamic traffic assignment methodologies. *Networks and Spatial Economics*, 15(3), pp. 1–22.
- Olofsson, J. (2014) Zombies ahead! A study of how hacked digital road signs destabilize the physical space of roadways. *Visual Communication*, 13, pp. 75–93.
- Peeta, S., Ramos, J. and Pasupathy, R. (2000) Content of variable message signs and on-line driver behavior. *Transportation Research Record: Journal of the Transportation Research Board*, 1725, pp. 102–108.
- Unnikrishnan, A. and Waller, S.T. (2009) User equilibrium with recourse. *Networks and Spatial Economics*, 9(4), pp. 575.
- Waller, S.T. and Ziliaskopoulos, A.K. (2002) On the online shortest path problem with limited arc cost dependencies. *Networks*, 40(4), pp. 216–227.

# MODELING LINK-BASED DYNAMIC USER EQUILIBRIA WITH DIFFERENTIAL COMPLEMENTARITY SYSTEMS

Xuegang (Jeff) BAN<sup>a</sup>

<sup>a</sup> *Department of Civil and Environmental Engineering,  
The University of Washington, US  
Email: [banx@uw.edu](mailto:banx@uw.edu)*

We introduce differential complementarity systems (DCS) and recent research efforts of applying DCS to model and solve link-based dynamic user equilibrium problems. Challenges and future research directions are also summarized and discussed.

**Keywords:** Dynamic User Equilibrium, Dynamic Traffic Assignment, Differential Complementarity Systems, Link-based Formulation, Solution Convergence

## 1. INTRODUCTION AND MOTIVATION

Dynamic user equilibrium (DUE) is one of the key problems of the Dynamic Traffic Assignment (DTA), which has received extensive attention and research for the last a few decades. This paper does not aim to provide a comprehensive review of the literature of DTA or DUE; readers can refer to [24, 20, 16, 6] for more detailed review of DUE/DTA. DUE, in its most basic form, needs to properly integrate two essential components: choice behavior of travelers to describe their mode, route, departure-time and other related choices, and traffic flow dynamics to describe the temporal and spatial evolution of dynamic traffic flow (such as how queues build up and dissipate) on transportation networks. The former can be usually formulated as optimization problems or equilibrium problems, while the latter can be often described as ordinary differential equations (ODEs) or partial differential equations (PDEs). Figure 1 depicts these two basic components of DUE. As indicated in Ma et al. [6], most DUE formulations in the literature started with continuous-time formulations, which were often quickly converted to their discrete-time counterparts, resulting in various finite dimension optimization problems such as nonlinear programming problems, variational inequalities (VI), and nonlinear complementarity problems (NCP). This essentially by passed rigorous investigations of the continuous-time formulations of DUE, as indicated by the dashed bold lines in Figure 1. This is partially due to the fact that a rigorous mathematical framework that can properly and systematically integrate the two components (i.e., the choice models and the traffic dynamics) in continuous time had not really existed, until very recently.

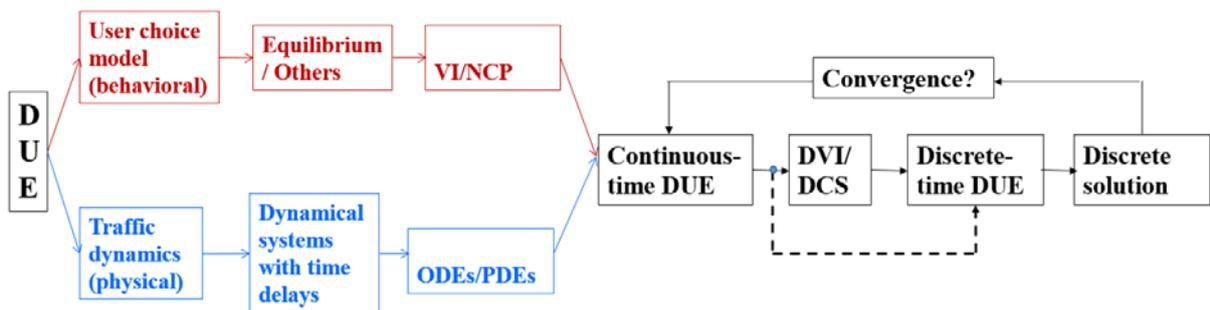


Figure 1: DUE Research Methods

As already been pointed out in the past, continuous-time and discrete-time DUE formulations are very different [8, 2]. Ignoring the formal investigation of continuous-time DUE problems can lead to a few theoretical or practical issues. First, there are potential inconsistencies between the continuous-time descriptions of DUE and its discrete-time solutions. That is, depending on how the discrete-time models are constructed, the obtained discrete-time solutions may or may not converge to solutions of continuous-time DUE models as time goes to infinitesimally small. More critically, since certain approximations are usually needed to develop discrete-time DUE formulations, the obtained discrete-solutions may indeed not be related at all to the continuous-time descriptions of the DUE problem.

Without a formal investigation of the continuous-time formulations, such convergence issues are almost impossible to be studied in a rigorous way. Secondly, studying the continuous-time DUE formulations and how to discrete them can also help provide insights on developing proper discrete-time DUE models, e.g., how to select the step size of discrete-time and the stability of the discretization scheme, etc. [2].

Recently, a new mathematical paradigm, called differential variational inequality (DVI), has been developed in the optimization and dynamic game fields [22, 3]. DVI aims to integrate two distinctly different components, namely the dynamic systems (often formulated mathematically as ODEs) and the variational inequality at each time instant. This provides a suitable and much-needed mathematical framework for continuous-time DUE, as shown in Figure 1. More recently, DVI and its special case DCS (differential complementarity systems) have been applied to model and solve DUE problems in continuous time, using both path-based formulations [4] and link-based formulation [2, 6].

The DVI (or DCS) based DUE formulations so far have been mainly focused on vehicular traffic, and in certain cases for specific DUE problems such as instantaneous DUEs [2] or DUEs on single-destination networks [6]. These models need to be extended to capture more general (such as multi-modal) DUE problems. More importantly, with the emergence of new technologies and systems in transportation (such as connected/automated vehicles (CAVs), shared mobility, and big data), DUE models need to capture/integrate those emerging systems in a systematic way. There have been initial attempts in those regards recently, but much more needs to be done. The purpose of this paper is to introduce DVI/DVS and provide a brief summary of what has been done in the last a few years regarding applying DVI/DCS to formulate link-based DUE problems. We will also provide some detailed discussions on a few challenging DUE related questions particularly concerning the application of DVI/DCS, which calls for collective research efforts from researchers in the DVI/DCS and DTA/DUE fields. Note that we focus on link-based DUE formulations in this paper and their specific issues. Path-based formulations have also been widely studied using DVI and one can refer to [4] for more details.

## 2. LITERATURE REVIEW

### 2.1 DVI and DCS

DVI can be defined in two ways by treating its two key components slightly differently. In [16], DVI is formulated as an infinite dimensional VI with ODEs cast as part of the constraint set. In [22], DVI is formulated as an ODE parameterized by an algebraic variable that is required to be a solution of a finite-dimensional state-dependent VI problem. In this paper, we follow the DVI definition in [22]; similar analysis may also be done if the DVI formulation in [16] is used. To define a DVI mathematically, we need to define a terminal time  $T > 0$ , and vector functions  $f : \mathbb{R}^{1+n+m} \rightarrow \mathbb{R}^n$  and  $F : \mathbb{R}^{1+n+m} \rightarrow \mathbb{R}^m$ . Then DVI can be defined as finding trajectories of (state variable)  $x : [0, T] \rightarrow \mathbb{R}^n$  and (control variable)  $u : [0, T] \rightarrow \mathbb{R}^m$  such that for almost all  $t \in [0, T]$ ,

$$\begin{aligned}
 \dot{x}(t) &\triangleq \frac{dx(t)}{dt} = f(t, x(t), u(t)) && \text{dynamics} \\
 u(t) &\in \text{SOL}(U, F(t, x(t), \bullet)) && \text{variational constraint} \\
 x(0) &= x^0 && \text{initial condition,}
 \end{aligned} \tag{1}$$

Here  $U$  is nonempty closed convex set in  $\mathbb{R}^m$ .  $\forall I(U, F)$  denotes the VI problem defined by the set  $U$  and function  $F$ , and  $\text{Sol}(U, F)$  denotes the solution set of  $\forall I(U, F)$ . Clearly, a DVI is to solve an (continuous-time) ODE and requires a finite dimensional VI holds at (almost each time instant)  $t$ . Thus DVIs are hard to solve and solutions to a DVI are typically at best piecewise differentiable. Therefore, one often seeks a weak solution of DVI in the sense of Caratheodory [22]. Specifically,  $(x, u)$  is a weak

solution of the DCS (1) if  $x$  is absolutely continuous and  $u$  is integrable on  $[0, T]$  such that  $x(0) = x_0$  and:

(a) the differential equation is satisfied for almost all  $t \in [0, T]$ , or equivalently, the integral condition:

$$x(t) = x(s) + \int_s^t f(\tau, x(\tau), u(\tau)) d\tau \text{ holds for all } 0 \leq s \leq t \leq T;$$

(b) the VI is satisfied for almost all  $t \in [0, T]$ , i.e.,  $\forall$  continuous  $\hat{u}(\tau) : [0, T] \rightarrow U$  such that

$$\int_0^T (\hat{u}(\tau) - \mu(\tau))^T F(\tau, x(\tau), u(\tau)) d\tau = 0.$$

A special case of DVI is the differential complementarity system (DCS), which can be defined as finding  $(x(t), u(t))$  such that the following hold:

$$\begin{aligned} \dot{x}(t) &\triangleq \frac{dx(t)}{dt} = f(t, x(t), u(t)) && \text{dynamics} \\ 0 \leq u(t) \perp F(t, x(t), u(t)) \geq 0 &&& \text{complementarity constraint} \\ x(0) &= x^0 && \text{initial condition,} \end{aligned} \quad (2)$$

Here  $u \perp w$  means that the vector  $u$  and  $v$  are perpendicular, i.e.,  $u^T v = 0$ . The connection between DCS and DVI is similar to the connection between VI and NCP. Especially as a special case of DVI, there are more specialized (and often more effective) methods to analyze and solve DCS. More analysis methods for DVI/DCS can be found in [11, 13, 14, 15, 17, 18, 19, 21, 23, 25, 26]. Those works particularly discussed how to discretize a DVI, how to obtain a discrete-time solution, how to construct a continuous-time solution from discrete-time solutions, and under what conditions the discrete-time solutions will convert to a continuous-time solution. In the full paper, some of the classical results will be summarized especially those related to the analysis and solution of DUE. Here it is easy to see that naturally the DVI/DCS fits the DUE problem well as it treats dynamics and variational constraints (or complementarity conditions) in an integrated framework.

## 2.2 Link-based DUE formulations

A general link-based DUE formulation, by applying the DVI/DCS framework, is given in [2]. It is given here with some minor changes. First, a transportation network is denoted as  $G(N, A)$ , where  $N$  and  $A$  are the sets of nodes and links respectively.  $S \in A$  is the set of destinations. Further denote  $x_{ij}^s(t)$  the amount of vehicular flow on link  $(i, j) \in L$  destined for node  $s \in S$ ,  $\tau_{ij}(t)$  the travel time of vehicular flow traversing link  $(i, j)$  when entering the link at time  $t$ ,  $p_{ij}^s(t)$  the rate of entry flow on link  $(i, j) \in L$  destined for node  $s \in S$ ,  $v_{ij}^s(t)$  the rate of exit flow on link  $(i, j) \in L$  destined for node  $s \in S$ , and  $\pi_{ij}^s(t)$  the minimum travel time from node  $i \in N$  to destination  $s$  at time  $t$ . Then the DCS-based DUE model (denoted as DUE-G for a general DUE model) can be presented as below

**DUE-G.** Find a weak solution satisfying for all  $s \in S$  and for almost all  $t \in (0, T]$ ,

- dynamics: for all  $(i, j) \in \mathcal{L}$ ,

$$\begin{cases} \dot{x}_{ij}^s(t) = p_{ij}^s(t) - v_{ij}^s(t) & \text{flow balance} \\ \dot{\tau}_{ij}(t) = \frac{p_{ij}^s(t)}{v_{ij}^s(t + \tau_{ij}(t))} - 1 & \text{flow propagation} \end{cases}$$

- Wardrop's route choice principle: for all  $(i, j) \in \mathcal{L}$ ,

$$0 \leq p_{ij}^s(t) \perp \tau_{ij}(t) + \pi_j^s(t + \tau_{ij}(t)) - \pi_i^s(t) \geq 0$$

- flow conservation: for all  $i \in \mathcal{N}$ :

$$0 \leq \pi_i^s(t) \perp \sum_{j:(i,j) \in \mathcal{L}_s} p_{ij}^s(t) - \sum_{k:(k,i) \in \mathcal{L}_s} v_{ki}^s(t) - d_i^s(t) \geq 0$$

- $\pi_s^s(t) = 0$
- link travel time function: for all  $(i, j) \in \mathcal{L}$ ,  $\tau_{ij}(t) = \tau_{ij}(x(t))$
- initial conditions: for all  $i \in \mathcal{N}$ ,  $x_{ij}^s(0) = \xi_{ij}^s$ .

Notice the time delayed terms in the flow propagation dynamics and route choice complementarity conditions, which introduce more challenges to the classical DVI/DCS problems. Also note that to properly model the link travel time function, certain traffic dynamics model needs to be employed. Depending on what specific traffic models are used, different DCS-based DUE formulations can be derived. Additional state or auxiliary variables may also need to be introduced such as queue lengths and the flow withholding variables [6]. For example, Ban et al. [2] applied the point-queue model for instantaneous DUE problems (i.e., route choices are determined by prevailing traffic conditions instead of predictive traffic conditions), leading to a DCS problem with constant time delays. Discretization schemes and methods to construct continuous-time solutions from discrete-time solutions were also provided in [2]. It was further shown that, under mild conditions, the discrete-time solutions can converge to a solution of the (original) continuous-time DCS-based DUE problem as the discrete time step goes to infinitesimally small. In [6], a physical queue model was applied to model predictive DUE problems on networks with single destinations. A special link transmission model, called the double-queue model was particularly used, which was first proposed in [10] and later extended in [7] and [5]. When physical queue model is applied, one challenge is to capture queue spillbacks, a phenomenon that happens when congestion from the downstream of a link propagates to the entrance of the link, potentially restricting inflows to the link (at a later time). By introducing flow withholding priorities and auxiliary variables, a node model was developed in [6] to resolve the merging/diverging and queue spillback problem. Certain approximation schemes were also introduced to reduce the time-delayed terms from time-varying, state-dependent delays to constant delays. The models and analysis in [6] however were only for single-destination transportation networks. Extending the results to multi-destination networks needs to resolve the first-in-first-out (FIFO) problems especially at junctions.

In the full paper, more results related to link-based DCS formulations for DUE will be provided, with more discussions on the major findings, implications, and challenges to solve more general problems. This will lay the foundation for research challenges and future research directions with respect to applying DVI/DCS to model and solve link-based DUE.

### 3. RESEARCH CHALLENGES AND FUTURE RESEARCH DIRECTIONS

We summarize a few challenges of applying the current DVI/DCS framework for DUE and suggest several future research directions. First, the classical DVI/DCS framework (e.g., [22]) does not deal with time-delayed terms directly, while time delays are an essential part of DUE. In order for the DVI/DCS framework to be more suitable for DUE, time-delayed DVI/DCS problems need to be studied rigorously,

particularly its solution methods and convergence analysis. It seems that DVI/DCS with constant delays may be a straightforward extension of DVI/DCS without delays as shown in Ban et al. [2]. Therefore, future research may focus on DVI/DCS with time-varying, state-independent delays first and then extend to time-varying, state-dependent delays. Secondly, the current link-based DCS model for DUE needs to be extended to multi-destination networks. For this, how to ensure FIFO especially in the node model is the key, as discussed above. Thirdly, DVI/DCS based DUE formulations so far have been mainly focused on vehicular traffic particularly passenger cars. The models need to be extended to include more general, realistic scenarios. These may include, e.g., multi-modal transportation networks that may include transit (buses, subways). For this, the methods of modeling the interactions of multi-modal dynamic traffic flow in [9] may be used, which captured the interactions at transferring stations or stops by modeling person queues as point queues. More importantly, as emerging modes, such as ridesourcing (e.g., Uber, Lyft, Didi, among others) and ridesharing, are becoming more popular in urban traffic networks, how to capture the behavior and impact of these new modes in transportation networks is crucial for dynamic transportation network modeling in the near future. This poses both opportunities and challenges when applying the DVI/DCS framework, which calls for further investigations. Last but not least, existing dynamic network models are mostly based on fairly restrictive mathematical principles (such as equilibrium etc.). As data are increasing available in transportation, dynamic network models need to systematically integrate various sources of data into the modeling framework to improve modeling accuracy and realism. The DVI/DCS framework may provide a platform for such data-driven modeling methods for DUE, which is an interesting future research direction that merits further in-depth investigations. In the full paper, we will provide more detailed discussions with respect to the above (and maybe other) identified challenges and future research directions. Thoughts will also be provided on how such challenges may be addressed.

## REFERENCES

- [1] X. Ban and J.S. Pang and H.X. Liu and R. Ma., 2012. Continuous-time point-queue models in dynamic network loading. *Transportation Research Part B* 46(3), 360–380.
- [2] X. Ban, J.S. Pang and H.X. Liu and R. Ma., 2012. Modeling and solving continuous-time instantaneous dynamic user equilibria: A differential complementarity systems approach. *Transportation Research Part B* 46(3), 389–408.
- [3] T.L. Friesz, 2010. *Dynamic Optimization and Differential Games*. Springer, New York.
- [4] T.L. Friesz, K. Han, P.A. Neto, T. Yao, 2013. Dynamic user equilibrium based on a hydrodynamic model. *Transportation Research Part B* 47 102–126. Springer, New York.
- [5] Ma, R., X. Ban, J.S. Pang., 2014. Continuous-time dynamic system optimum for single-destination traffic networks with queue spillbacks. *Transportation Research Part B* 68, 98–122.
- [6] Ma, R., X. Ban, J.S. Pang. 2017. A link-based differential complementarity system formulation for continuous-time dynamic user equilibria with queue spillbacks. *Transportation Science*, forthcoming.
- [7] Osorio, C., Flötteröd, G., Bierlaire, M., 2011. Dynamic network loading: A stochastic differentiable model that derives link state distributions. *Transportation Research Part B* 45 (9), 1410–1423.
- [8] Wie, B.W., R.L. Tobin, M. Carey. 2002. The existence, uniqueness and computation of an arc-based dynamic network user equilibrium formulation. *Transportation Research Part B* 36(10), 897–918.
- [9] Yang, X., Ban, X., Mitchell, J., 2017. Modeling multimodal transportation network emergency evacuation considering evacuees cooperative behavior. *Transportation Research Part A*, accepted.
- [10] Yperman, I. 2007. *The link transmission model for dynamic network loading*. (Doctoral dissertation). Retrieved from <https://lirias.kuleuven.be/bitstream/1979/946/2/phd-final-all.pdf>.
- [11] M.K. Camlibel. Popov-Belevitch-Hautus type tests for the controllability of linear complementarity systems. *Systems and Control Letters* 56 (2007) 381–387.
- [12] M.K. Camlibel, W.P.M.H. Heemels, and J.M. Schumacher. Algebraic necessary and sufficient conditions for the controllability of conewise linear systems *IEEE Transactions on Automatic Control* 53 (2008) 762–774.
- [13] M.K. Camlibel, J.S. Pang, and J. Shen. Conewise linear systems: Non-Zenoness and observability. *SIAM Journal on Control and Optimization* 45 (2006) 1769–1800.

- [14] M.K. Camlibel, J.S. Pang, and J.L. Shen. Lyapunov stability of linear complementarity systems, *SIAM Journal on Optimization* 17 (2006) 1056–110.
- [15] M.K. Camlibel and J.M. Schumacher. Existence and uniqueness of solutions for a class of piecewise linear dynamical systems. *Linear Algebra and its Applications* 351–352 (2002) 147–184.
- [16] T.L. Friesz and T. Kim and C. Kwon and M.A. Rigdon. Approximate network loading and dual-time-scale dynamic user equilibrium. *Transportation Research Part B* 45 (2010) 207–229.
- [17] L. Han and J.S. Pang. Non-Zenoness of a class of differential quasi-variational inequalities, *Mathematical Programming, Series A* 121 (2010) 171–199.
- [18] L. Han, A. Tiwari, M.K. Camlibel, and J.S. Pang. Convergence of time-stepping schemes for passive and extended linear complementarity systems. *SIAM Journal on Numerical Analysis* 47 (2009) 1974–1985.
- [19] W.P.M.H. Heemels, J.M. Schumacher, and S. Weiland. Linear complementarity systems. *SIAM Journal on Applied Mathematics* 60 (2000) 1234–1269.
- [20] S. Peeta and A.K. Ziliaskopoulos. Foundations of dynamic traffic assignment: the past, the present and the future. *Networks and Spatial Economics* 1(3-4) (2001) 233–265.
- [21] J.S. Pang and J. Shen. Strongly regular differential variational systems. *IEEE Transactions on Automatic Control* 52 (2007) 242–255.
- [22] J.S. Pang and D.E. Stewart. Differential variational inequalities, *Mathematical Programming, Series A* 113 (2008) 345–424.
- [23] J.S. Pang and D.E. Stewart. Solution dependence on initial conditions in differential variational inequalities. *Mathematical Programming, Series B* 116 (2009) 429–460.
- [24] B. Ran and D.E. Boyce. *Modeling Dynamic Transportation Networks: An Intelligent Transportation Systems Oriented Approach*. Springer-Verlag (New York 1996).
- [25] J.L. Shen and J.S. Pang. Semicopositive linear complementarity systems, *International Journal on Robust and Nonlinear Control* 17 (2007) 1367–1386.
- [26] J.L. Shen and J.S. Pang. Linear Complementarity systems: Zeno states, *SIAM Journal on Control and Optimization* 44 (2005) 1040–1066.

# A SCENARIO-BASED APPROACH FOR THE RELIABILITY-BASED USER EQUILIBRIUM PROBLEM IN DYNAMIC STOCHASTIC NETWORKS CONSIDERING TRAVEL TIME CORRELATIONS AND HETEROGENEOUS USERS

A. ZOCCAIE<sup>a</sup>, H. S. MAHMASSANI<sup>b</sup> and F. FAKHRMOOSAVI<sup>c</sup>

<sup>a</sup> Assistant Professor, Department of Civil and Environmental Engineering,  
Michigan State University, US

Email: [zockaiea@egr.msu.edu](mailto:zockaiea@egr.msu.edu)

<sup>b</sup> William A. Patterson Distinguished Chair in Transportation,  
Northwestern University Transportation Center, US

Email: [masmah@northwestern.edu](mailto:masmah@northwestern.edu)

<sup>c</sup> Doctoral Researcher, Michigan State University, US

Email: [moosavi@msu.edu](mailto:moosavi@msu.edu)

Travel time reliability needs to be considered in transportation network planning, as it is a major factor in travelers' route choice. Furthermore, each traveler's route choice decision, considering the reliability measure, affects the reliability of travel time for other travelers in the network. This calls for considering the reliability-based user-equilibrium (RBUE) problem in stochastic networks. Studies to date regarding RBUE problem formulations and solution algorithms have been limited to simplified assumptions or very small networks. This study aims to propose a methodology to solve RBUE problem in stochastic time-varying networks considering travel time correlations and heterogeneous users with actual applications in realistic large-scale networks. The presented methodology updates link travel time distributions and correlations based on flow-dependent relations. Multiple simulation runs are conducted considering scenarios based on real-world observations. An iterative solution procedure is presented in this study for the RBUE traffic assignment problem, which minimizes a defined gap function. The approach is implemented and tested on an actual large-scale network of Chicago, where 86 scenarios are simulated to generate the initial travel time distributions and correlations. The numerical results show satisfactory application of the algorithm and its sensitivity to the valuation of reliability by users and the correlation structure.

**Keywords:** Reliability, User Equilibrium, Stochastic, Heterogeneous Users, Time Varying Networks

## 1. INTRODUCTION

Travel time reliability plays a pivotal role in transportation network planning and traffic operations. Therefore, transportation policymakers need to incorporate the travel time variability in addition to the average travel time in their models. The sources of travel time variability in transportation networks are related to both supply-side and the demand-side (Nie 2011; Kim et al. 2013). Different users respond differently to the resulting uncertainty in travel times, reflecting heterogeneity in individual preferences and risk attitudes. Furthermore, link travel time distributions (TTD) are related to link flows in transportation networks. Thus, users' path selection decisions considering the reliability measures affect the experience of other travelers in the network. This calls for considering the reliability-based user equilibrium (RBUE) problem in stochastic networks.

When users consider a reliability measure in their path choices in stochastic time-varying (STV) networks, their decisions would naturally impact the congestion patterns that, in turn, determine travel times and their distributional properties. Therefore, studies to date regarding RBUE problem formulations and solution algorithms have been limited to simplified assumptions or very small networks (Shao et al. 2006; Lo et al. 2006; Chen et al. 2011; Nie 2011; Ng & Waller 2012; Wang et al. 2014; Xu et al. 2017). None of the aforementioned studies consider time-dependent networks. Furthermore, all above studies with the exception of Chen et al. (2011) require complete path enumeration, which makes the solution procedure non-tractable for large-scale networks, and do not consider link travel time correlations for the path finding sub-problem. More recently, Zockaie et al. (2018) presented a formulation framework for the RBUE problem that is capable of capturing the link travel time correlations for multiple classes of users. This study simplifies the relation between congestion patterns and link TTDs to an analytical relation (Kim 2014). Moreover, this study uses a

hypothetical relation to update link correlations. Therefore, there is still the need to account for the probabilistic nature of time-dependent travel times in the RBUE models more realistically.

The objective of this study is to propose a novel methodology to solve RBUE problem in STV networks considering travel time correlations and heterogeneous users with actual applications in realistic large-scale networks. This methodology is capable of updating link TTDs and correlations based on flow-dependent relations. To do so, multiple simulation runs are conducted considering scenarios (based on real-world observations, weather conditions, incidents, demand factors, and different adaptive drivers shares) with different weights and occurrence probabilities, in order to get time-dependent TTDs and correlations as well as updating them within each iteration of the descent direction method.

The non-linear minimization problem defined in Zockaie et al. (2018) is used in this study to represent the RBUE traffic assignment problem. In this problem, link travel times are congestion-dependent random variables that vary over time and exhibit general correlation patterns, reflecting spatial and temporal network flow processes. An iterative solution procedure is presented in this paper for this problem, which minimizes a defined gap function. A core element of this procedure is a path-finding algorithm that relies on a Monte-Carlo sampling approach, which finds the optimal path for multi-class users considering a given link TTD and correlation structure. The optimal path and its associated utility compared with the current used paths and their experienced utilities are used to update the traffic assignment at each iteration. Repeating traffic simulations for multiple scenarios relates congestion patterns (link flows) to link TTDs and correlations. The approach is implemented and tested on an actual large-scale network of Chicago, where 86 scenarios are simulated to generate the initial TTDs and correlations. To reduce the solution time, a randomly selected sub-set of scenarios are simulated again using the latest path flow assignment to update link TTDs and correlations at higher iterations. The numerical results show satisfactory application of the algorithm and its sensitivity to the valuation of reliability by users and the correlation structure.

## 2. PROBLEM STATEMENT

An RBUE state is attained when individual travelers cannot increase their utility by unilaterally changing their path. The utility function needs to be specified consistently with a certain reliability rule. Travel times along network links are random variables that follow a general non-stationary multivariate distribution with time-varying properties that depend on the network flow pattern. The utility of a traveler is defined based on a reliability measure depending on the path travel time distribution. When no better path can be found for any individual user in the network, the user equilibrium state is achieved that can be formulated as the following fixed-point problem:

$$\mathbf{r}^* = \text{GF}(\text{PF}(U(\mathbf{r}^*))) \quad (1)$$

$\mathbf{r}^*$  is the path flow assignment at the reliability-based user equilibrium state,  $U(\mathbf{r}^*)$  gives the experienced utility vector considering  $\mathbf{r}^*$  path flow assignment,  $\text{PF}$  finds the optimal paths with the maximum utilities considering experienced utilities  $U$ , and  $\text{GF}$  reassigns the path flow vector based on the optimal paths found by  $\text{PF}$ . Based on (Zockaie et al. 2018), a gap function can be defined for the equivalent VI formulation of the fixed-point problem to substitute the main problem with a non-linear optimization problem as follows:

$$\text{Minimize} \left[ \sum_{\forall o \in O} \sum_{d \in D} \sum_{\tau \in T} \sum_{c \in C} \sum_{p \in P(o,d,\tau,c)} \mathbf{r}_{\text{tcp}}^{\text{od}} \times (U_{\text{tc}}^{\text{od}}(\mathbf{r}) - U_{\text{tcp}}^{\text{od}}(\mathbf{r})) \right] \quad (2)$$

$$\sum_{p \in P(o,d,\tau,c)} \mathbf{r}_{\text{tcp}}^{\text{od}} = \mathbf{q}_{\text{tc}}^{\text{od}} \quad (3)$$

$$\forall o \in O, d \in D, \tau \in T, c \in C$$

$$\mathbf{r}_{\text{tcp}}^{\text{od}} \geq 0, U_{\text{tc}}^{\text{od}}(\mathbf{r}) - U_{\text{tcp}}^{\text{od}}(\mathbf{r}) \geq 0 \quad (4)$$

$$\forall o \in O, d \in D, \tau \in T, c \in C, p \in P(o, d, \tau, c)$$

$q_{\tau c}^{od}$  is the time-dependent demand for users' class  $c$ , traveling from origin  $o$ , to destination  $d$ , at departure time interval  $\tau$ ;  $P(o,d,\tau,c)$  is the set of all existing paths, and  $U_{\tau c}^{*od}$  is the optimal utility for all users of class  $c$ , traveling from  $o$  to  $d$  at  $\tau$ , when we have the equilibrated path flow assignment,  $r^*$ .

### 3. SOLUTION ALGORITHM

The proposed solution algorithm in this study consists of several components; a reliability-based optimal path finding procedure, an algorithmic gap-based process, and a simulation-based network loading tool. The minimum travel time budget path (MTTBP) reliability rule (Zockaie et al. 2013 and 2014) is considered for the path finding sub-problem. Furthermore, a gap-based process for iteratively redistributing user choice outcomes is required to achieve the desired equilibrium state. To this end, travelers need to shift from non-optimal paths to the corresponding optimal path through a descent search direction. The amount of flow redirected to the “optimum” path(s) at each iteration is determined using efficient variants of the method of successive averages (MSA) (Sbayti et al. 2007). Finally, a simulation-based network loading method, DYNASMART-P (Mahmassani 2001), is required to get link flows from the path flow assignment considering congestion patterns and associated stochastic properties of travel times. The main contribution of this paper is to consider a scenario-based approach to relate link flows and link TTDs and correlations. This approach starts with the initial link travel time distributions and correlations resulted from simulating multiple scenarios. In each iteration, the path flow assignment simulation and network loading are repeated for a sub-set of main scenarios to update the link TTDs and correlations considering the link flow variations.

**Step 0. Initialization** (set  $k=0$ ): Repeat simulations for multiple scenarios and get the initial time-dependent TTDs and correlations in addition to the initial path flow assignment ( $r^k$ ).

**Step 1. Link travel time realization:** Use the simulation results and the spatial and temporal correlation structure to generate the time-dependent joint TTD for each link and time interval.

**Step 2. Experienced path travel time distribution**

**Step 2-1.** For each link and time interval, draw random numbers from the joint TTD.

**Step 2-2.** Search over all the travelers and create the simulated paths set for each  $o$ ,  $d$ ,  $\tau$ , and  $c$ .

**Step 2-3.** Using the reliability rule and the random numbers generated in Step 2-1, calculate the utility of each path and its TTD for each  $o$ ,  $d$ ,  $\tau$ , and  $c$ .

**Step 2-4.** Find the path with the maximum utility and store its utility.

**Step 3. Calculate the average gap and check the convergence:** If it is met, stop and report the optimal path assignment; otherwise go to Step 4.

**Step 4. Update path flow assignment for all destination zones (Descent direction)**

**Step 4-1.** Find time-dependent optimal paths for each  $o$ ,  $\tau$ ,  $c$  and add them to the used paths set.

**Step 4-2.** Calculate the utilities for all existing paths in the set.

**Step 4-3.** Search over all vehicles and apply a path swapping rule to decide to switch the vehicle current path to the corresponding optimal path or not. Set  $k$  as  $k+1$  and update  $r^k$

**Step 5. Simulate vehicles based on the path flow assignment:** Use  $r^k$  and repeat simulation of all vehicles under multiple scenarios to capture the congestion pattern. Add the new observations to the observations in prior iterations to generate updated link TTDs and correlations. Go to Step 1.

### 4. NUMERICAL RESULTS

The proposed scenario-based methodology is implemented for a large-scale network of Chicago with 1,578 nodes and 4,805 links. Time-dependent link TTDs and correlations are explored through simulation of 86 scenarios based on the real-world information. The simulation results are used as the initial TTDs and correlations (Zockaie et al. 2016). To reduce calculation time, 10% of scenarios are selected randomly in each iteration to be repeated and update TTDs and correlations. For temporal correlations, consecutive time intervals, and for the spatial correlation, the neighborhood order of 2 is considered. Note that the MTTBP reliability rule is considered for the implementation of this study with 0.3, 0.7, and 1 as the weights of path travel time standard deviation relative to the mean travel time. The shares of users are 0.25, 0.5, and 0.25. Three correlation levels are compared; independent, spatial, and

both temporal and spatial. Figure 1 demonstrates a strictly decreasing pattern in terms of average gap and disutility for different correlation considerations, which confirms the convergence of the solution algorithm. Note that as expected, the independent case has the smallest values in all iterations and the other two cases have similar results. Performance measures of different classes of users for the case with just spatial correlation are reported in Figure 2. A decreasing pattern can be observed for all user classes, and a higher reliability valuation results in larger performance measures.

## 5. CONCLUSION

In this paper, a methodology is presented and implemented for the RBUE problem in a large-scale stochastic time-varying network. Spatial and temporal correlations of link travel times are considered for different user classes in terms of the reliability valuation. The MTTBP reliability rule is applied to the path finding problem. The main contribution includes a scenario-based approach to relate congestion patterns of a path flow assignment to link TDDs and correlations. The numerical results confirm the convergence of the algorithm. Comparing the results with the analytical approach in (Zockaie 2018) reflects the more realistic reliability consideration due to the strictly decreasing pattern in the average gap value. This study addresses an important gap in the literature by coming up with a novel methodology to solve RBUE problem in STV networks considering heterogeneity of users, realistic travel time distributions, as well as spatial and temporal link travel time correlations.

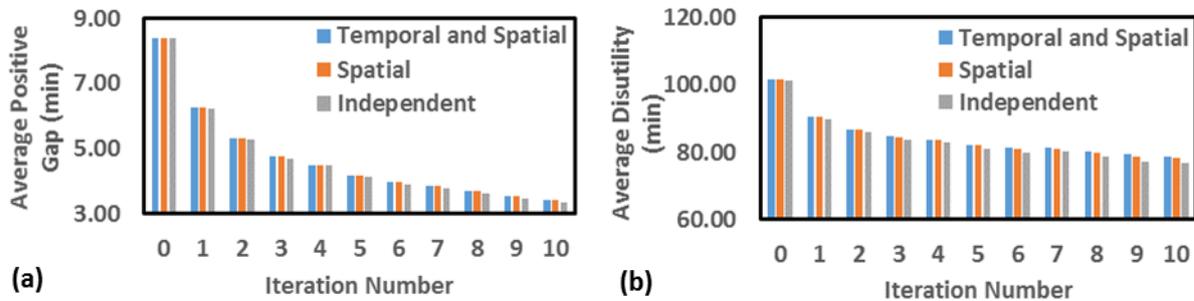


Figure 1. Different correlation patterns: (a) average gap, (b) average disutility

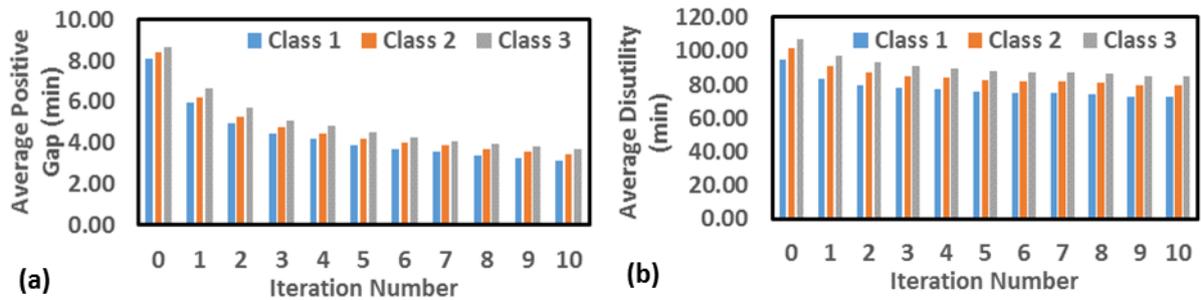


Figure 2. Multiple user classes with spatial correlations: (a) average gap, (b) average disutility

## REFERENCES

- Chen, B.Y. et al., 2011. An efficient solution algorithm for solving multi-class reliability-based traffic assignment problem. *Mathematical and Computer Modelling*, 54(5), pp.1428–1439.
- Kim, J. et al., 2013. Scenario-based approach to analysis of travel time reliability with traffic simulation models. *Transportation Research Record: Journal of the Transportation Research Board*, (2391), pp.56–68.
- Kim, J., 2014. *Travel Time Reliability of Traffic Networks: Characterization, Modeling and Scenario-based Simulation*. Northwestern University.
- Lo, H.K., Luo, X.W. & Siu, B.W.Y., 2006. Degradable transport network: travel time budget of travelers with heterogeneous risk aversion. *Transportation Research Part B: Methodological*, 40(9), pp.792–806.
- Mahmassani, H.S., 2001. Dynamic network traffic assignment and simulation methodology for advanced

- system management applications. *Networks and spatial economics*, 1(3), pp.267–292.
- Ng, M. & Waller, S.T., 2012. A dynamic route choice model considering uncertain capacities. *Computer-Aided Civil and Infrastructure Engineering*, 27(4), pp.231–243.
- Nie, Y.M., 2011. Multi-class percentile user equilibrium with flow-dependent stochasticity. *Transportation Research Part B: Methodological*, 45(10), pp.1641–1659.
- Sbayti, H., Lu, C.-C. & Mahmassani, H., 2007. Efficient implementation of method of successive averages in simulation-based dynamic traffic assignment models for large-scale network applications. *Transportation Research Record: Journal of the Transportation Research Board*, (2029), pp.22–30.
- Shao, H. et al., 2006. Demand-driven traffic assignment problem based on travel time reliability. *Transportation Research Record: Journal of the Transportation Research Board*, (1985), pp.220–230.
- Wang, J.Y.T., Ehrgott, M. & Chen, A., 2014. A bi-objective user equilibrium model of travel time reliability in a road network. *Transportation Research Part B: Methodological*, 66, pp.4–15.
- Xu, X. et al., 2017. A link-based mean-excess traffic equilibrium model under uncertainty. *Transportation Research Part B: Methodological*, 95, pp.53–75.
- Zockaie, A. et al., 2013. Impacts of correlations on reliable shortest path finding: a simulation-based study. *Transportation Research Record: Journal of the Transportation Research Board*, (2334), pp.1–9.
- Zockaie, A. et al., 2018. Reliability-based user equilibrium formulation in dynamic stochastic networks with correlated travel times and heterogeneous user preferences. In *Transportation Research Board 97th Annual Meeting 2018*, (No. 18-06741).
- Zockaie, A., Mahmassani, H.S. & Nie, Y., 2016. Path Finding in Stochastic Time Varying Networks with Spatial and Temporal Correlations for Heterogeneous Travelers. *Transportation Research Record: Journal of the Transportation Research Board*, (2567), pp.105–113.
- Zockaie, A., Nie, Y. & Mahmassani, H., 2014. Simulation-based method for finding minimum travel time budget paths in stochastic networks with correlated link times. *Transportation Research Record: Journal of the Transportation Research Board*, (2467), pp.140–148.

**Session A3**  
**Network Loading and Simulation I**

# HANDLING NON-UNIQUE FLOWS IN MACROSCOPIC FIRST-ORDER INTERSECTION MODELS - AN EQUILIBRIUM THEORY

Y.-H. CHEN<sup>a</sup> and C. M. J. TAMPÈRE<sup>b</sup>

<sup>a</sup> *Department of Civil and Environmental Engineering, University of Maryland at College Park, College Park, MD 20742, USA*

*Email: [yenchen@terpmail.umd.edu](mailto:yenchen@terpmail.umd.edu)*

<sup>b</sup> *Centre for Industrial Management / Traffic & Infrastructure, Katholieke Universiteit Leuven, Belgium*

*Email: [chris.tampere@kuleuven.be](mailto:chris.tampere@kuleuven.be)*

An important component of macroscopic network traffic models is the node model, representing traffic flow interactions at intersections and bottlenecks. In state-of-the-art node models, it has been shown that a unique solution for the turning flows cannot be guaranteed, which is problematic when applied in deterministic simulations.

In this article, an equilibrium theory is proposed to relax this non-uniqueness problem. Existing approaches use empirical ratios between turning flows involved in the same conflict (which are determined by drivers' compliance with priority rules). Whenever the same flows are involved in multiple conflicts with different ratios, non-unique solutions can occur. The new approach solves this by considering the ratios not rigid but rather as attractors defining a search direction. It is shown that this can solve the non-uniqueness in many (though not all) known, formerly problematic cases.

**Keywords:** First-Order Macroscopic Node Model, Internal Supply Constraint, Holding-Free Solution (HFS)

## 1. INTRODUCTION

In recent years, theory and models of nodes (intersections, bottlenecks) for macroscopic traffic simulation have been developed extensively, from merge/diverge (Daganzo, 1995; Jin and Zhang, 2003; and Ni and Leonard, 2005) to general applicability (Corthout *et al.*, 2012; Flötteröd and Rohde, 2011; Gibb, 2011; Jabari, 2016; Lebacque and Khoshyaran, 2005; Smits *et al.*, 2015; and Tampère *et al.*, 2011). Under reasonable assumptions of how different turning flows at intersections compete for limited supply (in outgoing links or internal conflict points), a unique solution for the turning flows cannot be guaranteed. This non-uniqueness issue is not desired for a deterministic model. To deal with this issue, pre- or post-processing approaches have been suggested (Corthout *et al.*, 2012), which are however rather pragmatic, lacking theoretical and behavioural ground. In this paper, an equilibrium theory is introduced to tackle this non-uniqueness problem.

## 2. BACKGROUND

### 2.1 The Feasible Region

Define  $I$  and  $O$  as set of incoming and outgoing flow, respectively. Algorithms have been developed to find the solution point within the feasible region, which is defined by constraints of non-negativity, demand, conservation of turning fraction (CTF), and external and internal supply (Tampère *et al.*, 2011).

$$q_i \geq 0 \quad \forall i \in I \quad (1)$$

$$q_i \leq \delta_i \quad \forall i \in I \quad (2)$$

$$\sum_i f_{i,o} \cdot q_i \leq \sigma_o \quad \forall o \in O \quad (3)$$

$$f_{i,o} = (q_{i,o} / q_i) = (\delta_{i,o} / \delta_i) \quad \forall i \in I \quad (4)$$

$$q_{i,o} = f_{i,o} q_i \leq F_{i,o} \left( \sum_{(i',o') \forall P_{i',o'} \succ_p P_{i,o}} w_{(i,o),(i',o')} q_{i',o'} \right) \quad \forall (i,o) \in L \quad (5)$$

Constraint (1) represents non-negativity: incoming flow of link  $i$  cannot be negative by its physical

nature. Furthermore, it cannot be larger than the demand  $\delta_i$  of the incoming link (constraint (2)). The (external) supply constraints (3) state that the outgoing flow  $q_o = \sum_i f_{i,o} \cdot q_i$  (where  $f_{i,o}$  is the turning fraction from link  $i$  to  $o$ , representing road users' route choice) should never exceed the supply constraint  $\sigma_o$  of outgoing link  $o$ . Constraint (4) states that road users should obey first-in-first-out (FIFO) otherwise they violate realistic physical properties on the roads. FIFO can be shown to equivalent to conservation of turning fraction, or CTF (Tampere *et al.*, 2011).

Considering only external supply constraint is not enough to describe interactions within the node. Traffic streams in a node influence each other. There is additional capacity lost when two movements interact at conflict points that can only be occupied by one flow at a time. For instance, left turn vehicles yield to the oncoming traffic stream, and when this is the case, the left turn flow is not only determined by the capacity of the target link and demand/capacity of the subject link but also the magnitude of the oncoming traffic. Drawing the relationship between the left turn flow and the oncoming flow, a decreasing and convex curve can be derived. This curve represents the *internal supply constraint*, also called *potential capacity* in HCM 2010 (Transportation Research Board, 2010). If defining  $F_{i,o}(\cdot)$  as the potential capacity function of stream  $(i,o)$ , which is convex and decreasing then the following equation holds:

$$q_{i,o} \leq F_{i,o}(\dots q_{i',o'} \dots) \quad \forall (i,o) \in L \quad (6)$$

where stream  $(i',o')$  has the priority over the minor stream  $(i,o)$  and  $L$  is the set of streams having priority lower than others and constrained by *internal supply constraint*. If we further define the relation as  $P_{i',o'} \succ_p P_{i,o}$ , where  $P_{i,o}$  denotes the priority of the stream and  $\succ_p$  denotes *higher priority than*, then the general form is (5), where  $w_{(i,o),(i',o')}$  is the weighting parameter in HCM.

## 2.2 The Objective Function

To derive the turning flows  $q_{i,o}$  considering all complex interactions between vehicles in incoming links and outgoing links, there are various methods. In other words, various objective functions appear in literature. In Smits *et al.* (2015), three models in the family reach Pareto optimal; in Flötteröd and Rohde (2011) and Tampere *et al.* (2011), algorithms stop at certain criteria, which are fulfilled if all incoming links are constrained by at least one of the constraints on itself or its turning flows. In this paper, holding-free solution (HFS) in Jabari (2016) is selected as the general objective function since it is the most general form covering most of the cases in the literature. The algorithm in Flötteröd and Rohde (2011) terminates at HFS with proof given in Jabari (2016). In addition, the algorithm of Tampere *et al.* (2011) is equivalent to Flötteröd and Rohde (Smits *et al.*, 2015), thus it also terminates at HFS. Three family members in Smits *et al.* (2015) also terminate at HFS. Given its generality, let us set the HFS as the objective function. Original HFS in Jabari (2016) did not consider the case of *internal supply constraint*, so let us show the modified version of HFS as:

$$(\delta_i - q_i) \cdot \prod_{o, f(i,o) > 0} (\sigma_o - q_o) \cdot G_i = 0 \quad \forall i \in I \quad (7)$$

where  $G_i$  is the modification considering *internal supply constraint* additionally,

$$G_i = \prod_{\forall (i',o'), P_{i,o} <_p P_{i',o'}} (F_{i,o}(\cdot) - q_{i,o}) \cdot \prod_{\forall (i'',o''), P_{i,o} >_p P_{i'',o''}} (F_{i'',o''}(\dots q_{i,o} \dots) - q_{i'',o''}) \quad \forall i \in I \quad (8)$$

This constraint guarantees that each incoming link  $i$  is constrained by at least one of the (demand, internal or external supply) constraints on itself or its turning flows. If not, one of the flows would be holding traffic for no reason, which conflicts with the entropy maximizing principle of Ansorge (1990).

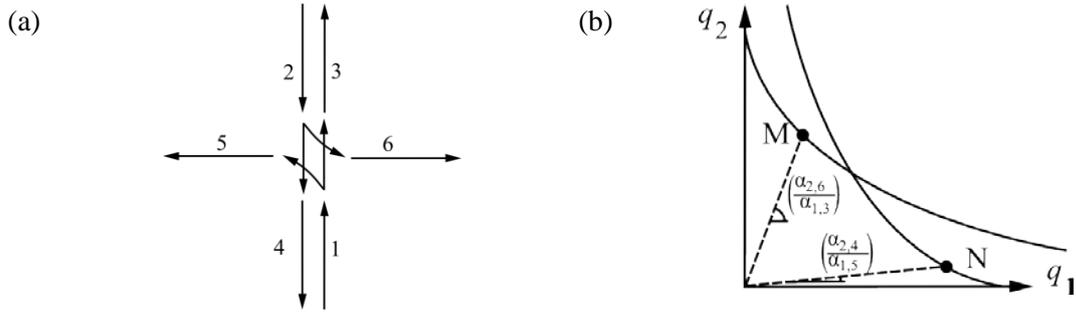


Figure 1. The two-dimensional Example (a) Movements (b) On the space formed by incoming flows



Figure 2. (a) Intersection of the ray and internal supply constraint (b) Equilibrium point of view

## 2.3 The Model

The complete general problem P1 is: find (7) subject to (1)-(5). Reaching HFS means reaching a Pareto optimum. The Pareto optimum is a frontier with (possibly) many solutions: most constraint sets define a feasible region containing only a single point; however, Corthout *et al.* (2012) identified conditions where the feasible region contains multiple solutions. Without additional specification, it can then not be identified which solution is the more appropriate one.

## 2.4 Non-Uniqueness Case

One of the cases in Corthout *et al.* (2012) is depicted in figure 1(a). Left turners yield to the oncoming traffic. If – as is commonly specified – turning flows need to be proportional to the priority parameters  $\alpha_{2,6}/\alpha_{1,3}$  and  $\alpha_{1,5}/\alpha_{2,4}$  for the minor flow over the major flow, respectively, then the solution is ambiguous. Observe figure 1(b), the ray representing priority ratio crossing the *internal supply constraint* is at M or N. Both M or N satisfy problem P1.

## 3. EQUILIBRIUM POINT OF VIEW

### 3.1 Role of Priority Constants: defining the solution or defining an attractor?

As illustrated in Figure 1, techniques in literature for finding the solution on the Pareto frontier is to introduce priority constants (Daganzo, 1995). As shown in figure 2(a), priority constants define a ray in flow space slope with slope  $\alpha_{\text{minor}}/\alpha_{\text{major}}$ , which intersects with the curve of potential capacity at  $\mathbf{q}^* \equiv [q_{\text{Major}}^*, q_{\text{minor}}^*]^T$ . The position of  $\mathbf{q}^*$  (i.e. ratio  $\alpha_{\text{minor}}/\alpha_{\text{major}}$ ) depends on empirical evidence. We propose a formulation compatible to this concept, yet yielding different results for cases like Figure 1. The key is to consider the point  $\mathbf{q}^*$  defined by the empirical priority ratios as an attractor, rather than a ray defining the solution. Once limited by the *internal supply constraint*, the flow  $\mathbf{q}$  would tend to reach an equilibrium point  $\mathbf{q}^* = \mathbf{q}^E$ . Non-equilibrium flow ratios tend to induce behavior trying to restore equilibrium; minor flows being suppressed too strongly get impatient and force themselves onto the conflicts (whether allowed by traffic regulations or not), and/or major flow drivers feel pity and let minor vehicles in. That is: unless pull towards other attractors prevent us from reaching the equilibrium.

Applied to cases with unique solutions like Figure 2, the equilibrium point  $\mathbf{q}^*$  generates the same

solution as the conventional point of view. No matter beginning at any arbitrary point  $\bar{q}$  and  $\hat{q}$  on the *internal supply constraint*, by following the chevrons one converges to the equilibrium (priority rate  $\alpha_{\text{minor}} / \alpha_{\text{Major}}$ ). The theory however also solves former non-unique cases.

### 3.2 Re-Examining the 2-D Case

Let us apply the equilibrium theory to the two-dimensional pathological case of Figure 1. In Figure 3(a), point E is the equilibrium attractor when there only the single priority relation between the subject flow  $q_{i,o}$  and the priority flow  $q_{i',o'}$  would exist. However, if another priority relation, denoted by  $F_{i'',o''}(\dots q_{i,o''})$ , exists, or, at least one of the movements from  $i''$  has the priority over arbitrary movements from  $i'$ , then attractor E falls into the non-feasible region. Ray  $\alpha_{i,o} / \alpha_{i',o'}$  is meaningful only when E is feasible, or when it intersects with  $F_{i,o}(\dots q_{i',o'})$ . Thus, attractor E cannot be reached. In addition, N does not have physical meaning either, as  $\alpha_{i,o} / \alpha_{i',o'}$  has no physical meaning when intersecting a curve  $F_{i',o'}(\dots q_{i,o''})$  belonging to another conflict point.

### 3.3 Achieving the Unique Solution

In figure 3(b), if we would aim for  $E_1$  on constraint  $F_1$ , we find flows to be constrained stronger at M by constraint  $F_2(\cdot)$ , which is attracted towards  $E_2$ , and the search direction is shown by chevrons. When reaching S, no other feasible sets exist by searching along the chevron(s). The same reasoning applies to N. No matter starting from M or N or other frontier on the feasible region, S is the termination point. Point S has the following properties: on one hand, it is in feasible region, which is the shaded area in figure 3(b); on other hand, it satisfies HFS, or (7). A side note of solution point S is: it shows that the reasonable point showing symmetry should be the solution, instead of asymmetric M or N which would induce equilibrium-seeking behavior.

## 4. SUMMARY AND FUTURE WORKS

In this paper, we generalized HFS and re-interpreted previously proposed priority ratios at internal (or external) conflicts as attractors towards an empirical equilibrium behavior. This allowed us to find a unique solution in pathological cases that formerly had multiple solutions on a Pareto front.

Our equilibrium theory can solve some conditions with multiple solutions but not all. General conditions remain to be investigated. Other open issues include capability of the theory to obtain a unique solution for higher-dimensional cases (such as yield-to-the-right crossings), and the development of algorithmic solutions. We leave these issues to be discussed in the full paper and in future works.

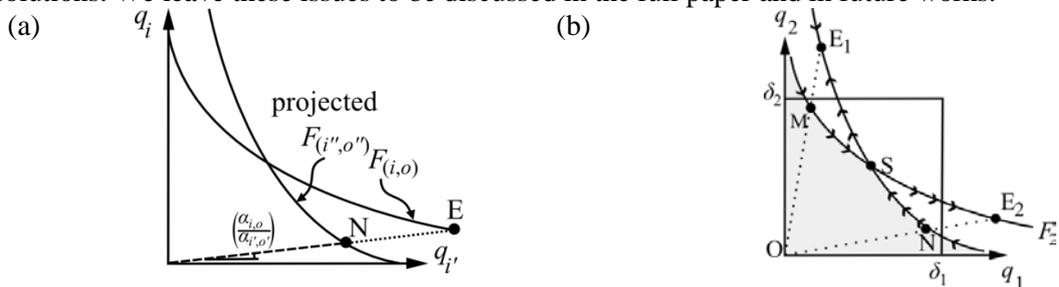


Figure 3 (a) Point E not feasible and N without physical meaning (b) Solution S

## REFERENCES

- Ansorge, R., (1990). What does the entropy condition mean in traffic flow theory? *Transportation Research Part B* 24 (2), 133–143.
- Corthout, R., Flötteröd, G., Viti, F., Tampère, C.M. (2012) Non-unique flows in macroscopic first-order intersection models. *Transportation Research Part B: Methodological* 46 (3), 343–359.

- Daganzo, C. F., 1995. The cell transmission model, part II: network traffic. *Tr. Res. B* 29(2), pp. 79-93.
- Flötteröd, G., Rohde, J. (2011) Operational macroscopic modeling of complex urban road intersections. *Transportation Research Part B* 45, pp. 903-922.
- Gentile, G., Meschini, L., & Papola, N. (2007) Spillback congestion in dynamic traffic assignment: a macroscopic flow model with time-varying bottlenecks. *Transportation Research Part B*, 41(10), pp. 1114-1138.
- Gibb, J. (2011) Model of traffic flow capacity constraint through nodes for dynamic network loading with queue spillback. *Transportation Research Record* 2263, pp. 113–122.
- Jabari, S.E. (2016) Node modeling for congested urban road networks. *Tr. Res. B* 91, pp. 229–249.
- Jin, W. L., and Zhang, H. M. (2003) On the distribution schemes for determining flows through a merge. *Transportation Research Part B*, 37(6), pp. 521-540.
- Lebacque, J., Khoshyaran, M. (2005) First order macroscopic traffic flow models: intersections modeling, network modeling. *In: Proceedings of the 16th ISTTT*, pp. 365-386.
- Ni, D., and Leonard, J. D. (2005) A simplified kinematic wave model at a merge bottleneck. *Applied Mathematical Modelling*, 29(11), pp. 1054-1072.
- Smits, E.S., Bliemer, M. C.J, Pel, A. J., van Arem, B. (2015) A family of macroscopic node models. *Transportation Research Part B* 74 (1), pp. 20–39
- Tampère, C.M., Corthout, R., Cattrysse, D., Immers, L.H. (2011) A generic class of first order node models for dynamic macroscopic simulation of traffic flows. *Transportation Research Part B* 45 (1), pp. 289–309.
- Transportation Research Board, National Research Council (2010) *Highway Capacity Manual 2010*. Washington DC.

# RIEMANN SOLVERS AT A JUNCTION FOR TRAFFIC FLOW ON A ROAD NETWORK BY USING A CONSERVED HIGHER-ORDER MODEL

D. L. QIAO<sup>a</sup>, P. ZHANG<sup>a</sup>, K. ZHANG<sup>b</sup> and H. LIU<sup>c</sup>

<sup>a</sup> *Shanghai Institute of Applied Mathematics and Mechanics, Shanghai University, Shanghai, China*

*Email: [gdl821115@163.com](mailto:gdl821115@163.com); [pzhang@shu.edu.cn](mailto:pzhang@shu.edu.cn)*

<sup>b</sup> *Beijing Municipal Transport Operation Coordinate Center, Beijing, China*

*Email: [13911309925@139.com](mailto:13911309925@139.com)*

<sup>c</sup> *Beijing Transportation Information Center, Beijing, China*

*Email: [liuhao@bjjtw.gov.cn](mailto:liuhao@bjjtw.gov.cn)*

Traffic flow on networks is very important for the research of urban traffic. More and more network traffic flow models were proposed, such as the LWR network model (Holden and Risebro, 1995), the conserved higher-order model (CHO) network model (Lin et al., 2015), and so on. The Riemann solver and the corresponding Godunov flux for the LWR model at a junction are well-known (Lebacque, 1996; Coclite et al., 2005), but the LWR network model cannot simulate complex traffic phenomena (such as stop-and-go waves). However, the higher-order network model can overcome this shortcoming. Especially, the CHO model is easier to be extended to a traffic network because of its advantages, compared to other higher-order traffic flow models. In Lin et al. (2015), the Riemann solvers for the CHO model at a junction were exactly given by solving the optimization problem which the total pseudo-outflow on all upstream roads is maximized. And stop-and-go waves were repeated by the CHO network model using the corresponding Godunov scheme.

Based on the research work in Lin et al. (2015), the present paper intends to obtain extended Riemann solvers for the CHO model at a junction by making the linear combination of pseudo-outflow be maximized. In fact, the linear combination of pseudo-outflow is the total outflow on all upstream roads, which indicates that extended Riemann solvers are more reasonable.

**Keywords:** Riemann Problem, Optimization Problem, Diverging Junction, Merging Junction

## 1. INTRODUCTION

Traffic flow on networks is very important for the research of urban traffic. More and more network traffic flow models were proposed, such as the LWR network model (Holden and Risebro, 1995), the conserved higher-order model (CHO) network model (Lin et al., 2015), and so on. The Riemann solver and the corresponding Godunov flux for the LWR model at a junction are well-known (Lebacque, 1996; Coclite et al., 2005), but the LWR network model cannot simulate complex traffic phenomena (such as stop-and-go waves). However, the higher-order network model can overcome this shortcoming. Especially, the CHO model is easier to be extended to a traffic network because of its advantages, compared to other higher-order traffic flow models. In Lin et al. (2015), the Riemann solvers for the CHO model at a junction were exactly given by solving the optimization problem which the total pseudo-outflow on all upstream roads is maximized. And stop-and-go waves were repeated by the CHO network model using the corresponding Godunov scheme.

Based on the research work in Lin et al. (2015), the present paper intends to obtain extended Riemann solvers for the CHO model at a junction by making the linear combination of pseudo-outflow be maximized. In fact, the linear combination of pseudo-outflow is the total outflow on all upstream roads, which indicates that extended Riemann solvers are more reasonable.

## 2. MODEL AND ITS' RIEMANN SOLVERS

### 2.1 Model Equations

Let  $J_{n \times m}$  denote a junction which connects  $n$  upstream roads and  $m$  downstream roads. On each road  $k$  ( $k = 1, \dots, n + m$ ), the CHO model is written as follows:

$$(\rho_k)_t + (\rho_k V(w_k))_x = 0, \quad (1)$$

$$(w_k)_t + (w_k V(w_k))_x = \frac{V(w_k) - v_e(\rho_k)}{-\tau V'(w_k)}, \quad (2)$$

where  $\rho_k = \rho_k(x, t)$  and  $w_k = w_k(x, t)$  are the density and pseudo-density on road  $k$  in location  $x$  at time  $t$ . The system of (1) and (2) is called the CHO model at a junction.

## 2.2 Riemann Solvers

The homogeneous system of (1) and (2) is written as follows:

$$(\rho_k)_t + [f_k(\rho_k, w_k)]_x = 0, \quad (3)$$

$$(w_k)_t + [g_k(w_k)]_x = 0, \quad (4)$$

where  $f_k(\rho_k, w_k) = \rho_k V(w_k)$  and  $g_k(w_k) = w_k V(w_k)$  are the flow and pseudo-flow on road  $k$ , respectively, and  $f_k(\rho_k, w_k) = z_k^{-1} g_k(w_k)$ ,  $z_k = w_k / \rho_k$ . If the initial condition of the above homogeneous system is

$$\rho_k(x, 0) = \rho_{k0}, \quad w_k(x, 0) = w_{k0}, \quad k = 1, \dots, n + m, \quad (5)$$

where  $\rho_{k0}$  and  $w_{k0}$  are constants. Then, the system of (3)-(5) is called the Riemann problem at a junction.

We will seek for the Riemann solvers which are denoted by  $(\hat{\rho}_k, \hat{w}_k)$  at the junction for all roads. The flows at the junction are denoted by

$$\hat{f}_k = f_k(\hat{\rho}_k, \hat{w}_k), \quad \hat{g}_k = g_k(\hat{w}_k), \quad k = 1, \dots, n + m, \quad (6)$$

and  $\hat{f}_k = \hat{z}_k^{-1} \hat{g}_k$ ,  $\hat{z}_k = \hat{w}_k / \hat{\rho}_k (\geq 1)$ . With respect to the Riemann invariant  $z_i$  on upstream road  $i$  (see Lin et al. (2015) for detailed discussion), we have

$$\hat{z}_i = \frac{\hat{w}_i}{\hat{\rho}_i} = \frac{w_{i0}}{\rho_{i0}} = z_{i0}, \quad \text{or } \hat{f}_i = \frac{\rho_{i0}}{w_{i0}} \hat{g}_i, \quad i = 1, \dots, n. \quad (7)$$

According to the user equilibrium conditions at the instant, it is assumed that the inflow on road  $j$  coming from the outflow  $\hat{f}_i$  equals  $\alpha_{ij} \hat{f}_i$ , then by the flow conservation at the junction, we have

$$\hat{f}_j = \sum_{i=1}^n \alpha_{ij} \hat{f}_i, \quad j = n + 1, \dots, n + m; \quad \sum_{j=n+1}^{n+m} \alpha_{ij} = 1, \quad 0 \leq \alpha_{ij} \leq 1, \quad i = 1, \dots, n. \quad (8)$$

Similarly, we have

$$\hat{g}_j = \sum_{i=1}^n \beta_{ij} \hat{g}_i, \quad j = n + 1, \dots, n + m; \quad \sum_{j=n+1}^{n+m} \beta_{ij} = 1, \quad 0 \leq \beta_{ij} \leq 1, \quad i = 1, \dots, n. \quad (9)$$

Here,  $\alpha_{ij} = \beta_{ij}$  ( $i = 1, \dots, n, j = n + 1, \dots, n + m$ ), which ensure that the Riemann solvers are physically bounded. The pseudo-flow conservation at the junction gives

$$\sum_{i=1}^n \hat{g}_i = \sum_{j=1}^{n+m} \hat{g}_j. \quad (10)$$

Note that equation (4) is a scalar equation of  $w_k$ , there are two ways to obtain the Riemann solvers of equation (4) at the junction. The first way is that the total pseudo-outflow  $\sum_{i=1}^n \hat{g}_i$  is maximized, see the references (Coclite et al., 2005, Lin et al., 2015). The second way is that the total outflow  $\sum_{i=1}^n \hat{f}_i$  is maximized, or equivalently the linear combination of the pseudo-outflow  $\sum_{i=1}^n z_{i0}^{-1} \hat{g}_i$  is maximized. It is equivalent to solve optimization problem

$$\begin{aligned} & \max \left\{ \sum_{i=1}^n z_{i0}^{-1} \hat{g}_i \right\} \quad s.t. \\ & \begin{cases} 0 \leq \hat{g}_i \leq d_i, \quad i = 1, \dots, n, \\ 0 \leq \hat{g}_j = \sum_{i=1}^n \beta_{ij} \hat{g}_i \leq s_j, \quad j = n+1, \dots, n+m, \\ \sum_{i=1}^n \hat{g}_i = \sum_{j=1}^{n+m} \hat{g}_j \end{cases} \end{aligned} \quad (11)$$

where the demand function  $d_i = d_i(w_{i0})$  and supply function  $s_j = s_j(w_{j0})$  are defined as (Lebacque, 1996)

$$d_i(w_{i0}) = \begin{cases} g_i(w_{i0}), & w_{i0} \in [0, w^*] \\ g_i(w^*), & w_{i0} \in [w^*, w_m] \end{cases}, \quad s_j(w_{j0}) = \begin{cases} g_j(w^*), & w_{j0} \in [0, w^*] \\ g_j(w_{i0}), & w_{j0} \in [w^*, w_m] \end{cases}, \quad (12)$$

and  $g'_k(w^*) = 0$ ,  $w_m = \rho_{jam}$ , here  $\rho_{jam}$  is the jam density. According to the above discussions, the steps for Riemann solvers of system (3) and (4) are stated as follows:

1. Obtain  $\{\hat{g}_i\}_{k=1}^{n+m}$  by solving the optimization problem (11), then,  $\{\hat{w}_i\}_{k=1}^{n+m}$  are obtained by equation (6),  $\{\hat{f}_i\}_{i=1}^n$  and  $\{\hat{\rho}_i\}_{i=1}^n$  are obtained by equation (7);
2. Obtain  $\{\hat{f}_j\}_{j=n+1}^{n+m}$  by equation (8), and  $\{\hat{\rho}_j\}_{j=n+1}^{n+m}$  are obtained by equation (6),

$$\hat{\rho}_j = \frac{\sum_{i=1}^n \alpha_{ij} \hat{\rho}_i V(\hat{w}_i)}{V(\hat{w}_j)}, \quad j = n+1, \dots, n+m. \quad (13)$$

For convenience, we only give the solutions of the optimization problem (11) in two simple junctions.

1. For a diverging junction with  $n = 1$  and  $m = 2$  ( $\beta_{12} = 1 - \beta_{13}$ ), the solution of the optimization problem (11) is

$$\hat{g}_1 = \min\{d_1, s_2 / \beta_{12}, s_3 / \beta_{13}\}, \quad \hat{g}_2 = \beta_{12} \hat{g}_1, \quad \hat{g}_3 = \beta_{13} \hat{g}_1. \quad (14)$$

2. For a merging junction with  $n = 2$  and  $m = 1$  ( $\beta_{13} = \beta_{23} = 1$ ), the solution of the optimization problem (11) is divided into two cases:

- (i) If  $z_{10}^{-1} = z_{20}^{-1}$ ,  $\max \left\{ \sum_{i=1}^n z_{i0}^{-1} \hat{g}_i \right\}$  is equivalent to  $\max \left\{ \sum_{i=1}^n \hat{g}_i \right\}$ , the solution is

$$\hat{g}_1 = q \hat{g}_3, \quad \hat{g}_2 = (1-q) \hat{g}_3, \quad \hat{g}_3 = \min\{d_1 + d_2, s_3\}, \quad (15)$$

$$q = \begin{cases} d_1 / (d_1 + d_2), & d_1 + d_2 \leq s_3, \\ \gamma, & d_1 + d_2 > s_3, \quad \gamma s_3 \leq d_1, \quad (1-\gamma)s_3 \leq d_2, \\ d_1 / s_3, & d_1 + d_2 > s_3, \quad \gamma s_3 > d_1, \\ (s_3 - d_2) / s_3, & d_1 + d_2 > s_3, \quad (1-\gamma)s_3 > d_2. \end{cases} \quad (16)$$

where  $\gamma$  and  $1-\gamma$  are the weights of priority of cars on road 1 and 2 driving into road 3 respectively.

- (ii) If  $z_{10}^{-1} \neq z_{20}^{-1}$ , the solution is the same as form (15), but

$$q = \begin{cases} d_1 / (d_1 + d_2), & d_1 + d_2 \leq s_3, \\ (s_3 - d_2) / s_3, & d_1 + d_2 > s_3, \quad z_{10}^{-1} < z_{20}^{-1}, \quad s_3 \geq d_2, \\ 0, & d_1 + d_2 > s_3, \quad z_{10}^{-1} < z_{20}^{-1}, \quad s_3 < d_2, \\ d_1 / s_3, & d_1 + d_2 > s_3, \quad z_{10}^{-1} > z_{20}^{-1}, \quad s_3 \geq d_1, \\ 1, & d_1 + d_2 > s_3, \quad z_{10}^{-1} > z_{20}^{-1}, \quad s_3 < d_1. \end{cases} \quad (17)$$

### 3. NUMERICAL SIMULATION

Considering a simple small network which is consisted of two junctions and four roads, which the length of each road is  $L = 5000m$ , see Fig. 1.

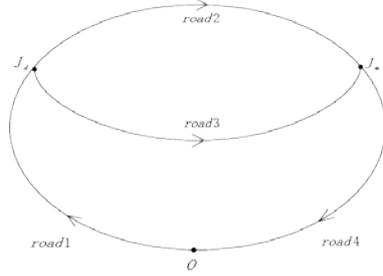


Figure 1. A simple small network

In system of (1) and (2), the velocity-density relationships are set as (Zhang et al., 2011)

$$V(\rho_k) = v_f \frac{1 - \rho_k / \rho_{jam}}{1 + b(\rho_k / \rho_{jam}) + a(\rho_k / \rho_{jam})^2}, k = 1, 2, 3, 4, \quad (18)$$

$$v_e(\rho_k) = v_f ((1 + \exp((\rho_k / \rho_m - 0.25) / 0.06))^{-1} - 3.72 \times 10^{-6}), k = 1, 2, 3, 4, \quad (19)$$

The initial conditions are set as

$$\begin{cases} \rho_1(x, 0) = \rho_{10} \rho_{jam} + 0.1 \rho_{jam} \{ \cosh^{-2}[160 / L(x - 3L/8)] - 1/4 \cosh^{-2}[40 / L(x - 13L/32)] \} \\ \rho_k(x, 0) = \rho_{k0} \rho_{jam}, k = 2, 3, 4, \\ w_k(x, 0) = V^{-1}(v_e(\rho_k(x, 0))), k = 1, 2, 3, 4, \end{cases} \quad (20)$$

where  $\rho_{10} = 0.22$ ,  $\rho_{20} = 0.20$ ,  $\rho_{30} = 0.24$  and  $\rho_{40} = 0.22$ . The periodic boundary conditions are applied at node  $O$ . The parameters in formulas (14) and (16) are chosen as  $\beta_{12} = \beta_{13} = 0.5$  and  $\gamma = 0.5$ . Corresponding parameters in formulas (18)-(20) are set as  $a = 4.0$ ,  $b = -0.8$ ,  $v_f = 25m/s$  and  $\rho_{jam} = 0.16l_k \text{ veh}/m$ , where  $l_k$  is the number of lanes on road  $k$ , and  $l_1 = l_4 = 2$ ,  $l_2 = l_3 = 1$ .

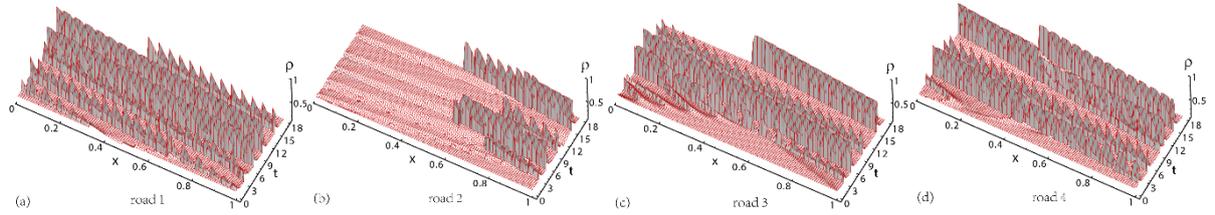


Figure 2. Evolution of densities on all roads corresponding to  $\max \left\{ \sum_{i=1}^n \hat{g}_i \right\}$

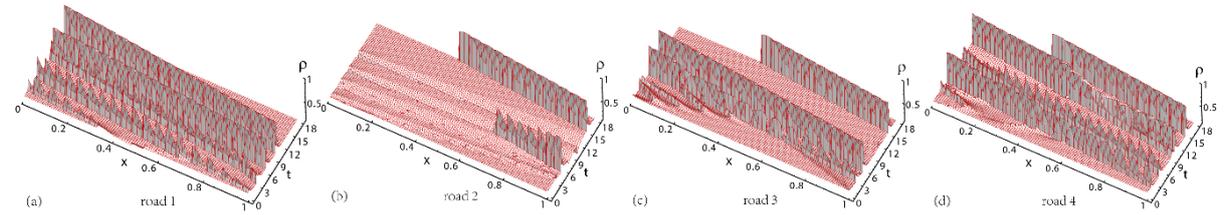


Figure 3. Evolution of densities on all roads corresponding to  $\max \left\{ \sum_{i=1}^n z_{i0}^{-1} \hat{g}_i \right\}$

The simulation results corresponding to  $\max \left\{ \sum_{i=1}^n \hat{g}_i \right\}$  and  $\max \left\{ \sum_{i=1}^n z_{i0}^{-1} \hat{g}_i \right\}$  are shown in Fig. 2 and 3, respectively. The densities are scaled by  $\rho_{jam}$  on all roads, and the simulation time  $t$  ( $0s \leq t \leq 3600s$ ) is scaled by 200s. Stop and go waves are observed in these phases, but the results are different corresponding to the same road in Fig. 2 and 3, that is because the  $\rho_{20}$  is not equal to

$\rho_{30}$ , which makes the solutions be different.

## ACKNOWLEDGMENTS

This study was jointly supported by grants from the National Natural Science Foundation of China (11672348, 11272199), the National Basic Research Program of China (2012CB725404).

## REFERENCES

- Holden, H. and Risebro, N.H. (1995) A mathematical model of traffic flow on a network of roads. *SIAM J. Math. Anal.*, 26 (4), pp. 999-1017.
- Lin, Z.Y., Zhang P., Dong L.Y., (2015) Traffic flow on a road network using a conserved higher-order model. *AIP Publishing LLC*, pp. 1862–1886.
- Coclite, G.M., Garavello, M. and Piccoli, B. (2005) Traffic flow on a road network. *SIAM J. Math. Anal.*, 36, pp. 1862–1886.
- Lebacque, J.P. (1996) The Godunov scheme and what it means for first order traffic flow models, in *Proceedings of the Thirteenth International Symposium on Transportation and Traffic Theory*, edited by J. B. Lesort, Lyon, France, pp. 647–677.
- Zhang, P., Qiao, D.L., Dong, L.Y., et al. (2011) A number of Riemann solvers for a conserved higher order traffic flow model, in *Fourth International Joint Conference on Computational Sciences and Optimization (CSO)*, IEEE, Lijiang, pp. 1049–1053.

# UNDERSTANDING THE TRADE OFFS BETWEEN DTA MODELS REALISM AND ROBUSTNESS: THE IMPACT OF SPILLBACK MODELING

Stephen D. BOYLES<sup>a</sup> and Natalia Ruiz JURI<sup>b</sup>

<sup>a</sup> *Assistant Professor, Department of Civil, Architectural and Environmental Engineering, The University of Texas at Austin*

*Email: [sboyles@austin.utexas.edu](mailto:sboyles@austin.utexas.edu)*

<sup>b</sup> *Research Associate, Center for Transportation Research, The University of Texas at Austin*

*Email: [nruizjuri@mail.utexas.edu](mailto:nruizjuri@mail.utexas.edu)*

This paper examines the trade offs between model realism and robustness to input errors, in the context of dynamic network loading for dynamic traffic assignment (DTA) implementations. We focus on the model capability to represent congestion spillbacks, which are responsible for a significant portion of the observed congestion in real transportation networks. In practice, input data are never known with exact precision due to measurement, estimation, and forecasting errors. Therefore, it is not obvious that the more "realistic" model is preferable to one which is more robust to these errors. Rather, one or the other model may be preferable depending on (1) the level of uncertainty in the input data, (2) the relative sensitivities of the two models, (3) and the magnitude of the error introduced by ignoring spillback.

In this paper we explore the impacts of the error introduced by ignoring queue spillback to the error introduced from incorrect input data when spillbacks are explicitly modeled. Our approach includes conceptual tests in a simple setting, and numerical examples in large real-world networks.

**Keywords:** Intersections, Spillback, Point Queue, Uncertainty, Dynamic Network Loading

## 1. INTRODUCTION

Bringing advanced models to practice often involves moving from simpler, stylized models to more realistic ones by relaxing assumptions. While more realistic models are *ceteris paribus* preferable to less realistic ones, realism is rarely the only criterion for model selection. Realism must be balanced against tractability, availability of data for calibration and validation, mathematical regularity, and robustness to errors in the input data. Practitioners must make modeling choices taking all these factors into account, given their relevance in a specific application.

This paper specifically examines the trade offs between model realism and robustness to input errors, in the context of dynamic network loading for dynamic traffic assignment (DTA) implementations. We focus on the model capability to represent congestion spillbacks, which are responsible for a significant portion of the observed congestion in real transportation networks.

While capturing spillbacks is clearly a critical feature of a practical DTA model, spillbacks may lead to excessive congestion and gridlocks, which typically result in poorer convergence and less stable models. The latter is often observed in future year models for which inputs are inaccurate, and traffic control at intersections can be only coarsely approximated. Further, in practice, input data are never known with exact precision due to measurement, estimation, and forecasting errors. Therefore, it is not obvious that the more realistic model is preferable to one which is more robust to these errors. Rather, one or the other model may be preferable depending on (1) the level of uncertainty in the input data, (2) the relative sensitivities of the two models, (3) and the magnitude of the error introduced by ignoring spillback.

In this paper we explore the impacts of the error introduced by ignoring queue spillback (systematic error) to the error introduced from incorrect input data (random error) when spillbacks are explicitly modeled. Our approach includes conceptual tests in a simple setting, and numerical examples in large real-world networks.

Preliminary findings in a simple freeway interchange suggest that in cases of low data uncertainty the model with spillback is clearly preferred; in the case of high data uncertainty, the no-spillback model

often has the lower expected absolute error. The full paper will include similar analyses on larger-scale urban networks, and we will explore practical solutions to address issues related to unrealistic spillback impacts in future year models. The following sections illustrate the concepts previously described through a simple example.

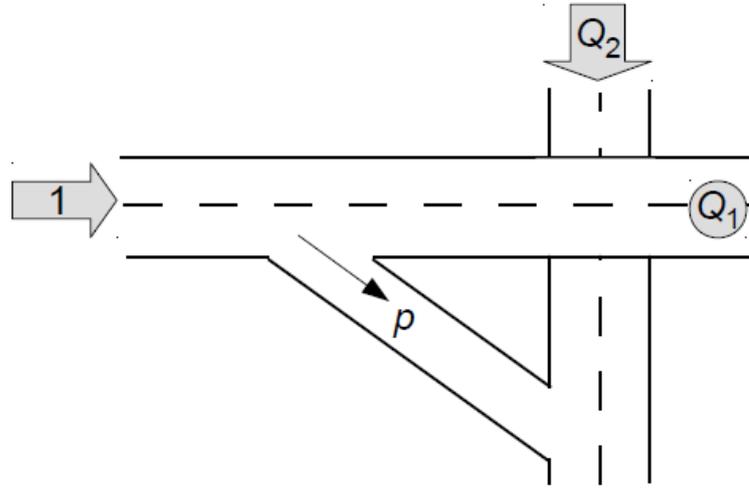


Figure 1. Freeway interchange for demonstration; mainlines have capacity 1 and ramp has capacity 1/2.

## 2. FREEWAY INTERCHANGE SCENARIO DESCRIPTION

Consider the freeway interchange shown in Figure 1. The two mainlines have equal capacity, and the ramp connecting them has half the capacity of the mainlines. For simplicity, assume that units are chosen such that the capacity of each mainline is equal to one, and that the inflow rate on the horizontal mainline is also one. There are two model parameters which must be estimated: the proportion  $p$  of flow on the horizontal link choosing the ramp, and the flow  $Q_2$  on the vertical mainline link. In this demonstration both are assumed stationary in time. We wish to estimate the steady-state flow rate on the horizontal link downstream of the diverge, shown by the circular detector in Figure 1, after any initial transient conditions have subsided.

We use the standard merge and diverge equations in dynamic network loading, assuming that in highly congested conditions the merge allocates flow to approaches proportionate to their capacity, and that diverges respect the first-in, first-out principle. Since the inflow to the ramp is initially  $p$ , if  $p + Q_2 \leq 1$  the merge is uncongested and all flow can move freely. If  $p + Q_2 > 1$ , three cases are possible: if  $p < 1/3$ , a queue forms on the vertical link but not the ramp; if  $Q_2 < 2/3$ , a queue forms on the ramp but not the vertical link; and if  $p \geq 1/3$  and  $Q_2 \leq 2/3$ , queues form on both approaches. In a point queue model, the presence of queues at the merge has no bearing on behavior at the diverge, and the total outflow from the diverge is given by  $\min\{1, 1/(2p)\}$ , with the detector registering

$$Q_1^{NS} = (1 - p) \min\{1, 1/(2p)\}. \quad (21)$$

In the case of spillback, equation (1) only holds if there is no queue on the ramp. Otherwise, at steady-state, the ramp outflow (and thus its inflow) is given by  $1 - Q_2$  if the queue is only on the ramp, and by  $1/3$  if queues exist on both the ramp and vertical link. The corresponding flow at the detector is obtained by multiplying the ramp flow by  $(1-p)/p$ . These results are summarized in Figure 2, and are denoted by the mappings  $Q_1^{NS}(Q_2, p)$  and  $Q_1^S(Q_2, p)$  for the no-spillback and spillback cases, respectively. Full derivations are omitted here for reasons of space.

### 3. PROCEDURE

The possible values of inputs  $Q_2$  and  $p$  lie within the unit square  $[0, 1]^2$ . Within this range, twenty evenly-spaced values of  $Q_2$  and  $p$  were combined to produce four hundred scenarios for analysis. For each of these scenarios, the following procedure was performed.

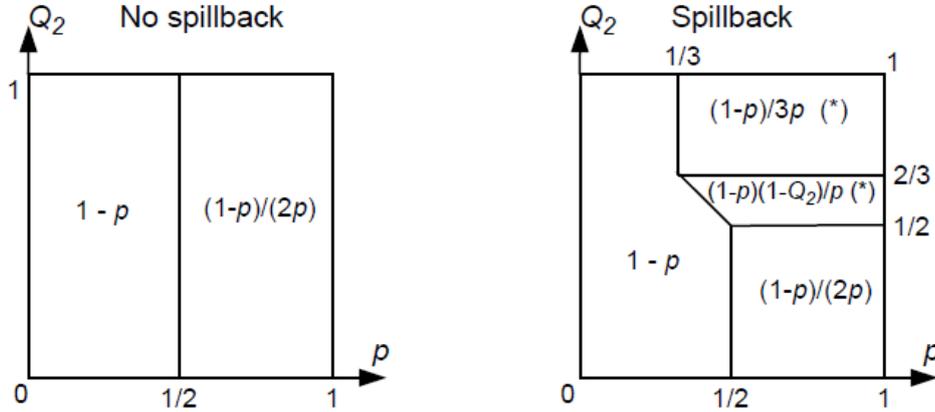


Figure 2. Steady-state  $Q_1$  values for no-spillback and spillback cases. Asterisks indicate regions where a ramp queue restricts inflow.

1. Let  $\hat{Q}_2$  and  $\hat{p}$  denote the values corresponding to this scenario. These are assumed to be the true values of these parameters, corresponding to a true flow rate of  $Q_1^S(\hat{Q}_2, \hat{p})$ . (That is, the spillback model is presumed completely accurate if given the true  $Q_2$  and  $p$  values.)
2. Generate  $n$  sampled values of  $Q_2$  and  $p$ , using independent normal distributions with respective means  $\hat{Q}_2$  and  $\hat{p}$ , and a provided standard deviation.

3. For each sample, the error associated with the no-spillback model is calculated as

$$\epsilon^{NS} = \left| Q_1^{NS}(Q_2, p) - Q_1^S(\hat{Q}_2, \hat{p}) \right|, \quad (22)$$

while the error associated with the spillback model is

$$\epsilon^S = \left| Q_1^S(Q_2, p) - Q_1^S(\hat{Q}_2, \hat{p}) \right|. \quad (23)$$

4. Based on the sampled values of  $\epsilon^{NS}$  and  $\epsilon^S$ , calculate the additional expected error in the no-spillback model to be

$$\delta = E[\epsilon^{NS} - \epsilon^S] \quad (24)$$

along with the standard deviation  $s$  of this difference.

5. Calculate the  $t$  score  $t = \delta / (s / \sqrt{n})$ .

If the resulting  $t$  score is greater than a specified positive critical value, the error in the no-spillback model is significantly greater than that in the spillback model, and the spillback model is to be preferred. If it is smaller than a negative critical value, the error in the no-spillback model is significantly less than that of the spillback model, and the no-spillback model is preferred. Otherwise, there is no significant difference in the errors produced by the models.

#### 4. DISCUSSION AND CONCLUSIONS

Figure 3 presents the results of these simulations for three cases: when the standard deviation of the sampled  $Q_2$  and  $p$  values was small (0.01), moderate (0.1), and large (0.25). A sample size of  $n = 2500$  was used for each scenario, and critical values of 1:96 were used for the statistical test, corresponding to 5% significance. In this figure, an S denotes that the model with spillback produces less expected error, NS denotes that the no-spillback model produces less expected error, and '=' denotes no statistically significant difference in errors. Owing to the sample size, the  $t$  scores were typically quite large, averaging +87:4 across all scenarios where the spillback all model was preferred, and 13:9 across scenarios where the no-spillback model was preferred.

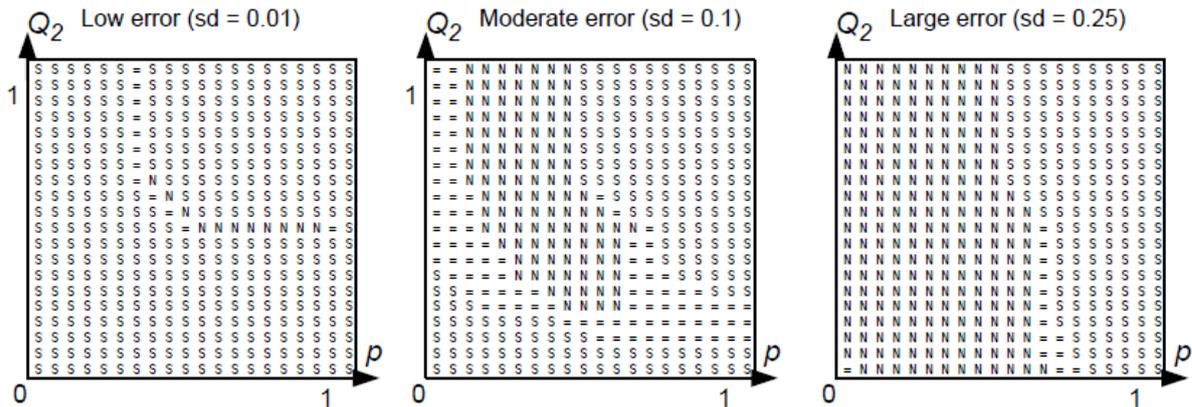


Figure 3. Model with less error for each scenario, if significant ( $p = 0:05$ ).

When input error is small, the model with spillback is almost always preferred, while the no-spillback model often produces less expected error when input errors are large, specifically when the diverge proportion is small. This demonstrates that the no-spillback model may be preferable to the model with spillback under certain circumstances: even though it contains systematic error, it is more robust to random error.

The full paper examines this idea in the context of a dynamic traffic assignment model of a realistic urban area. Specifically, this larger experiment will allow for equilibrium route choice, which may act as a "restoring mechanism" against errors in the input data. The analysis will explore whether, and under what conditions, including spillback can increase error. It will also consider practical solutions for mitigating spillback-induced errors by relaxing the queuing modeling approach at selected intersections.

**Session A4**  
**Dynamic Traffic Control I**

# CONTROLLING PEDESTRIAN FLOWS USING A DYNAMIC TRAFFIC MANAGEMENT SYSTEM

Nicholas MOLYNEAUX<sup>a</sup>, Riccardo SCARINCI<sup>a</sup> and Michel BIERLAIRE<sup>a</sup>

<sup>a</sup> *Transport and Mobility Laboratory,  
School of Architecture, Civil and Environmental Engineering,  
Ecole Polytechnique Fédérale de Lausanne*

*Email: {[nicholas.molyneaux](mailto:nicholas.molyneaux@epfl.ch), [riccardo.scarinci](mailto:riccardo.scarinci@epfl.ch), [michel.bierlaire](mailto:michel.bierlaire@epfl.ch)}@epfl.ch*

Forecasts predict a strong increase of the demand of public transport facilities in the next decades, close to doubling the number of users in the "Arc lémanique" area of Switzerland for example. This increase in demand can lead to congestion and poor level-of-service inside transportation hubs. We propose a control strategy to reduce the congestion levels inside a transportation hub. The objective is to guarantee adequate level-of-service throughout the infrastructure. This is achieved by influencing the flow of pedestrians arriving into the central areas of the hub. These measures are coupled with a dynamic assignment and demand model to simulate the effect on pedestrian flows. To quantify the impact of such management strategies, a simulation framework is developed and the results are compared to real-life data collected in the train station of Lausanne, Switzerland. The gating strategy proposed can efficiently reduce the congestion without excessively increasing the travel times of passengers.

**Keywords:** Pedestrian Traffic Simulation, Management Strategy, Gating

## 1. INTRODUCTION

Multiple forecasts predict a strong increase in demand for public transport (PT) over the next couple of decades. This increase will put pressure on the pedestrian infrastructures associated with public transport: public transport vehicles and platforms, waiting areas, corridors or even the services available to customers (shops, ticket machines). The problems which can arise concern multiple actors. On one hand, users wish to move around freely and safely inside the infrastructure, implying low levels of congestion. On the other hand, the public transport operators try to maximize the number of passengers who use their services while minimizing the operational costs. Although these objectives seem antagonist to each other, there is one common aspect between them: avoiding poor user experience. In order to attract users to their system, PT operators must ensure that the level-of-service experienced by the passengers is acceptable. To achieve this, the operators must either limit the demand levels such that the infrastructure can support it, or increase the capacity of the infrastructure to cope with a larger demand.

This challenge of improving the efficiency of the infrastructure has been vastly studied in vehicular traffic over the past decades. Addressing this challenge gave birth to Dynamic Traffic Management Systems (DTMS) which aim at increasing the capacity of road networks by exploiting the fundamental relationships of traffic. Around the 90s and 00s, quite a few frameworks were designed for evaluating dynamic traffic management systems in the context of vehicular traffic. For example, DYNASMART is an evaluation tool for advanced traffic management systems (ATMS), while DYNAMIT aims at providing prediction-based guidance for route choice. For an in-depth review, see (Mahmassani, 2001) and (Ben-Akiva, et al., 2003). The traffic management systems have various strategies available. In vehicular traffic, some examples are ramp-metering, variable message signs, traffic signals or perimeter control. Management strategies for vehicular traffic have proven useful and they have been extensively studied. On the contrary, management strategies for pedestrian flows are still largely unexplored.

We propose a framework which is capable of evaluating and generating optimal management strategies for pedestrians. The objective is to design a flexible framework similarly to frameworks proposed for vehicular traffic, where the specificities of pedestrian traffic are taken into account. We develop, similarly to perimeter control, a gating strategy which can be used to control pedestrian flows. The calibration of this controller relies on two assumptions: 1) the existence of an aggregate fundamental diagram and 2) a linear relationship between the outflow from an area and the generalized flow within that area. Both of these assumptions have been verified based on empirical data collected in the train station of Lausanne (Switzerland), location chosen for the case study.

## 2. METHODOLOGY

The present section presents the main components of a framework aimed at evaluating and generating management strategies. A given management strategy can be split into various parts, for which we defined the following terminology. The first aspect concerns the action globally and is called the management/control *strategy*, while the second aspect refers to the actual operation of the given action and is called the management/control *policy*. The third is the measures used to apply the control policy and is called the management/control *devices*. As an illustrative example, let's consider gates used to control the flow of pedestrians. The control strategy is gates, with their position and number. The operational aspects, i.e. the control policy specifies the flow rate allowed through the gates. Finally, the control devices are the gates themselves which enforce the control policy.

### 2.1 Simulation laboratory

The framework which is proposed for simulating and generating optimal control strategies is designed with real-time applications as the ultimate goal. This decision implies that the framework should be usable with either input from a real-life environment or a simulated environment. In vehicular traffic, such a simulation of real life is sometimes called a “plant” or a “simulation laboratory” (Ben-Akiva, et al., 2003). We define three key components for such a framework:

#### Pedestrian traffic

The core element of many traffic systems is the interaction between demand and supply. For pedestrian traffic, the supply is the walkable space available to the pedestrians and is called the infrastructure. This infrastructure is *loaded* by the pedestrian demand. This pressure on the supply will in turn influence the motion of pedestrians as they might decide to change route or perform a different set of activities due to congestion. When a specific control strategy needs to be evaluated this component can be reproduced by simulation.

#### Traffic controller

As the management and control of the pedestrian's movements is a central objective, the integration of a controller into the pedestrian traffic is required. The primary objective of this controller is to take decisions regarding the control devices and apply the management strategies. To achieve this, an evaluation of the state of the system is required, based on “real time” data from the pedestrian traffic. Even within a pure simulation environment, data from the pedestrian traffic must be transferred as if the pedestrian simulator was in fact real. The controller “does not need to know” whether the data is coming from a simulator or from a real situation. The traffic controller is divided into two sub-components. First the “state evaluation” which computes, based on the data transferred from the “Pedestrian traffic”

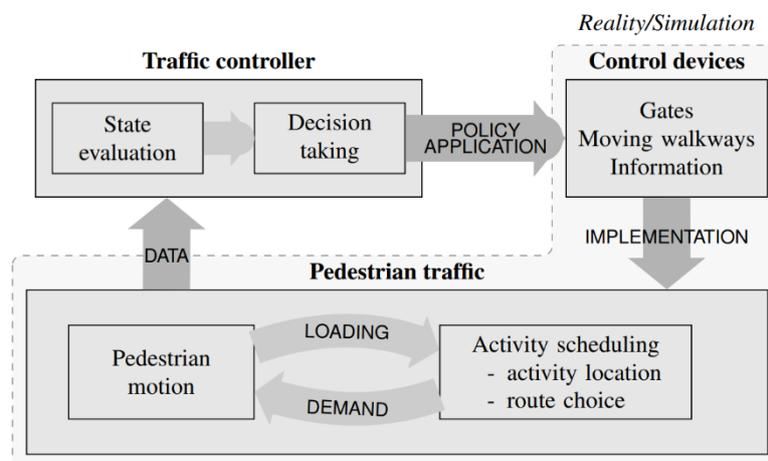


Figure 3: Interactions between the three main components in the framework. The dashed box surrounds the elements which can be either from “reality” or from a simulation.

element, an estimation of the state of the system (for example density). Secondly, based on this state evaluation, the management policies are used to decide how to update the control devices by fixing the values of the control variables.

### Control devices

To apply the management polices defined by the operator, some control devices must be installed. These are manipulated by the controller in order to reach some prespecified objective. This can be a target pedestrian density, travel time or transfer success rate for example.

These three major components are organized in a cycle, as presented in Figure 1. The data passed to the traffic controller (simulated or measured) depends the KPIs required to take the decisions. This can be pedestrian density, flow or travel time for example. The evaluation of the state of the system can be done based on any combination of the following: data from the current time, data from short term history or historical data. After the state of the system is estimated, the controller will compute, based on the control strategy which is used, the values of the various control variables. These values are then passed to the control devices, to be implemented by these devices.

### 2.2 Controller

The framework previously presented can be used with many different management strategies. To test the framework we use gates to control the flow of pedestrians entering a prespecified area, similarly to perimeter control for vehicular traffic (Keyvan-Ekbatani, et al., 2012). The development of the proportional-integral (PI) controller relies on the following two assumptions: 1) the existence of an aggregate fundamental diagram and 2) a linear relationship between the outflow from an area and the generalized flow within that area. Thanks to the empirical data collected in Lausanne's main train station, both of these assumptions have been verified.

The PI controller derived from the two assumptions stated previously and the conservation of pedestrian flow takes the following form:

$$q_{in}(k) = q_{in}(k - 1) - K_P[\rho(k) - \rho(k - 1)] + K_I[\hat{\rho} - \rho(k)] \quad (1)$$

where  $q_{in}(k)$  is the inflow during time interval  $k$ ,  $\rho(k)$  the density measured in zone A,  $\hat{\rho}$  the target density for zone A and  $K_P$  and  $K_I$  are the proportional and integral gains. Such “gains” represent the intensity with which the controller will react to the error between the measured and target densities. Both  $K_P$  and  $K_I$  must be estimated. This can be done either using the analytical expressions of these parameters, or a multivariate linear regression using least-squares on the empirical data. Figure 2 presents a simple case of gating where the flows entering the “junction” from two directions are controlled.

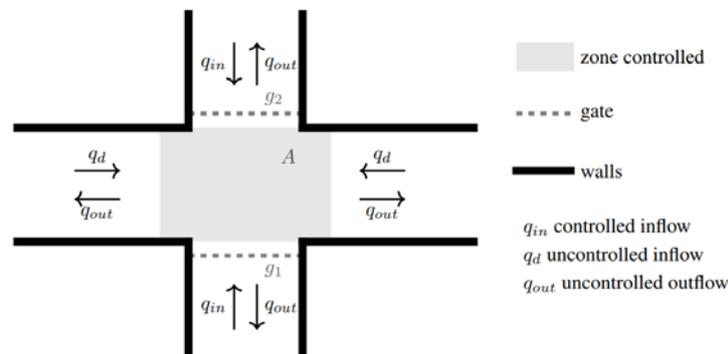


Figure 4: Schematic presentation of the gates controlling the inflow of pedestrians into a controlled zone denoted A. Each arrow represents a flow of pedestrians.

### 3. PRELIMINARY RESULTS

The case study considered for testing the framework and the gating controller is the train station of Lausanne, Switzerland. This location was chosen as individual tracking data has been collected in 2013 during the morning peak hours (7h to 8h30). This data is used for the calibration of the parameters  $K_P$  and  $K_I$  from Eq. (1).

Two pedestrian underpasses connect the platforms to each other and the main hall in Lausanne's train station. Simulations have been performed on one of the two pedestrian underpasses where gates are installed at the bottom of the access ramps to one of the pairs of platforms (the setup is similar to Figure 2). The framework has been implemented using state of the practice models for the pedestrian motion and route choice. The social force model is used to model the motion of pedestrians (Helbing, et al., 2005) and the route choice is modeled using the shortest path (Dijkstra, 1959). Although these models are relatively simple compared to more sophisticated approaches, the preliminary results are encouraging.

In order to evaluate the effectiveness of gating as a control strategy, two indicators are computed from the simulations. Firstly, the density inside the area under management is monitored throughout the simulation (which is also used by the PI controller). The second metric which is measured is the travel time of each pedestrian. We recall that the key objective of gating as a management strategy is to prevent the density from entering the flow-breakdown area. Therefore, in order to observe whether this objective is accomplished in a recurrent way, the mean density over multiple simulation runs is computed for two scenarios: one with gating and one without.

The scenario considered at this stage is the arrival of a train from which passengers alight. These individuals leave the platform and enter the pedestrian underpass and interact with other pedestrians already present. Figure 3 presents the results based on 100 simulation runs of the mean density inside the area under control (the junction at the bottom of the access ramp to the platform) alongside the mean travel time per simulation of all pedestrians in the system. The gating strategy effectively prevents the congestion from entering into the flow-breakdown stage, defined as a pedestrian density above  $1.0 \text{ pax}/\text{m}^2$ . Furthermore, the travel times of the pedestrians only increases by 8.5 seconds (18%). The extra travel time induced by congestion is limited.

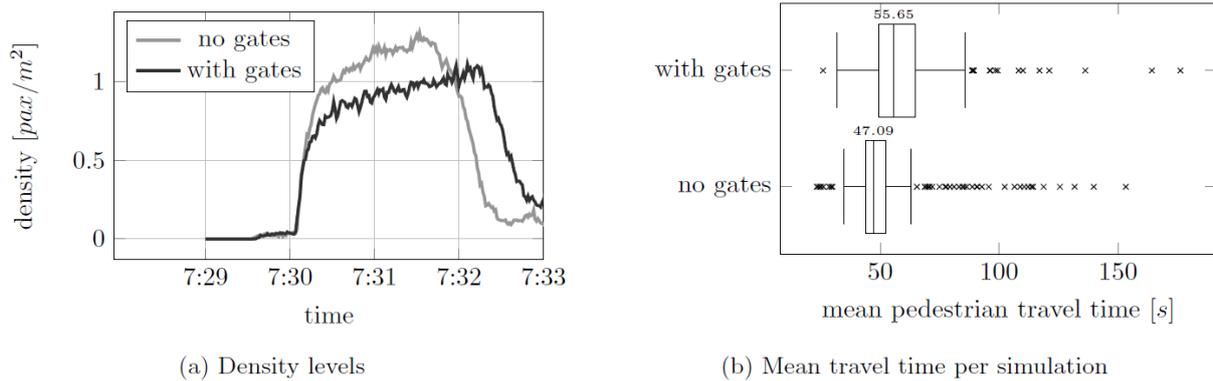


Figure 5: The objective of the controller to prevent the density from exceeding the threshold set at  $1.0 \text{ pax}/\text{m}^2$  is met. Furthermore, the mean travel times of all pedestrians in the system is only increased by 18% (based on the median of mean travel times). The value above the box plots is the median of the mean travel times.

### 4. CONCLUSION

We present a framework for controlling pedestrian flows using a dynamic traffic management system. This framework is made operational thanks to the usage of gating as a management strategy. The

objective consisting of preventing high densities is reached while the travel times are not significantly impacted. Furthermore, the application of the framework shows that investigating dynamic traffic management systems for pedestrians is promising. Many different problems can be addressed with a well-designed framework, such as analyzing pedestrian flows in shopping malls, the management of concert halls or conference centers or even evacuation scenarios.

In order to consolidate the preliminary results more advanced scenarios will be considered. The travel time of different groups of pedestrians will be analyzed for various scenarios and gating management policies. Alongside these simulations, more advanced models for pedestrian motion and route choice will be used.

Long term objectives include the utilization of different management strategies like information provision or a network of moving walkways. The second aspect which will be researched is the inclusion of the dynamic pedestrian traffic management system inside a simulation-based optimization framework for finding the optimal sets of parameters.

## **ACKNOWLEDGMENTS**

This research was performed as part of the TRANS-FORM (Smart transfers through unravelling urban form and travel flow dynamics) project funded by the Swiss Federal Office of Energy SFOE and Federal Office of Transport FOT grant agreement SI/501438-01 as part of JPI Urban Europe ERA-NET Cofound Smart Cities and Communities initiative. We thankfully acknowledge both agencies for their financial support.

## **BIBLIOGRAPHY**

- Ben-Akiva, M. et al., 2003. Evaluation of freeway control using a microscopic simulation laboratory. *Transportation Research Part C: Emerging Technologies*, 11(1), pp. 29-50.
- Dijkstra, E. W., 1959. A note on two problems in connexion with graphs. *Numerische mathematik*, 1(1), pp. 269-271.
- Helbing, D., Buzna, L., Johansson, A. & Werner, T., 2005. Self-Organized Pedestrian Crowd Dynamics: Experiments, Simulations, and Design Solutions. *Transportation Science*, 39(1), pp. 1-24.
- Keyvan-Ekbatani, M., Kouvelas, A., Papamichail, I. & Papageorgiou, M., 2012. Exploiting the fundamental diagram of urban networks for feedback-based gating. *Transportation Research Part B: Methodological*, 46(10), pp. 1393-1403.
- Mahmassani, H. S., 2001. Dynamic Network Traffic Assignment and Simulation Methodology for Advanced System Management Applications. *Networks and Spatial Economics*, Sep, 1(3), pp. 267-292.

# ARCHITECTURES OF URBAN TRAFFIC CONTROL WITH CONSIDERATION OF DRIVERS' ROUTE CHOICES

A. H. F. CHOW <sup>a</sup>

<sup>a</sup> *Department of Systems Engineering and Engineering Management,  
City University of Hong Kong, Hong Kong  
Email: [andychow@cityu.edu.hk](mailto:andychow@cityu.edu.hk)*

This paper investigates different architectures of control systems for managing network efficiency with consideration of users' responses. We formulate and analyse different control systems including centralised and decentralised (or distributed) systems. In a centralised system, all traffic information will be sent to a central controller or agent which is responsible for deriving and implementing all control actions. Following Diakaki et al. (2002), we formulate a centralised urban traffic control system based upon a cyclic store-and-forward traffic model. The signal settings are derived from a linear quadratic regulator (LQR) which aims to minimise the global network queues. The underlying store-and-forward traffic model in the optimal control formulation updates the traffic queues every signal cycle. In contrast, the decentralised or distributed systems delegate the control decisions to local controllers, with certain communication rules implemented between them. The study will explore various decomposition and decentralisation schemes, which include multi-agent based (de Oliveira and Camponogara, 2010) and alternating direction method of multipliers (ADMM).

**Keywords:** Adaptive Traffic Signals, Decentralisation, Linear Quadratic Regulator (LQR)

## 1. INTRODUCTION

Road networks are complicated systems that involve dynamics and interaction of a vast number of components including different types of junctions, vehicle classes, and road users. It is expected that efficiency and resilience of road networks can be significantly improved by coordinating all components under a centralised framework. Nevertheless, the complexity of urban networks makes it difficult to be managed by a single central system. This calls for a number of research effort on the more parsimonious and 'easy-to-run' distributed control paradigm in which the local components can derive their own control actions. It is understood that a centralised system will theoretically be able to derive a more effective policy than its distributed counterpart with better coordination. Nevertheless, the computational time involved when deriving such global optimal control plan increases exponentially with the size of the underlying transport networks and would eventually become intractable. One would argue that the computational efficiency gained by adopting distributed systems will have to come at the expense of the overall system-wide performance. However, a number of recent studies have suggested that the performance of a well-designed distributed system would not be significantly outperformed by a centralised one (see Chow, 2015; Chow and Sha, 2016).

This paper investigates different architectures of control systems for managing network efficiency with consideration of users' responses. We formulate and analyse different control systems including centralised and decentralised (or distributed) systems. In a centralised system, all traffic information will be sent to a central controller or agent which is responsible for deriving and implementing all control actions. Following Diakaki et al. (2002), we formulate a centralised urban traffic control system based upon a cyclic store-and-forward traffic model. The signal settings are derived from a linear quadratic regulator (LQR) which aims to minimise the global network queues. The underlying store-and-forward traffic model in the optimal control formulation updates the traffic queues every signal cycle. In contrast, the decentralised or distributed systems delegate the control decisions to local controllers, with certain communication rules implemented between them. The study will explore various decomposition and decentralisation schemes, which include multi-agent based (de Oliveira and Camponogara, 2010) and alternating direction method of multipliers (ADMM).

The control systems are implemented on the SUMO simulation platform over different network topologies and demand settings. SUMO (Simulation of urban mobility) is an open-source stochastic microscopic traffic flow model that is able to capture fine details and stochastic nature of traffic

dynamics. It also provides an application program interface (API) that can incorporate various external features including different kinds of signal control schemes and dynamic traffic routing and assignment algorithms. To capture the likely responses of drivers with respect to the variations in prevailing traffic condition, we further implement a dynamic (re-)routing algorithm to represent the responses of users with respect to the prevailing traffic conditions and signal settings. The results reveal that the decentralised systems could indeed perform as well as their centralised counterparts if the users are provided sufficient information and able to adjust their travel decisions with respect to prevailing traffic conditions appropriately. This study contributes to the state-of-art of traffic control design in a cooperative framework between supply and demand of transport system.

## 2. METHODOLOGY

In this paper we develop and investigate different optimal control formulations ranging from centralised to decentralised architectures with consideration of drivers' responses and uncertainties in prevailing traffic conditions. In a centralised traffic control system, all traffic measurements will be sent to a central controller or agent which is responsible for deriving and implementing all control actions. The centralised controller is formulated by using linear quadratic regulator (LQR) formulation. This centralised control system is formulated based upon a cyclic store-and-forward traffic model with an objective aiming to minimise the global network queues by adjusting the green splits  $\mathbf{g} = [g_i(k)]$  for all links  $i$  in signal cycle  $k$ . The underlying store-and-forward traffic model in the LQR formulation updates the traffic queues every signal cycle  $c$  given the green splits  $\mathbf{g}$ .

The LQR optimal control problem is formulated as:

$$\min_{\mathbf{g}} Z = \frac{1}{2} \sum_{k=0}^{\infty} (\mathbf{x}_k^T \mathbf{S} \mathbf{x}_k + \Delta \mathbf{g}_k^T \mathbf{R} \Delta \mathbf{g}_k), \quad (1)$$

where  $\mathbf{x}_k$  is a vector of all number of vehicles, usually regarded as all queue lengths in traffic engineering (Aboudolas, et al., 2009)  $x_i(k)$  on all links  $i$  by the end of signal cycle  $k$ ,  $\Delta \mathbf{g}_k$  is a vector of all  $\Delta g_i(k) = g_i(k) - g_i^N$  allocated to all links  $i$  in signal cycle  $k$  in which  $g_i^N$  is the associated nominal green time for link  $i$ . The control variable  $\Delta \mathbf{g}_k$  is introduced into the objective function (1) to regulate the changes in control  $\mathbf{g}$  and hence the 'aggressiveness' of the controller. The matrices  $\mathbf{S}$  and  $\mathbf{R} = r\mathbf{I}$  are non-negative definite, diagonal weighting matrices representing the trade-off between queue minimisation and rate of change of the control variable  $\mathbf{g}$ . The diagonal elements of  $\mathbf{S}$  are set to be the reciprocals of the admissible queue length  $x_i^*$  for all links  $i$ . The value of  $x_i^*$  depends on the physical dimension (e.g. length, width, etc) of the associated link  $i$ . The choice of  $r$  in  $\mathbf{R}$  nevertheless will be a trial-and-error process so as to achieve the most satisfactory control performance. It is noted that the original green split regulator is a centralised optimisation model as it considers all links  $i$ .

The dynamics of  $x_i(k)$  in (1) is govern as follows. If link  $i$  is a source link (i.e. connecting to an origin node), we have

$$x_i(k+1) = x_i(k) + c \left[ \lambda_i(k) - s_i \frac{g_i(k)}{c} \right], \quad (2)$$

where  $\lambda_i(k)$  is the demand rate (in [veh/time]) flowing link  $i$  from its upstream origin during cycle  $k$ ,  $g_i(k)$  is the total green time allocated to link  $i$  in cycle  $k$ ,  $s_i$  is the saturation flow of link  $i$ , and  $c$  is a predefined cycle time. If link  $i$  is an intermediate link, we have:

$$x_i(k+1) = x_i(k) + c \left[ \sum_{\forall j \in J(i)} \beta_{ji} s_j \frac{g_j(k)}{c} - s_i \frac{g_i(k)}{c} \right], \quad (3)$$

where  $J(i)$  is the set of all links  $j$  upstream of  $i$ ,  $\beta_{ji}$  specifies the proportion of flow in link  $j$  flowing

into  $i$ . The state equations (2) and (3) can be summarised in matrix form as:

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{B}\mathbf{g}_k + c\boldsymbol{\lambda}_k, \quad (4)$$

where  $\boldsymbol{\lambda}_k$  is a vector of all demand flows to the source links in the network,  $\mathbf{B}$  is a matrix derived from summarising the set of state equations (2) and (3), which is usually a sparse matrix of ‘minus’ saturation flows for each link depending on the network topology and turning ratios.

The optimality condition of (1), subject to (4) can be derived in the form of the following feedback control law on green time  $g_i(k)$  allocated to each link  $i$  over each cycle  $k$  as:

$$\mathbf{g}_k = \mathbf{g}_k^N - \mathbf{L} \cdot \mathbf{x}_k, \quad (5)$$

where  $\mathbf{L}$  is the corresponding control gain matrix associated with each link  $i$  in the network. The control gain matrix will be dependent on  $\mathbf{S}$ ,  $\mathbf{R}$ , and  $\mathbf{B}$ . The gain matrix  $\mathbf{L}$  can be determined by solving a Riccati equation (see Chow and Sha, 2016).

Diakaki et al (2002) consider an infinite horizon in cycle  $k$  where the control gain  $\mathbf{L}$  is calculated offline as its steady state solution. Consequently, the control law (5) can be operated readily with feeding information of queue lengths  $\mathbf{x}$  in real time without needing to run any optimisation algorithm in real time. This makes the centralised controller (5) computationally effective for real time operations. A weakness of this steady state approximation is that the corresponding control system will not be adaptive to prevailing traffic conditions and travellers' responses in real time. The study adopts a model predictive control formulation which implements the centralised control formulation (1) on a rolling horizon framework in which the network demand and split ratios will be estimated and fed into the control system in real time. To improve its computational effectiveness, the study will explore various decomposition and decentralisation schemes of (1), which include multi-agent based (de Oliveira and Camponogara, 2010) and alternating direction method of multipliers (ADMM) (Timotheou, et al, 2015). Contrasting with its centralised counterpart, the decentralised systems delegate some control decisions to local controllers, with certain communication rules implemented between them. Specifically, the local control actions do not require knowledge of global network inflow, and it adjusts local signal settings based upon local queue length measurements around each intersection (see e.g. Varaiya, 2013; Le et al., 2015; Chow, 2015).

### 3. PRELIMINARY RESULTS AND DISCUSSION

As an illustration, we apply different control systems on a microscopic simulation platform, called SUMO, over different network topologies. SUMO (Simulation of urban mobility) is an open-source stochastic microscopic traffic flow model that is able to capture fine details and stochastic nature of traffic dynamics. It also provides an application program interface (API) known as TraCI (Traffic Control Interface) that can incorporate various external features including different signal controllers and dynamic traffic (re-)routing algorithms. In addition to different networks, we also investigate performances of the control systems with different levels and distributions of traffic demand. The control systems considered herein include the centralised formulation (1) - (4), and a decentralised version of it by only keeping the diagonal elements of  $\mathbf{L}$  in (5). Results associated with more sophisticated decentralisation method of (1) - (4) will be presented in the full paper. The control systems are implemented into SUMO through this TraCI API.

The control systems are tested on a two-dimensional three-by-three grid network (see Figure 1) which contains turning movements and route choices over a two-dimensional plane. Each link in the network has a free-flow travel time of 10 sec and the saturation flow 1800 veh/h. The cycle time is set and fixed to be 90 sec. To incorporate the potential route changes of drivers with respect to prevailing traffic conditions, we implement an iterative routing algorithm as follows: given an origin-destination matrix, each vehicle will first proceed following the shortest path connecting its origin-destination pair. As the vehicle and other vehicles around it proceeds, travel times along each link in the network varies and hence updates of travel times will be (and can be) made known to each vehicle regularly, which is every second in the present study. Given the updates of link travel times, each vehicle will re-calculate its

shortest path and revise its routing decision whenever it reaches a node. The performances of different controllers are shown in Figure 2 over different network wide degree of saturation  $\gamma$  (see discussion in Chow and Sha, 2016) in which we include both with and without re-routing cases. The results first reveal that the centralised setting performs better than the decentralised one. Nevertheless, preliminary results reveal that the performance difference between the centralised and decentralised controllers can be reduced significantly when drivers are allowed to make appropriate re-routing decisions, say with provision of prevailing traffic conditions as in the numerical experiment. The observations suggest that, although the centralised controllers could outperform the decentralised ones as expected, the performance of the decentralised controllers can indeed be enhanced and brought close to their centralised counterpart with appropriate responses (e.g. re-routing) from the demand side. This reveals the potential of decentralised control integrated with demand management under a cooperative management framework.

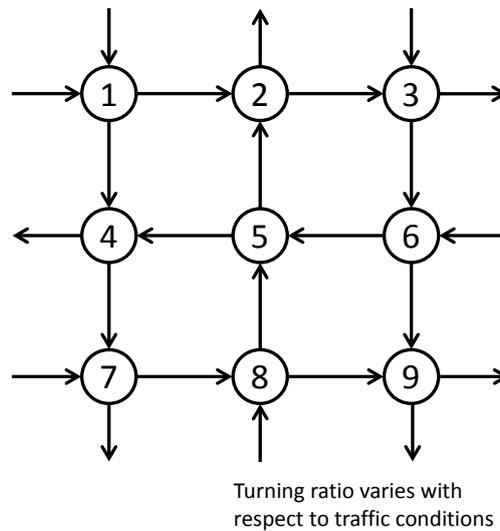


Figure 1: Example network

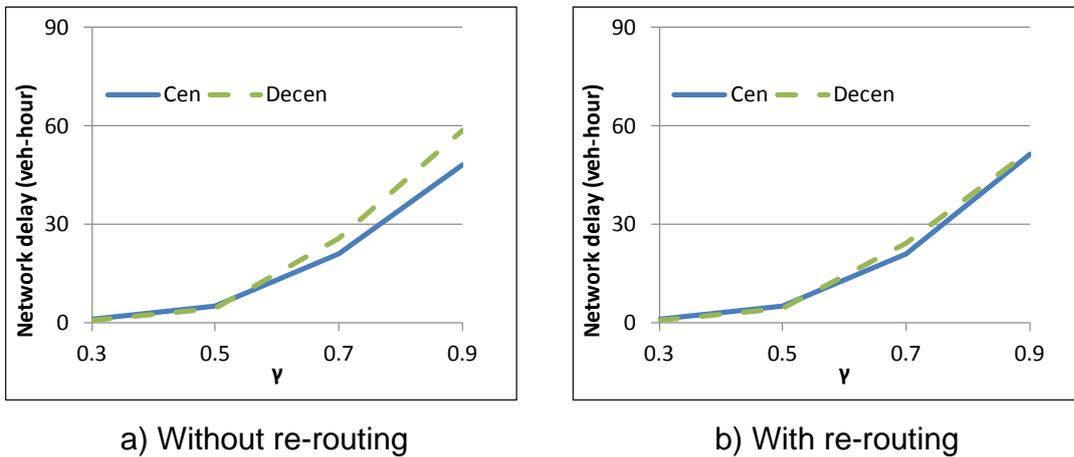


Figure 2: Experiment results

**REFERENCES**

Aboudolas, K., M. Papageorgiou, and E. Kosmatopoulos, Store-and-forward based methods for the signal control problem in large-scale congested urban road networks. *Transportation Research Part C*, Vol. 17, 2009, pp. 163–174.

Aboudolas, K., M. Papageorgiou, A. Kouvelas, and E. Kosmatopoulos, A rolling-horizon quadratic-programming approach to the signal control problem in large-scale congested urban road

- networks. *Transportation Research Part C*, Vol. 18, 2010, pp. 680–694.
- Chow, A.H.F., Optimisation of dynamic motorway traffic via a parsimonious and decentralised approach. *Transportation Research Part C*, Vol. 55, 2015, pp. 69–84.
- de Oliveira, L. and E. Camponogara, Multi-agent model predictive control of signaling split in urban traffic networks. *Transportation Research Part C*, Vol. 18, 2010, pp. 120–139.
- Diakaki, C., M. Papageorgiou, and K. Aboudolas, A multivariable regular approach to traffic responsive network-wide control. *Control Engineering Practice*, Vol. 10, 2002, pp. 183–195.
- Kouvelas, A., K. Aboudolas, M. Papageorgiou, and E. Kosmatopoulos, A hybrid strategy for real-time traffic signal control of urban road networks. *IEEE Transactions on Intelligent Transportation Systems*, Vol. 12, 2011, pp. 884–894.
- Le, T., P. Kovacs, N. Walton, H. Vu, and L. Andrew, Decentralised signal control for urban road networks. *Transportation Research Part C*, Vol. 58, 2015, pp. 431–450.
- Varaiya, P., Max pressure control of a network of signalized intersections. *Transportation Research Part C*, Vol. 36, 2013, pp. 177–195.

# DYNAMIC TRAFFIC SIGNAL OPTIMIZATION CONSIDERING NETWORK EQUILIBRIUM FLOWS

Wei HUANG<sup>a</sup> and Hong K. LO<sup>a</sup>

<sup>a</sup>*Department of Civil and Environmental Engineering,*

*The Hong Kong University of Science and Technology, Hong Kong, China*

*Email: [weihuang@ust.hk](mailto:weihuang@ust.hk); [cehklo@ust.hk](mailto:cehklo@ust.hk)*

This paper formulates a two-level control optimization problem, which integrates an anticipatory traffic control optimization and a real-time dynamic signal control optimization. The anticipatory network traffic control calculates optimal steady-state performance based on an equilibrium assignment model, and provides a reference plan for the local dynamic signal settings. In order to seek consistent long-term and short-term control decisions, a coordination algorithm is proposed. This study shows how a rolling horizon approach is tailored for the local control to both accommodate varying traffic conditions and track the strategic reference plan. Moreover, an iterative learning scheme is introduced to update the reference plan, using observations from the operational level. The iterative learning control scheme elevates the traffic system to its optimal steady-state performance despite the presence of modeling error inherent in the equilibrium model used to anticipate travelers' responses.

**Keywords:** Adaptive Signal Control, Network Traffic Control, Hierarchical, Steady-State Reference, Rolling Horizon

## 1. INTRODUCTION

Traffic signal control is one important form of traffic control and management. While playing an important role in traffic management, traffic signals must be properly designed and operated to realize their full potentials. In fact, optimizing signal settings is considered a low cost approach to alleviate congestion (Srinivasa Sunkari, 2004). Traffic signal control can be broadly classified into fixed-time and traffic-responsive approaches. Being able to adapt to changing conditions, traffic-responsive methods can bring substantial benefits in reducing vehicle delays.

Most of the operating traffic controls assume exogenously determined traffic flows. From the perspective of network planning, signal control affects travelers' route choice as travelers generally factor signalized delays into their route choice decisions and settle into long-term equilibrium flows, which in turn influence the optimization of signal settings. Therefore, the route choice effect needs to be endogenized into the network signal control problem. An *anticipatory traffic control* was introduced to model the interaction between traffic assignment (route choice) and signal control (Taale, 2008): the controller anticipates traveler's route choice response to changes in signal settings, and determines signal settings to optimize a network performance, e.g., network total travel time. As one specific instance of the network design problem (NDP), anticipatory control is commonly formulated as a bi-level optimization problem, or mathematical programming with equilibrium constraints (Huang et al., 2016).

The network-wide anticipatory control enables signal control systems to explicitly incorporate the behavioral aspects of road users, permitting transport planning activities to be addressed by signal control adjustments. However, throughout the traffic signal control literature, there has been a separation from the concern of holistic transport management, which can be clearly seen by distinguishing models used by transport planners and signal control engineers (Smith et al., 2001). Moreover, planning and real-time control applications are usually characterized by different time resolutions. While the local control focuses on cycle-by-cycle dynamic signal timings, anticipatory control determines the control parameters at a coarse-resolution level, for instance, a longer-term average green split for a horizon of 30 minutes. As a result, there is an obvious inefficacy if anticipatory controls are implemented to manage changes in the prevailing traffic conditions; and the other way round, if dynamic signal controls are applied for network design that intends to encourage users' choices to be made in favor of steady-state performance. In order to seek consistent long-term and short-term decision-making, this study develops a two-level hierarchical control strategy to integrate the control decisions made at different timescales.

## **1.1 A brief hierarchical traffic signal control context**

A hierarchical structure has been adopted in many adaptive traffic control systems. UTOPIA (Mauro and Di Taranto, 1989) is a hybrid control system that is constructed with an area level and a local level. The area controller generates a reference plan for a longer time period (e.g. a horizon of 30 minutes), and the local controller adapts this reference plan and optimizes signal setting for a short horizon (e.g., 120s). RHODES traffic-adaptive signal control system consists of three hierarchical levels (Head et al., 1992). The highest level is a “dynamic network loading” model that captures the slow-varying characteristics of traffic and generates different demand patterns for its lower level. The middle level is referred to as “network flow control”, which allocates green times for each different demand pattern and each signal phase. Given the approximate green times, the “intersection control” at the third level determines phase switching based on observed and predicted arrivals at each intersection.

The basic concept of hierarchical control systems is to decompose the large-scale traffic control problem into several sub-problems that are interconnected at multiple resolution levels. In most of the operating traffic control systems, however, a strategic level that explicitly deals with the gaming of control settings and road users’ choice behavior is missing. Ignoring users’ route choice freedom may result in unexpected traffic states, which may at the end deteriorate the control performance. This motivates the development of a new control strategy that is able to lead the traffic system towards a desired traffic state thus avoiding future performance deterioration, while retaining the real-time adaptability features.

## **1.2 A feedback control context**

Anticipatory traffic control determines an optimal traffic control with user equilibrium constraints. Traffic control is consistent when the equilibrium flow predictions on which it is based are verified after users react to it. Ideally, an accurate equilibrium flow prediction model, also called flow response function, is used to determine the optimal anticipatory control. In view of the complexity of gaming between equilibrium flow and signal control, the prediction model is inevitably subject to errors and hence rendering suboptimal solutions. An effective anticipatory control requires explicitly correcting the modeling error by learning from the real user responses.

Using measurements to improve control performance has been extensively investigated for real-time control applications. A dynamic controller typically adjusts its control parameters according to the detected traffic. Heuristic feedback laws have been applied for freeway dynamic traffic control, such as ALINEA (Papageorgiou et al., 1991). Methods have also been introduced to adjust traffic propagation modeling, on both freeway and urban road networks, in a rolling-horizon approach (Aboudolas et al., 2010). In such real-time adaptive control, traffic observations are used to ‘reset’ the model predicted state, i.e., feedback serves to re-initialize a new prediction and repeat the model-based predictive control process. However, the flow propagation model does not change during the process.

In the context of steady-state control, the equilibrium flow prediction error may become one major cause of control performance deterioration, as the equilibrium model that captures user behavior over a relatively long period will have changed in response to a new control setting. In this regard, correcting the model approximation error contributes substantially to an effective control design. Different from the aforementioned feedback controls that adjust the controller’s knowledge of the current traffic state itself, this paper focuses on another feedback mechanism that explicitly corrects the controller’s prediction model.

The main contributions of this paper are stated as follows:

- 1) We formulate a two-level control optimization problem, which encapsulates an anticipatory traffic control optimization and a real-time dynamic signal control optimization.
- 2) A coordination algorithm is proposed to integrate the anticipatory control and dynamic signal settings under a consistent framework.
- 3) An adaptive adjustment scheme is proposed to update the anticipatory control, so that it adapts to the observations of the controller’s interaction with road users and traffic environment.

## 2. PROBLEM FORMULATION

An overview of the proposed two-level control system is shown in Figure 1. The anticipatory traffic control optimization focuses on steady-state control and locates at the strategic level. The time horizon is usually in hours. It determines long-term control plans, e.g. green split in a peak hour, while accounting for the optimal equilibrium performance. The long-term green split and the equilibrium flow pattern are then passed to the operational level as a reference plan. On the other hand, the local level control calculates dynamic traffic signal settings based on a flow propagation model, for example the cell transmission model for signal optimization (Lo, 2001). The time span at this level is typically in seconds or cycles. It determines cycle-by-cycle dynamic timing plans in response to changes in the prevailing traffic conditions following a rolling horizon approach. Moreover, the local signal control is required to track the desired reference plan provided by the strategic level. A coordination algorithm is proposed for the tracking control to ensure that the local control complies with the steady-state control. After implementation of the dynamic control in real traffic networks, the changes in traffic states (e.g., traffic flow patterns) are then communicated back to the strategic level to update the anticipatory control. Re-optimization of the steady-state reference is usually performed in days or weeks, for the travelers to experience the control changes and settle into a new equilibrium state after an epoch of days operation.

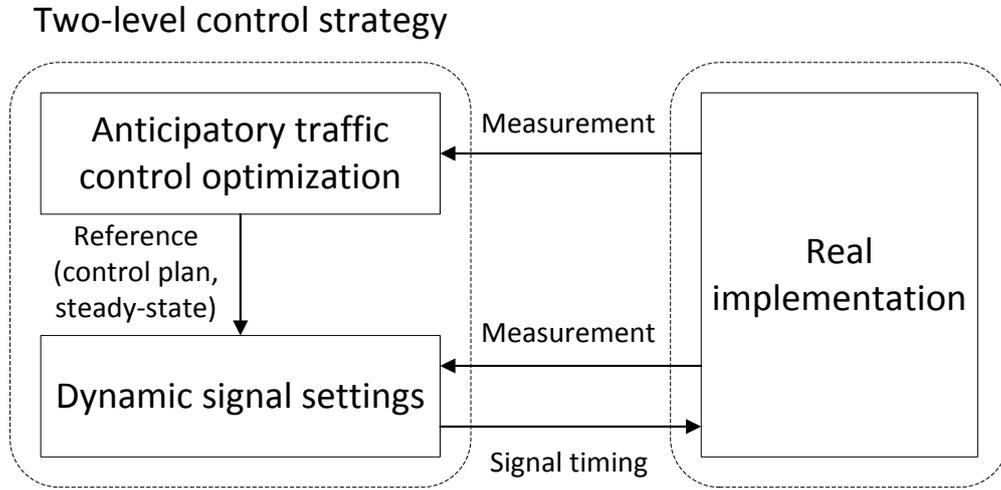


Figure 1. An overview of the two-level control system

The anticipatory traffic control optimization problem is formulated as follows:

$$\min_{\mathbf{g}^r} z(\mathbf{g}^r, \mathbf{f}^r) \quad (1)$$

subject to

$$\mathbf{f}^r = f(\mathbf{g}^r) \quad (2)$$

$$\underline{\mathbf{g}}^r \leq \mathbf{g}^r \leq \bar{\mathbf{g}}^r \quad (3)$$

where  $z(.,.)$  is the objective function, for instance, network total travel time.  $f(\mathbf{g}^r)$  is the user equilibrium flow obtained with any given control variable  $\mathbf{g}^r$ .  $\underline{\mathbf{g}}^r$  and  $\bar{\mathbf{g}}^r$  are lower bound and upper bound of the control variables, respectively.

Regarding the inherent structural modeling error in the equilibrium flow response function  $f(.,.)$ , an approximation function is used to explicitly account for model bias.

$$\tilde{\mathbf{f}} = \psi(f(\mathbf{g}^r), \mathbf{b}) \quad (4)$$

in which  $\mathbf{b}$  represents the structural model bias, we may perform a model bias correction whenever a new observation of equilibrium flow  $\tilde{\mathbf{f}}$  becomes available.

$$\Delta \mathbf{b}_{k+1} = b(\tilde{\mathbf{f}}_k, \mathbf{b}_k) \quad (5)$$

in which,  $\Delta \mathbf{b}_{k+1}$  is the model bias update between two consecutive iterations, i.e.,  $\mathbf{b}_{k+1} - \mathbf{b}_k$ . The updated model bias is then fed into the anticipatory control optimization problem to re-optimize the steady-state reference. Different learning paradigms can be designed to supervise the update of model bias. An iterative learning scheme has been developed for the model bias correction in Huang et al. (2016).

The local control level considers online dynamic signal settings. We denote the simulation time step as  $t_s$  and the control time step as  $t_c$ . Let the prediction horizon be  $N_p$ , and the control time interval is an integer multiple  $I$  of the simulation time interval. The dynamic signal setting optimization problem can be expressed as:

$$\min_{\mathbf{g}(t_c)} J(\hat{\mathbf{n}}(t_s), \mathbf{g}(t_c)) \quad (6)$$

subject to

$$\hat{\mathbf{n}}(t_s) = \mathbf{h}(\mathbf{n}(t_s), \mathbf{g}(t_c)) \quad (7)$$

$$\underline{\mathbf{g}}(t_c) \leq \mathbf{g}(t_c) \leq \overline{\mathbf{g}}(t_c) \quad (8)$$

$$\mathbf{f}(\hat{\mathbf{n}}(t_s), N_p) = \mathbf{f}^r \quad (9)$$

in which  $J(.,.)$  is the control objective that models the accumulated performance over the prediction horizon.  $\hat{\mathbf{n}}(t_s)$  refers to the traffic state at simulation time step  $t_s$ , predicted with a traffic flow model  $\mathbf{h}(.,.)$ .  $\mathbf{n}(t_s)$  represents the real or measured state. The decision variables are signal settings over the prediction horizon.

$$\hat{\mathbf{n}}(t_s) = [\hat{\mathbf{n}}(t_s + 1|t_s) \ \hat{\mathbf{n}}(t_s + 2|t_s) \ \dots \ \hat{\mathbf{n}}(t_s + N_p I|t_s)]^T \quad (10)$$

$$\mathbf{g}(t_c) = [\mathbf{g}(t_c|t_c) \ \mathbf{g}(t_c + 1|t_c) \ \dots \ \mathbf{g}(t_c + N_p - 1|t_c)]^T \quad (11)$$

A rolling horizon scheme is applied which enables the controller to obtain feedback from the real traffic network and adapt to the disturbances of the traffic environment. Equation (9) shows that, at each rolling step, a constraint on the steady-state flow pattern is considered, i.e., the aggregate link flow over the prediction horizon should comply with the equilibrium flow pattern. Therefore, the dynamic optimization problem is solved in such a way that the signal settings track the optimal reference passed from the strategic level.

### 3. CONCLUSION

This paper proposes a two-level control strategy integrating adaptive anticipatory traffic control and cycle-by-cycle dynamic signal settings. An anticipatory control calculates optimal steady-state performance and provides a reference plan for the operational signal settings. This study shows how a rolling horizon approach is tailored for the local control to accommodate varying traffic conditions and track the desired reference plan. Moreover, an iterative learning control scheme is introduced to update the reference plan, so that the system seeks to achieve optimal steady-state performance despite the presence of modeling error.

### REFERENCES

- Aboudolas, K., Papageorgiou, M., Kouvelas, A., Kosmatopoulos, E. (2010) A rolling-horizon quadratic-programming approach to the signal control problem in large-scale congested urban road networks. *Transportation Research Part C*, 18(5), pp. 680-694.
- Head, K.L., Mirchandani, P.B., Sheppard, D. (1992) Hierarchical framework for real-time traffic control. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1360, pp. 82-88.
- Huang, W, Viti, F., Tampère, C.M.J. (2016) Repeated anticipatory network traffic control using iterative optimization accounting for model bias correction. *Transportation Research Part C*, 67, pp. 243-265.

- Lo, H.K. (2001) A cell-based traffic control formulation: strategies and benefits of dynamic timing plans. *Transportation Science*, 35(2), pp. 148-164.
- Mauro, V., Di Taranto, C. (1989) UTOPIA. In *Proceedings of the 6th IFAC/IFIP/IFORS symposium on control computers and communication in transportation*, Paris, France, pp. 575-597.
- Papageorgiou, M., Hadj-Salem, H., Blosseville, J.M. (1991) ALINEA: A local feedback control law for on-ramp metering. U.S. Dept. Transp., Washington, DC, *Transportation Research Record* 1320.
- Smith, M.J., Clegg, J., Yarrow, R. (2001) Modeling traffic signal control. In: Button KJ, Hensher DA (eds) *Handbook of Transport Systems and Traffic Control*, Elsevier Science Ltd, Kidlington, Oxford, pp. 503-525.
- Srinivasa Sunkari, P.E. (2004) The benefits of retiming traffic signals. *ITE Journal*, pp. 26-29.
- Taale, H. (2008) *Integrated Anticipatory Control of Road Networks. A game theoretical approach*. Ph.D. thesis, TRAIL thesis series T2008/15, Delft University of Technology, Delft, The Netherlands.

# SAFETY, STABILITY AND SMOOTHNESS IN CONTROL OF CONNECTED AUTONOMOUS VEHICLES

B. G. HEYDECKER <sup>a</sup> and Y. D. CHEN <sup>a</sup>

<sup>a</sup> *Centre for Transport Studies, University College London, UK*

*Email: [b.heydecker@ucl.ac.uk](mailto:b.heydecker@ucl.ac.uk)*

The increasing availability of information and communications technology (ICT) offers advantages for the management and control of autonomous vehicles through connectivity within the fleet. This can be used to share information at the distinct levels of actions about to be undertaken as the result of decisions made information that will influence joint decision making intentions for travel beyond the current state (position, speed, direction).

Different degrees of cooperation among the vehicle fleet are possible according to which of these levels is implemented. In the present paper, we investigate the effect on cooperation through communication on the control of autonomous vehicles. In particular, we consider the effects of leading vehicles communicating their acceleration and braking decisions with following vehicles that are equipped to use this to anticipate behaviour of their leading vehicles. This concept of local broadcast of driving actions through vehicle ad-hoc networks (Vanets) is established as an effective means to stabilise the flow of vehicles, and represents one form of cooperation among vehicles in the fleet.

**Keywords:** CAV (Connected Autonomous Vehicles), Vehicle Following, ITS, Stability, Smoothness

## 1. INTRODUCTION

The increasing availability of information and communications technology (ICT) offers advantages for the management and control of autonomous vehicles through connectivity within the fleet. This can be used to share information at the distinct levels of

actions about to be undertaken as the result of decisions made  
information that will influence joint decision making  
intentions for travel beyond the current state (position, speed, direction).

Different degrees of cooperation among the vehicle fleet are possible according to which of these levels is implemented. In the present paper, we investigate the effect on cooperation through communication on the control of autonomous vehicles. In particular, we consider the effects of leading vehicles communicating their acceleration and braking decisions with following vehicles that are equipped to use this to anticipate behaviour of their leading vehicles. This concept of local broadcast of driving actions through vehicle ad-hoc networks (Vanets) is established as an effective means to stabilise the flow of vehicles, and represents one form of cooperation among vehicles in the fleet.

In this work, we consider a vehicle-following model that includes as stimulus each of relative speed and deviation from a speed-dependent desired spacing, with response being the acceleration (or braking) of the following vehicle. We manipulate this model into a second-order differential equation in spacing, with a forcing term that comprises the desired spacing for the vehicles speed plus the communicated acceleration of the lead vehicle. We investigate various properties of this formulation in terms of the choice of parameters that represent sensitivity of response to the stimuli, temporal lag in response to stimulus, choice of formulation for desired spacing, and influence of the lead vehicle acceleration. The analysis is developed to identify criteria for the relationships among the parameters to achieve good dynamic behaviour in non-equilibrium conditions. We present analyses of safety that provide lower bounds on sensitivity in order to achieve safe (*i.e.* collision-free) following. We then consider stability in the senses of platoon stability and string stability. These properties are determined by the complementary solutions to the differential equation: the presence of forcing terms influences the particular solutions but does not affect stability. Instead, we consider the effect of the form of the forcing terms on the smoothness of following in a string of vehicles in the sense of the propagation of perturbations in trajectory through the string of following vehicles. We show that if each following vehicle's acceleration is forced by their lead vehicle's, then disturbances will be propagated without attenuation, leading to qualitative lack of smoothness in the flow of traffic.

## 2. MODEL FORMULATION

We develop a car-following model for dense traffic that is formulated as a feedback control towards the set point of zero speed relative to the lead vehicle with spacing dependent on speed. In the linear case, this model has two sensitivity parameters, controlling respectively the response to the deviations in speed and in spacing. The model form accommodates a choice of equilibrium speed-spacing relationships, and also a choice of reference speed for calculation of the spacing. We identify requirements on the model functions and parameter values for rational driving and macroscopic behaviour.

We suppose that the following vehicle adapts to the speed of the lead vehicle, and seeks a spacing that is desirable according to its speed. The response is acceleration of the following vehicle, which can be expressed as:

$$\ddot{x}(t+T) = f_v [\dot{x}_{n-1} - \dot{x}_n(t)] + f_s [x_{n-1}(t) - x_n(t) - D_n [\dot{x}_n(t)]],$$

where  $x_n(t)$  is the position of vehicle  $n$  at time  $t$ ,  $\dot{x}$  is the derivative of  $x$  etc with respect to time, the function  $f_v(\cdot)$  represents the sensitivity to deviations from 0 relative speed, and  $f_s(\cdot)$  represents the sensitivity to deviations from speed-dependent desired spacing  $D_n(v_n)$ . Neglecting the lag  $T$  and expressing this in terms of spacing, this can be put in the form of a second-order differential equation with forcing term corresponding to the acceleration of the lead vehicle:

$$\ddot{s}_n(t+T) + a_v \dot{s}_n(t) + a_s s_n(t) = \ddot{x}_{n-1}(t+T) + a_s D_n \{\dot{x}_n(t)\},$$

where  $s_n(t) = x_{n-1}(t) - x_n(t)$  is the spacing in front of vehicle  $n$  at time  $t$ .

The desired velocity is considered as the velocity of the vehicle in front and the desired spacing is calculated from the classical macroscopic velocity density relationships. The platoon stability and string stability of the model (no lag) are initially analysed theoretically using the modelling frame developed by Wilson and Ward (2011). It is found that as long as the rational driving constraints are satisfied, the model will always be platoon stability. Sensitivity  $X$  of equilibrium spacing to variation in velocity is expressed in forms for different macroscopic relationships.

We analyse further the case that the sensitivity of the following vehicle is proportional to difference in speed and deviation from desired spacing, so that the functions  $f(\cdot)$  are linear with constant of proportionality. We show that this model will be string stable if two parameters  $\alpha_1$  and  $\alpha_2$  satisfy  $\alpha_1 > \frac{1}{X} - \frac{X}{2} \alpha_2$  when the speed of the following vehicle is used to calculate the desired spacing ( $\tilde{v} = v_n$ ). The model will be string stable if two parameters satisfy  $\alpha_1 > \frac{1}{X} + \frac{X}{2} \alpha_2$  when the velocity of the lead vehicle is used to calculate the desired spacing ( $\tilde{v} = v_{n-1}$ ). Various macroscopic relationships are investigated for each case. We show that Underwood's macroscopic relationship with  $\tilde{v} = v_n$  shows the most appropriate behaviour in high density traffic conditions, though Edie's conjunct relationship for high density condition with  $\tilde{v} = v_n$  also shows good behaviour. This model is also investigated numerically with increasing the time lag  $\tau$  and time increment  $T$  of solution. It is also found that either increase in either of time increment  $T$  or time lag  $\tau$  will worsen the platoon stability. It is also found increase in  $\alpha_1$  value requires decrease in  $\alpha_2$  for the model to remain platoon stable. The choice of these two parameters has no lower bound provided that they are strictly positive. However, in order to ensure stability the upper bound value for both parameters will decrease as time lag increases, as is indicated by Herman, Montroll, Potts and Oliver (1959).

We proceed to investigate numerically the effect of providing following vehicles with the acceleration of the lead vehicle, as could be achieved through a vanet or other local communication. If the following vehicle adopts all of the lead vehicle's acceleration in addition to its response to deviation from desired state, then its trajectory will include all fluctuations of the lead vehicle. This is illustrated by the

trajectories of vehicles 1-6 shown in Figure 1, where the undisturbed trajectory of vehicle 1 is taken as the reference. In this case, vehicle 7 alone does not adopt the acceleration of its lead vehicle (vehicle 6). Thus for vehicle  $n = 7$

$$\ddot{s}_n(t+T) + a_v \dot{s}_n(t) + a_s s_n(t) = a_s D_n \{ \dot{x}_n(t) \}$$

The effect that its trajectory is substantially smoother: this qualitative smoothness is then transferred to all subsequent vehicles, which in this case do again adopt the acceleration of their lead vehicle.

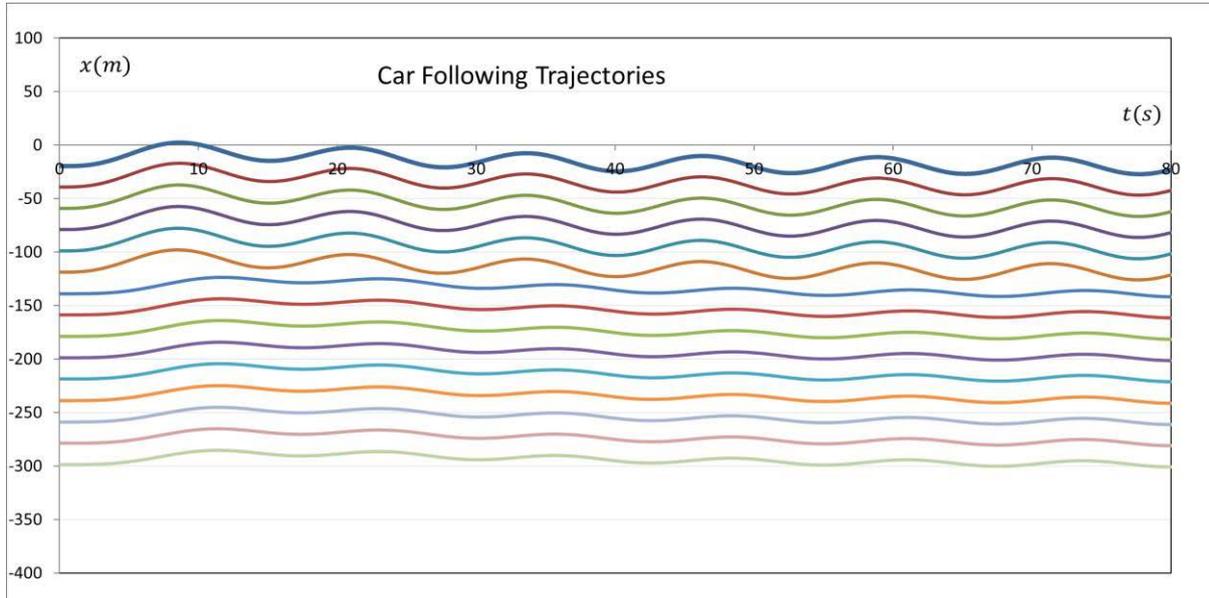


Figure 1: Normalised trajectories of following vehicles.

### 3. CONCLUSIONS

The analysis presented here establishes properties of car-following models depending on their sensitivity to plausible stimuli and information provided from vehicles downstream. The various desirable properties of stability and smoothness depend on the sensitivities of the following rules. In particular use of information about the acceleration of the lead vehicle in calculating that of a following vehicle can result in propagation of fluctuations through a platoon, so that this should be factored down. These results will be relevant to automated vehicle following mechanisms if they are to be used throughout the fleet in order that the macroscopic behaviour of traffic remains acceptable.

### REFERENCES

- Herman, R., Montroll, W., Potts, R.B. and Rothery, R.W. (1959) Traffic dynamics: Analysis of stability in car following. *Operations Research*, 7, pp. 86-106.
- Wilson, R.E. and Ward, J.A. (2011) Car-following models: Fifty years of linear stability analysis – A mathematical perspective. *Transportation Planning and Technology*, 34(1), pp. 3-18.

**Session B1**  
**Traffic and Demand Management I**

# TRAFFIC NETWORK PARTITIONING FOR HIERARCHICAL MACROSCOPIC FUNDAMENTAL DIAGRAM APPLICATIONS BASED ON FUSION OF GPS PROBE AND LOOP DETECTOR DATA

Kang AN<sup>a</sup>, Xianbiao HU<sup>b\*</sup> and Xiaohong CHEN<sup>c</sup>

<sup>a</sup> *Key Laboratory of Road and Traffic Engineering of the State Ministry of Education, School of Transportation Engineering, Tongji University, Shanghai 201804, China*

*Email: [1310061@tongji.edu.cn](mailto:1310061@tongji.edu.cn)*

<sup>b\*</sup> *Department of Civil, Architectural and Environmental Engineering, Missouri S&T, Rolla, MO 65401, USA*

*Email: [xbhu@mst.edu](mailto:xbhu@mst.edu)*

<sup>c</sup> *Key Laboratory of Road and Traffic Engineering of the State Ministry of Education, School of Transportation Engineering, Tongji University, Shanghai 201804, China,*

*Email: [chenxh@tongji.edu.cn](mailto:chenxh@tongji.edu.cn)*

Macroscopic Fundamental Diagram (MFD) enables researchers and practitioners to model, monitor and control traffic at an aggregate way without the need to go into detailed link information. Previous network partitioning methods assume complete link-level traffic information over the entire network, which makes them inapplicable when it comes to real-world setting. In a previous research by the authors, an approach based on region growing technique and lambda-connectedness definition was proposed to partition a network while allowing certain degree of data missing ratio. The result shows traffic engineers still need to install loop detectors on at least 60% of the links in order to get good quality network partitioning and calibration result, which is still expensive to set up and maintain. In this paper, we propose a method which, based on fusing Probe Vehicle Data (PVD) and loop detector data (LDD), extracts the locally homogeneous subnetworks with a grid-level network approach instead of dealing with detailed link-level network. With such approach, the key requirement on data size of having sufficient probe vehicles at any given link will be relaxed to having a reasonable number of probe vehicles in the vicinity to capture the local traffic dynamics, which becomes more feasible and practicable.

**Keywords:** Network Partition, Macroscopic Fundamental Diagram, Data Fusion, Region Growing, Loop Detector Data, GPS Probe Vehicle Data

## 1. INTRODUCTION

Macroscopic Fundamental Diagram (MFD), as a graphical method to characterize traffic state in a homogeneous road network, enables researchers and practitioners to model, monitor and control traffic at an aggregate way without the need to go into detailed link information. Various network partitioning methods were proposed in prior studies to partition a heterogeneous network into multiple homogeneous sub-networks. These approaches are mostly based on a normalized cut mechanism, which takes the traffic statistics of each link (e.g. link volume, speed or density) as input to calculate the degree of similarity between two links, and perform graph cut to divide a network into two sub-networks at each iteration when the traffic dynamics between links are dramatically different. These methods assume complete link-level traffic information over the entire network, e.g. the accurate measurement of the traffic conditions exist for every single link in the network, which makes them inapplicable when it comes to real-world setting. In a previous research by the authors (An et al. 2016, An et al. 2016), an approach based on region growing technique and lambda-connectedness definition was proposed to partition a network while allowing certain degree of data missing ratio. The sensitivity analysis result shows traffic engineers still need to install loop detectors on at least 60% of the links in order to get good quality network partitioning and calibration result, which is still expensive to set up and maintain.

Probe vehicles with GPS devices have received much attention for its potential as a new sensing technology in the past decade. While probe sensors can record instantaneous vehicle position and speed information, which allows us to track vehicle trajectory and reflects traffic dynamics along the traffic network over both time and space dimensions, they're mostly used to derive link-level traffic dynamics such as average travel time, delay and speed (Feng et al. 2014, Noh et al. 2016), Having sufficient GPS probe vehicle data (PVD) coverage over the entire network to ensure a good amount of sample size on

any link in the network at any time becomes a prerequisite of performing analytics and research with this approach. As a result, partitioning traffic network with GPS probe vehicle data as the sole data source becomes challenging, if practicable at all. The other main issue associated with PVD utilization in MFD related research is that previous studies usually assume a homogeneous distribution of the probe vehicles (Leclercq et al. 2014), i.e. an equal market penetration rate over the entire network, which is rather unrealistic.

In this paper, we propose a method which, based on fusing PVD and loop detector data (LDD), extracts the locally homogeneous subnetworks with a grid-level network approach instead of dealing with detailed link-level network. By fusing the two data sources, we take advantage of both better coverage from PVD and the full-size detection from LDD. The concept we're trying to explore and prove in this research is that, as opposed to the traditional approach of aggregating data to generate link-level traffic statistics and then perform network partition, can we approach this problem with a grid-level solution? Here, a grid is defined as the basic unit in the traffic network, which comprises multiple roadway nearby segments and can be merged to a large subnetwork or further divided into smaller grid if needed. The basic idea is that by fusing available PVD and LDD within this grid, an estimate of traffic dynamics will be computed, and then network partitioning algorithm will be further developed to perform network growing and partitioning based on the similarities between adjacent grids. With such approach, the key requirement on data size of having sufficient probe vehicles at any given link will be relaxed to having a reasonable number of probe vehicles in the vicinity to capture the local traffic dynamics, which becomes more feasible and practicable.

Another natural merit of this modeling approach is that the penetration rate of the probe vehicles, which is a key variable in measuring and calibrating MFD (Edie 1965) and is calculated by the ratio of number of probe vehicles to the total traffic counts from loop detectors, can be computed for each grid at different time slots, instead of assuming a single value for the entire network. It's also worth noting that for the data from loop detectors, only the volume is used but not the speed or density information, which makes this proposed model generically applicable to not only freeway but also arterial network, regardless of which part of road segment are they installed.

## 2. METHODOLOGY

The objective of this paper is to explore a way to derive traffic dynamics with the fusion of GPS probe vehicle data and loop detector data and further utilize this information to partition the traffic network into homogeneous subnetworks to ensure the existence of macroscopic fundamental diagrams. To be specific, this paper seeks an approach with the fusion of the two data sources that can achieve the following objectives.

- 1) Implement the grid definition to help take advantages of both GPS probe vehicle data and loop detector data
- 2) Derive an estimate of traffic dynamics with both GPS probe vehicle data and loop detector data at the grid level considering varying market penetration rates
- 3) Extract homogeneous subnetworks with similar density values to ensure the existence of MFD
- 4) Build a smooth boundary for each subnetwork to ease the implementation of traffic controls

By taking advantage of both better coverage from GPS probe vehicle data and the full-size detection from loop detector data, and instead of dealing with link-level traffic dynamics, this paper tries to estimate the grid-level traffic state information with the fusion of GPS probe vehicle data and loop detector data and further utilize this information to partition the traffic network into homogeneous subnetworks with similar traffic states to ensure the existence of macroscopic fundamental diagram. To achieve this goal, four modules are designed to realize the approach and they are illustrated in the following paragraphs and Figure 1.

**Module 1 – Grid Definition.** To avoid the requirements of detailed link-level information, this module tries to cluster the nearby road segments into one grid and further to implement the fusion method of GPS probe vehicle data and loop detector data to estimate the traffic dynamic at the grid level. This

module is designed to meet the objective #1.

**Module 2 – Traffic Dynamics Estimation with Varying Market Penetration Rates.** There're two main steps in this module. The first one is probe vehicle penetration rate estimation and the second one estimates traffic dynamics within the grid with a formulation modified based on Edie's equation (Edie 1965). With the first step, the penetration rate of the probe vehicles can be computed for each grid at different time slots, instead of assuming a single value for the entire network.

**Module 3 – Network Conversion.** Once the grids are defined and their dynamics computed, this step will perform a network conversion so that the grids in the original network will become nodes in the new network, and the connectivity between grids in the original network will become links in the new network. The length, capacity and traffic properties of the links in the new network will also be defined accordingly.

**Module 4 – Iterative Region Growing.** This module contains multiple iterations of different steps that partition the traffic network into homogeneous subnetworks to meet the several requirements of objective #3 and #4. This module utilizes the region growing algorithm to take advantage of the heuristics and high computation efficiency. In each iteration, the penetration rate and traffic density of each subnetwork is updated using the same definitions in module 2 and further the region growing technique is performed to grow these subnetworks from last iteration to new clusters based on the similarity definition. At the end of each iteration, the subnetworks are checked to see if certain stopping criteria are met or not.

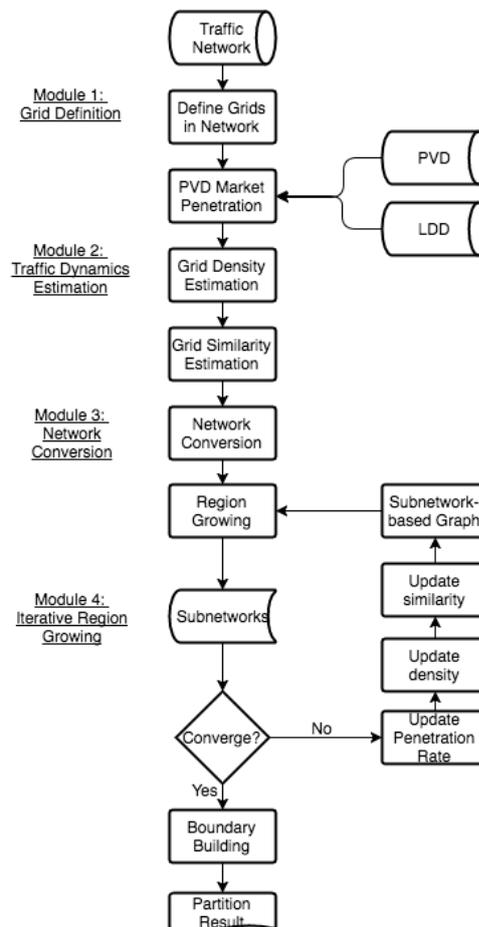


Figure 6. General Framework

### 3. PRELIMINARY ANALYSIS

Figure 2 shows the network partition result. While further analysis is under way, preliminary analysis shows that compared with previous approach that requires a single data source of 60% loop detector

data, the proposed new approach is able to achieve comparable result with a scenario of data availability as low as 15%.

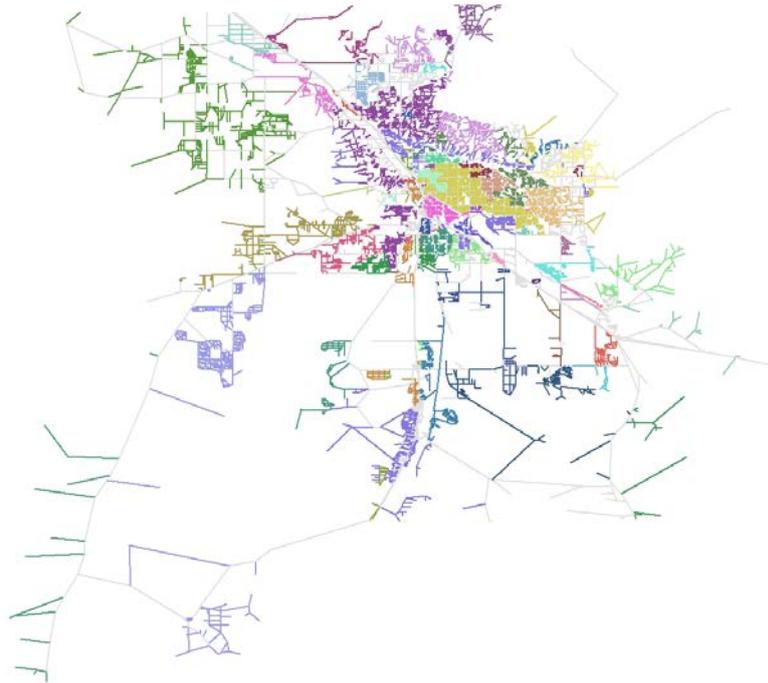


Figure 7 network partition result

Other numeric analysis in progress include analysis of minimum grid size, heuristics of grid definition, sensitive analysis with regard to data availabilities and other aspects.

## REFERENCES

- An, K., Y.-C. Chiu, X. Hu and X. Chen (2016). A Network Partitioning Algorithm to Support Macroscopic Fundamental Diagram Based Hierarchical Traffic Network Control and Management. *IEEE Transactions on Intelligent Transportation Systems*.
- An, K., X. Hu, Y.-C. Chiu and X. Chen (2016). A Network Partitioning Algorithm to Support Macroscopic Fundamental Diagram Based Network Modeling. 96th Annual Meeting of the Transportation Research Board. Washington DC.
- Edie, L. (1965). Discussion of Traffic Stream Measurement and Description. Proc. Second Intern. Symp. on the Theory of Road Traffic Flow, J. Almond, Ed., OECD, Paris.
- Feng, Y., J. Hourdos and G. A. Davis (2014). "Probe vehicle based real-time traffic monitoring on urban roadways." *Transportation Research Part C: Emerging Technologies* 40: 160-178.
- Leclercq, L., N. Chiabaut and B. Trinquier (2014). "Macroscopic Fundamental Diagrams: A cross-comparison of estimation methods." *Transportation Research Part B: Methodological* 62(Supplement C): 1-12.
- Noh, H., A. Sun, X. Hu and A. Rehan (2016). Development of a regional network performance measurement model for planning application based on high-frequency GPS data. *Transportation Research Record: Journal of the Transportation Research Board*.

# AN OPTIMAL CONTROL FRAMEWORK FOR MULTI-REGION MACROSCOPIC FUNDAMENTAL DIAGRAM SYSTEMS CONSIDERING ROUTE CHOICE AND DEPARTURE TIME CHOICE

R. X. ZHONG<sup>a\*</sup>, Y. P. HUANG<sup>a</sup>, J. H. XIONG<sup>a</sup>, N. ZHENG<sup>b</sup>, William LAM<sup>c</sup> and A. SUMALEE<sup>c</sup>

<sup>a</sup> *Guangdong Key Laboratory of Intelligent Transportation Systems, School of Engineering, Sun Yat-sen University, Guangzhou, China*

*Email: [zhrenxin@mail.sysu.edu.cn](mailto:zhrenxin@mail.sysu.edu.cn), [huangyp9@mail2.sysu.edu.cn](mailto:huangyp9@mail2.sysu.edu.cn), [xiongh3@mail2.sysu.edu.cn](mailto:xiongh3@mail2.sysu.edu.cn)*

<sup>b</sup> *School of Transportation Science and Engineering, Beihang University, Beijing, China*

*Email: [nan\\_zheng@buaa.edu.cn](mailto:nan_zheng@buaa.edu.cn)*

<sup>c</sup> *Department of Civil and Environmental Engineering,*

*The Hong Kong Polytechnic University, Hong Kong SAR, China*

*Email: [cehklam@polyu.edu.hk](mailto:cehklam@polyu.edu.hk), [asumalee@gmail.com](mailto:asumalee@gmail.com)*

Macroscopic fundamental diagram (MFD) has been used for aggregate modeling of urban traffic network dynamics under stationary traffic assumption for dynamic taxi dispatching, vehicle relocation and dynamic pricing schemes to tackle the dimensionality problem of microscopic approaches. A city is assumed to be partitioned into several regions with each admits a well-defined MFD. To overcome the limitation of inconsistent lags, the MFD model with time delay, wherein the regional travel time function is modeled as an endogenous time-varying delay, is adopted to describe the traffic dynamics within a region. On the other hand, it is necessary to enable simultaneous route choice and departure time choice under the MFD framework for various applications such as vehicle dispatching and relocation. This paper presents an optimal control framework to model dynamic system optimum (DSO) with simultaneous route choice behavior and departure time choice for general urban networks. Necessary conditions for the DSO are analytically derived through the lens of Pontryagin minimum principle. In contrast to existing analytical methods, the proposed method is applicable for general MFD systems without approximation schemes of the equilibrium solution. Numerical examples by time discretization are conducted to illustrate the characteristics of DSO and corresponding dynamic external costs.

**Keywords:** Macroscopic Fundamental Diagram, Nonlinear Travel Time Function, Dynamic System Optimal

## 1. INTRODUCTION

Macroscopic fundamental diagram (MFD), establishing a mapping from the network flow accumulation to the trip completion rate, has been widely used for aggregate modeling of urban traffic network dynamics under stationary traffic assumption (Geroliminis and Sun, 2011; Saberi et al., 2014; Laval and Castrilln, 2015). The travel time is regarded as a time delay in the control input by Keyvan-Ekbatani et al. (2012); Haddad and Mirkin (2016) wherein the control input delays are presented and assumed to be constant in the traffic flow dynamics for perimeter control design to simplify the model description. However, since the network accumulation is a dynamic process, travel time, as a function of the network accumulation, is thus time-varying and endogenous. How to handle such kind of endogenous time-varying delay subject to state and input constraints is still missing in transportation literature.

For vehicle dispatching and relocation, it is necessary to enable route choice under the MFD framework. Recently, efforts have been dedicated to combine the dynamic traffic assignment (DTA) and the MFD framework by incorporating route choices or preferences in case of heterogeneous networks with multiple MFD regions because drivers have the ability to choose a different sequence of regions to decrease their travel times (Haddad et al., 2013; Yildirimoglu and Geroliminis, 2014; Yildirimoglu et al., 2015). Haddad et al. (2013) developed a dynamic simple route choice model to account for the change in route choice decision in response to the new control policy wherein travelers choose the route with minimum instantaneous travel time while Yildirimoglu and Geroliminis (2014) investigated the feasibility of MFD modeling integrated with aggregated route choice dynamics and developed a region-based route choice model. And later on, the UE and SO traffic conditions for MFD-based traffic models with route choice behavior aggregation were presented in Yildirimoglu et al. (2015). However, these existing methods are either algorithmic or numerical (Laval et al., 2017). Laval et al. (2017) considered

the delay in the dynamic user equilibrium (DUE) model analytically for a simple network with a freeway as an alternative to the city street network modeled by the MFD dynamics wherein several definitional constraints on the state and input are ignored. Route choices/guidance for general network under the MFD framework is still missing in the MFD literature that addresses the aforementioned endogenous delay in a systematical manner. Therefore, the dynamic system optimum (DSO) analysis under the MFD framework remains an interesting challenge.

## 2. MFD-BASED NETWORK LOADING MODEL

An urban traffic network is partitioned into  $M$  regions on condition that the traffic is homogeneously distributed within each region and thus each admits a well-defined MFD, which establishes a concave mapping from the network output (or exit)  $G_i(n_i)$  (veh/s) to network accumulation state  $n_i(t)$  (veh). By definition, the average network travel time function of  $R_i$  is then calculated as  $h_i(t) = n_i(t) / G_i(n_i(t))$ . We then adopt the travel time function in the MFD system flow conservation. The relationship between the inflow and outflow rate is  $(1 + \dot{h}_i(t))G_i(t + h_i(t)) = q_i(t)$ . Since both  $G_i(\cdot)$  and  $q_i(\cdot)$  are nonnegative,  $1 + \dot{h}_i(t) \geq 0$ . If  $1 + \dot{h}_i(t) > 0$ , then  $G_i(t + h_i(t)) = q_i(t) / (1 + \dot{h}_i(t))$ .

To enable route choice in the multi-region MFD system, we define sequential movements along different MFD regions, i.e., a path/route, by  $p \doteq \{R_1, R_2, \dots, R_{m(p)}\}$ ,  $p \in P$ , where  $m(p)$  is the number of regions used by path  $p$ ;  $P$  is the set of all paths in the network. To enable route choice and departure time choice for the MFD framework,  $q^p(t)$  denotes departure rate to a particular path  $p$  at a particular time instant  $t$ . Along the path, flow exiting from the preceding upstream region is taken as input to the downstream region. The flow propagation dynamics for the regions along path  $p$  is given as:

$$\mathbf{d}n_{R_i}^p(t) / \mathbf{d}t = q^p(t) - G_{R_i}^p(t), \quad \mathbf{d}n_{R_i}^p(t) / \mathbf{d}t = G_{R_{i-1}}^p(t) - G_{R_i}^p(t), \quad \forall p \in P, i \in [2, m(p)]$$

where  $n_{R_i}^p(t)$  denotes the network accumulation state of  $R_i$  traveling on path  $p$  at time  $t$ ;  $q^p(t) = G_{R_0}^p(t)$ .

$\tau_{R_i}^p(t)$ , the corresponding exit time for vehicles entering region  $R_i$  at time  $\tau_{R_{i-1}}^p(t)$ , is then proceeded as:

$$\tau_{R_i}^p(t) = t + h_{R_i}(n_{R_i}(t)) = t + n_i(t) / G_i(n_i(t)), \quad \tau_{R_i}^p(t) = \tau_{R_{i-1}}^p(t) + h_{R_i}(n_{R_i}(\tau_{R_{i-1}}^p(t))), \quad \forall p \in P, i \in [2, m(p)]$$

Journey time of a path is evaluated by the nested delay operator as

$$h_p(t, \mathbf{n}) := \sum_{i=1}^{m(p)} \delta_{R_i}^p \Phi_{R_i}(t, \mathbf{n}), \quad \forall p \in P, \quad \mathbf{n} = (n_{R_i} : \forall R_i \in \mathcal{R}), \quad \delta_{R_i}^p = \begin{cases} 1, & \text{if } R_i \in p \\ 0, & \text{otherwise} \end{cases}$$

$$\Phi_{R_i}(t, \mathbf{n}) = h_{R_i}(n_{R_i}(t)), \quad \Phi_{R_i}(t, \mathbf{n}) = h_{R_i}(n_{R_i}(t + \Phi_{R_1} + \dots + \Phi_{R_{i-1}})) = h_{R_i}(n_{R_i}(t + \sum_{j=1}^{i-1} \Phi_{R_j})), \quad \forall i \in [2, m(p)]$$

Combining early/late arrival penalty  $\kappa(\chi)$  with the journey time achieves the effective path delay/cost operators, i.e.,  $\Psi_p(t, \mathbf{n}) = h_p(t, \mathbf{n}) + \kappa(\chi)$ ,  $\forall p \in P$  wherein  $\chi = t + h_p(t, \mathbf{n}) - t^*$ ,  $t^* < T$ .

## 3. OPTIMAL CONTROL FORMULATION AND EQUILIBRIUM CONDITION FOR DSO

The DSO equilibrium cost provides a bound on the best performance of a traffic network. For a set of connected regions, consider a finite time planning horizon  $t \in [0, T]$ ,  $T > 0$ , each origin-destination (OD) pair  $w \in W$  is with a fixed total amount of demand  $Q_w$  to be served. An optimal control formulation of the DSO that seeks an optimal route inflow profile  $q^{p*}(t)$  to minimize the total system travel cost within the study period  $[0, T]$  is given as:

$$\min \sum_{\forall p \in P} \int_0^T \Psi_p(t, \mathbf{n}) q^p(t) \mathbf{d}t$$

subject to

$$\mathbf{d}n_{R_i}^p(t) / \mathbf{d}t = q^p(t) - G_{R_i}^p(t), (\alpha_{R_i}^p) \quad \forall p \in P \quad (1a)$$

$$\mathbf{d}n_{R_i}^p(t) / \mathbf{d}t = G_{R_{i-1}}^p(t) - G_{R_i}^p(t), (\alpha_{R_i}^p) \quad \forall p \in P, i \in [2, m(p)] \quad (1b)$$

$$\mathbf{d}E_w(t) / \mathbf{d}t = \sum_{p \in P_w} q^p(t), (\rho_w) \quad \forall w \in W \quad (1c)$$

$$G_{R_i}^p(t + h_{R_i}(n_{R_i}(t))) (1 + h'_{R_i}(n_{R_i}(t)) \cdot \dot{n}_{R_i}(t)) = q^p(t), (\beta_{R_i}^p) \quad \forall p \in P \quad (1d)$$

$$G_{R_i}^p(t + h_{R_i}(n_{R_i}(t))) (1 + h'_{R_i}(n_{R_i}(t)) \cdot \dot{n}_{R_i}(t)) = G_{R_{i-1}}^p(t), (\beta_{R_i}^p) \quad \forall p \in P, i \in [2, m(p)] \quad (1e)$$

$$-q^p(t) \leq 0, (\gamma_{R_i}^p) \quad \forall p \in P \quad (1f)$$

$$-n_{R_i}^p(t) \leq 0, (\lambda_{R_i}^p) \quad \forall p \in P \quad (1g)$$

$$n_{R_i}(t) \leq n_{R_i}^{jam}, (\eta_{R_i}) \quad \forall R_i \in \mathcal{R} \quad (1h)$$

$$E_w(T) = Q_w, (\phi_w) \quad \forall w \in W \quad (1i)$$

$$n_{R_i}(0) = 0, \quad \forall R_i \in \mathcal{R}, \quad E_w(0) = 0, \quad \forall w \in W, \quad q^p(0) = 0, \quad \forall p \in P \quad (1j)$$

where

$$n_{R_i}(t) = \sum_{p \in P} n_{R_i}^p(t) \delta_{R_i}^p, \quad \forall R_i \in \mathcal{R}; \quad Q_w = \sum_{p \in P_w} \int_0^T q^p(t) \mathbf{d}t, \quad \forall w \in W$$

$n_{R_i}(t)$  is the total accumulation state of  $R_i$  at time  $t$ ; ' is differentiation to function argument, and  $\cdot$  is differentiation to time  $t$ .  $P_w$  is the set of paths connecting OD pair  $w$  and  $E_w(t)$  is an extended state. (1a)-(1b) are network traffic dynamics. (1d)-(1e) are flow propagation constraints. Note for any given  $q^p(t) \geq 0, \forall p \in P$ , we have  $G_{R_i}^p(t) \geq 0, \forall p \in P, \forall R_i \in \mathcal{R}$ . (1c) and (1i) are flow conservation constraints, while (1j) specifies the zero initial conditions. (1h) is saturation constraint imposed by MFD. The variables in brackets of (1a)-(1i) are Lagrange multipliers.

**Proposition 3.1.** The equilibrium condition for DSO for general MFD system with simultaneous route choice and departure time choice is:

$$q^p(t) \begin{cases} > 0 \Rightarrow \Psi_p(t, \mathbf{n}^*) + \frac{\partial \Psi_p(t, \mathbf{n}^*)}{\partial q^p(t)} q^p(t) + l_p(t) = \phi_w, \\ = 0 \Rightarrow \Psi_p(t, \mathbf{n}^*) + \frac{\partial \Psi_p(t, \mathbf{n}^*)}{\partial q^p(t)} q^p(t) + l_p(t) > \phi_w, \end{cases} \quad \forall p \in P_w, w \in W$$

where  $\mathbf{n}^*$  is the optimal state of the DSO problem for OD pair  $w$ ,  $\phi_w$  is the travel cost under the DSO equilibrium condition, specified by the fixed total travel demand  $Q_w$ . The dynamic external cost includes

two parts, i.e.,  $\frac{\partial \Psi_p(t, \mathbf{n}^*)}{\partial q^p(t)} q^p(t)$ , and  $l_p(t)$ . The first part is the sensitivity value of the total system

travel cost with respect to a perturbation in the path flow  $q^p(t)$ .

$$l_p(t) = \sum_{i=0}^{m(p)-1} \int_{\tau_i^p(t)}^{\tau_{i+1}^p(t)} (\zeta_{i+1}^p(u) + \eta_{i+1}(u) \delta_{i+1}^p) \mathbf{d}u, \quad \zeta_i^p(u) = q_p(t) \frac{\partial \Psi_p(t, \mathbf{n})}{\partial h_p(t, \mathbf{n})} \frac{\partial h_p(t, \mathbf{n})}{\partial n_i^p(u)}, \quad u \in [\tau_{i-1}^p(t), \tau_{i+1}^p(t)]$$

represents the second part of the dynamic external cost, where  $\int_{\tau_{i-1}^p(t)}^{\tau_i^p(t)} \eta_i(u) \delta_i^p \mathbf{d}u$  attains the cost to access the restricted region  $R_i$  with  $\eta_i$  denoting the Lagrange multiplier associated with the saturation

constraint imposed on  $R_i$ , and  $\int_{\tau_{i-1}^p(t)}^{\tau_i^p(t)} \zeta_i^p(u) \mathbf{d}u$  denotes another part of the dynamic external cost imposed on travelers caused by their presence on  $R_i$  along path  $p$ .

#### 4. NUMERICAL EXAMPLES

Figure 1 depicts the Braess' network consisting of 4 regions with a single OD pair tied by 2 paths. The OD pair from  $R_1$  to  $R_4$  is connected by 2 paths,  $R_1 \rightarrow R_2 \rightarrow R_4$  (Path 1) and  $R_1 \rightarrow R_3 \rightarrow R_4$  (Path 2)

while the configuration of each region is specified in Table 1. The total travel amount between  $R_1$  and  $R_4$  is  $Q_{od} = 1500$  (units). The planning horizon is  $T=80$ (unit-times) with regional inflow rate restricted to be less than or equal to 100 (unit/time). Expected arrival time interval is specified as  $[t_{de}, t_{dl}] = [50, 70]$ . The DSO is solved with a linear early/late arrival penalty as

$$\kappa[\chi] = \begin{cases} 2(t + h(n) - t_{de}), & t + h(n) < t_{de} \\ 0, & t_{de} \leq t + h(n) \leq t_{dl} \\ 4(t + h(n) - t_{dl}), & t + h(n) > t_{dl} \end{cases}$$

Path inflow profiles are plotted against the generalized travel cost profiles in Figure 2(a). None of the path is utilized at the very beginning since the cost exceeds the equilibrium cost. Both paths then become active when the equilibrium travel cost is reached. Since the travel time expected to traverse Path 1 is smaller than that of Path 2, the departure time window of Path 1 is a little wider than that of Path 2. As indicated in the figure, dynamic external costs arise along with the inflow rates. Travelers also choose their paths to balance the equilibrium costs of the two paths to avoid more dynamic external cost. Corresponding accumulation states of the four regions are depicted in Figure 2(b).

Note that the generalized travel cost oscillates even within the departure window. In fact, similar phenomenon was also observed in the DTA literature with the simple linear travel time function, see e.g., Chow (2009). Chow (2009) speculated two reasons for this: a) the larger congestible portion induces more complex traffic dynamics and hence complicates the calculation process of dynamic externality and thus equilibrium condition; b) the effect of time discretization: finer discretization could induce a smoother inflow profile at the price of the curse of dimensionality for optimization. The focus of this paper is on the properties of the DSO under the MFD framework rather than on the numerical treatments. We speculate that the disequilibrium happens may be also due to the continuity of the effective cost function for general multi-region networks. When the link travel time function (or the regional travel time function in our case) is linear, Han and Friesz (2017) proved that the effective cost function is continuous. This continuity guarantees the existence of optimal solutions to the DSO. For the MFD system with nonlinear travel time function, the continuity of the effective cost function remains unknown.

Table 1: Network configurations: Nonlinear network travel time function

| Region $i$ | Network configuration                 |                                       |                          |
|------------|---------------------------------------|---------------------------------------|--------------------------|
| $R_1$      | $a_1 = 20 \cdot 1.4877 \cdot 10^{-7}$ | $b_1 = 20 \cdot 2.9815 \cdot 10^{-3}$ | $c_1 = 20 \cdot 15.0912$ |
| $R_2$      | $a_2 = 20 \cdot 1.4877 \cdot 10^{-7}$ | $b_2 = 30 \cdot 2.9815 \cdot 10^{-3}$ | $c_2 = 21 \cdot 15.0912$ |
| $R_3$      | $a_3 = 20 \cdot 1.4877 \cdot 10^{-7}$ | $b_3 = 20 \cdot 2.9815 \cdot 10^{-3}$ | $c_3 = 20 \cdot 15.0912$ |
| $R_4$      | $a_4 = 20 \cdot 1.4877 \cdot 10^{-7}$ | $b_4 = 20 \cdot 2.9815 \cdot 10^{-3}$ | $c_4 = 20 \cdot 15.0912$ |

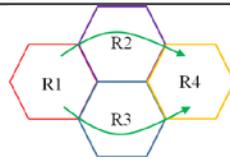
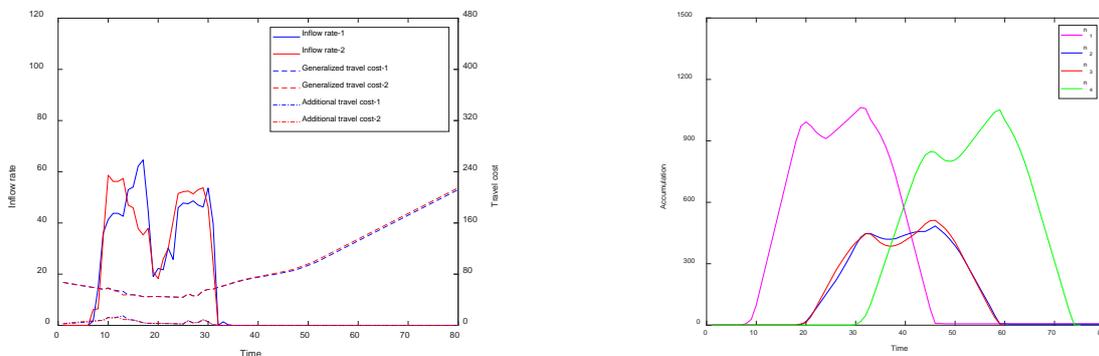


Figure 1. An example of the Braess' network



(a) Inflow rate and travel cost under DSO

(b) Accumulation state under DSO

Figure 2. An example of the Braess' network with nonlinear travel time function

## REFERENCES

- Chow, A. H., 2009. Dynamic system optimal traffic assignment – a state-dependent control theoretic approach. *Transportmetrica*, 5(2):85–106.
- Geroliminis, N. and Sun, J., 2011. Hysteresis phenomena of a macroscopic fundamental diagram in freeway networks. *Transportation Research Part A*, 45(9):966–979.
- Haddad, J. and Mirkin, B., 2016. Adaptive perimeter traffic control of urban road networks based on MFD model with time delays. *International Journal of Robust and Nonlinear Control*, 26(6):1267–1285.
- Haddad, J., Ramezani, M., and Geroliminis, N., 2013. Cooperative traffic control of a mixed network with two urban regions and a freeway. *Transportation Research Part B*, 54:17–36.
- Han, K. and Friesz, T. L., 2017. Continuity of the effective delay operator for networks based on the link delay model. *Networks and Spatial Economics*, 17(4):1095–1110.
- Keyvan-Ekbatani, M., Kouvelas, A., Papamichail, I., and Papageorgiou, M., 2012. Exploiting the fundamental diagram of urban networks for feedback-based gating. *Transportation Research Part B*, 46(10):1393–1403.
- Laval, J. A. and Castrillon, F., 2015. Stochastic Approximations for the Macroscopic Fundamental Diagram of Urban Networks. *Transportation Research Part B*, 81:904–916.
- Laval, J. A., Leclercq, L., and Chiabaut, N., 2017. Minimal parameter formulations of the dynamic user equilibrium using macroscopic urban models: Freeway vs city streets revisited. *Transportation Research Part B*. doi: <https://doi.org/10.1016/j.trb.2017.08.027>.
- Saberi, M., Mahmassani, H. S., Hou, T., and Zockaie, A., 2014. Estimating Network Fundamental Diagram Using Three-Dimensional Vehicle Trajectories: Extending Edies Definitions of Traffic Flow Variables to Networks. *Transportation Research Record*, (2422):12–20.
- Yildirimoglu, M. and Geroliminis, N., 2014. Approximating dynamic equilibrium conditions with macroscopic fundamental diagrams. *Transportation Research Part B*, 70:186–200.
- Yildirimoglu, M., Ramezani, M., and Geroliminis, N., 2015. Equilibrium analysis and route guidance in large-scale networks with MFD dynamics. *Transportation Research Part C*, 59:404–420.

**Session B2**  
**Day-to-Day Dynamics I**

# DAY-TO-DAY MULTIMODAL DYNAMIC TRAFFIC ASSIGNMENT: IMPACTS OF THE LEARNING PROCESS IN CASE OF NON-UNIQUE SOLUTIONS

Mostafa AMELI<sup>a,b</sup>, Jean Patrick LEBACQUE<sup>a</sup> and Ludovic LECLERCQ<sup>b</sup>

<sup>a</sup> *University of Paris-Est, IFSTTAR, GRETTIA, France*

*Email: [mostafa.ameli@ifsttar.fr](mailto:mostafa.ameli@ifsttar.fr)*

*Email: [jean-patrick.lebacque@ifsttar.fr](mailto:jean-patrick.lebacque@ifsttar.fr)*

<sup>b</sup> *Univ. Lyon, IFSTTAR, ENTPE, LICIT, France*

*Email: [ludovic.leclercq@ifsttar.fr](mailto:ludovic.leclercq@ifsttar.fr)*

Unicity of the solution is an important issue in the Dynamic Traffic Assignment (DTA) problem. A key argument for unicity is strictly monotone path travel time function with respect to the number of travelers that use the path. There is no guaranty to hold the monotonicity condition in the multimodal urban transportation network. This study aims to investigate the result of the evolution of a transportation system over a long-term day to day learning process. Specifically, and mathematically speaking, we are going to address the impact of initial solutions in the day-to-day multimodal DTA. The analysis the impact of initialization will be carried out in the concept of trip-based day-to-day DTA. This work considers a large-scale multimodal network with flexible opening over time of several public transport facilities and attempts to find deterministic equilibria by simulation-based optimization. To analyze the impact of non-unicity, we are going to work on the learning process when the facilities are introduced in a different order during the day-to-day process. The results show that not only do we have non-unicity but also, that we can save a lot of hours on total travel time by opening the public transportation facilities in the optimal order.

**Keywords:** Day-To-Day, Dynamic Traffic Assignment, Unicity Analysis, Trip-Based, Multimodal

## 1. INTRODUCTION

Beckman *et al.* (1956) indicate the importance of unicity and stability of the DTA solution. The conditions of unicity have been properly reviewed by Iryo (2013) for flow-based DTA by considering the different approach to solving the DTA problem. Moreover, many researches have been done to prove the unicity of DTA solutions on flow-based dynamic traffic model with several assumptions and limitations on traffic network model (Mounce, 2007; Mounce and Smith, 2007; Iryo 2011; Iryo 2015; Iryo and Smith, 2017). This study adapts the conditions to trip-based DTA wherein each traveler is defined as an individual. A key characteristic for unicity is strictly monotone path travel time function with respect to the number of travelers that use the path. However, for some systems like multimodal urban transportation networks (Mounce, 2001), the monotonicity condition simply not hold. In such cases, the final equilibrium depends on the initial network states and the convergence process. This study aims to investigate the end result of the evolution of a transportation system over a long-term day to day learning process. Mathematically speaking, we are going to address the impact of initial solutions in the day-to-day multimodal DTA.

The goal is first, to prove the non-unicity of the problem with theory and experiments and then analyze the impact of initialization on solving process of trip-based day-to-day DTA problems. In this work, we consider a trip-based multi-modal approach to network equilibrium. We assume that mode and path choice are carried out at the same level, therefore travel time (TT) depends on travel path and the mode(s) attributes of travelers. Travelers in the traffic network attempt to minimize their own TT. Therefore, the solution of assignment problem based on Wardrop's first principle (Wardrop, 1952) is called User Equilibrium (UE). If travelers have an indifference bound for minimization, the solution will be Bounded Rational User Equilibrium (BRUE) (Di and Liu, 2016). In order to find the equilibrium, this study considers a day-to-day optimization process wherein the assignment pattern is changed day by day in order to find UE or BRUE. Smith and Wisten (1995) show that under UE conditions there is an equilibrium state of this dynamical procedure. This work considers a large-scale network with fixed demand but flexible opening over time of several public transport facilities and attempts to find deterministic equilibria by simulation-based optimization.

The learning curve of travelers plays a big role in day-to-day DTA to guide the optimization process to find a stable equilibrium. Learning curve defines as a guide to determine the direction of changing the status of the network to find the equilibrium (Peeta and Ziliaskopoulos, 2001). It contains the predictive TT (by dynamic traffic model) and perceived TT (by simulator) based on the topology of the network and available modes for users. It helps the user to select the path for the next day by the experiments of traveling in following day (Smith *et al.* 2014). This study explores the impact of the user learning curve on the unicity of user network equilibrium when the new transportation facility is added to the network. Normally for BRUE, we will have the non-unicity of solutions in each scenario with mono-class users because of the initial conditions (BRUE setting) but for UE we will investigate the equilibrium for both cases mono and multi-class travelers in a multimodal network. For non-unicity of UE, the most likely situation is when we have heterogeneous travel behaviors, e.g. distributed values of time. Because generally multi-class network equilibrium is expressed as a nonmonotone problem (Marcotte and Wynter, 2004). In order to analyze the impact of non-unicity, we are going to work on learning process when the facilities are introduced in the different order during the day-to-day process.

## 2. PROBLEM STATEMENT

Multi-modality permits travelers to use different transportation facilities. Urban public transportation facilities are costly to create and have impacts on traffic network equilibrium. Opening the new metro line or closing some metro stations for maintenance are some examples of changing equilibrium in the network. These events change the access of travelers to some transportation modes. In other words, the path set is changed for several origin-destination (OD) pairs. In case of maintenance process, the transportation network comes back to the initial situation. The question is, do we have the same equilibrium as before? For the second case, if we want to add three metro lines, do we have a same equilibrium with a different order for opening metro lines? Here, we want to answer this question by a numerical experiment on the large-scale network. We are looking for the impact of the opening frequency of three metro lines on unicity of UE and BRUE. We can open three metro lines at the same time and calculate the equilibrium or successively open one metro line every 100 days and look for equilibrium in the day-to-day process. In optimization point of view, it means we change the initial assignment pattern to find the equilibrium. This study is going to be mostly based on simulation, but we are currently investigating also if we can derive the same pattern from analytical derivation. There are four main questions that we are going to figure out in this study:

- Is the network UE/BRUE unique when we have different scenarios for opening the metro lines (different initial point) with homogenous users?
- Do we have unicity in case of UE conditions and multi-class users?
  - we are also investigating analytical derivation to find UE.
- The practical question would be also if there is no unicity, which order has the minimum total travel time? Are we precise enough when we predict total travel time by simulator?

## 3. NUMERICAL EXPERIMENTS

In this work, we use Symuvia as a trip-based simulator for calculating the needed variables in the network. Symuvia has been developed by the LICIT laboratory in IFSTTAR. It is a microscopic simulator based on the Lagrangian resolution of the LWR model (Leclercq *et al.*, 2007). The day-to-day DTA is applied to the large-scale network of Lyon 6e + Villeurbanne with 1,883 Nodes, 3,383 Links, 94 Origins, 227 Destinations and 54,190 trips. Walking, buses and private cars are the initial available transportation modes in the network. There are three metro lines (A, B and C) and 25 metro stations in the network (figure 1).

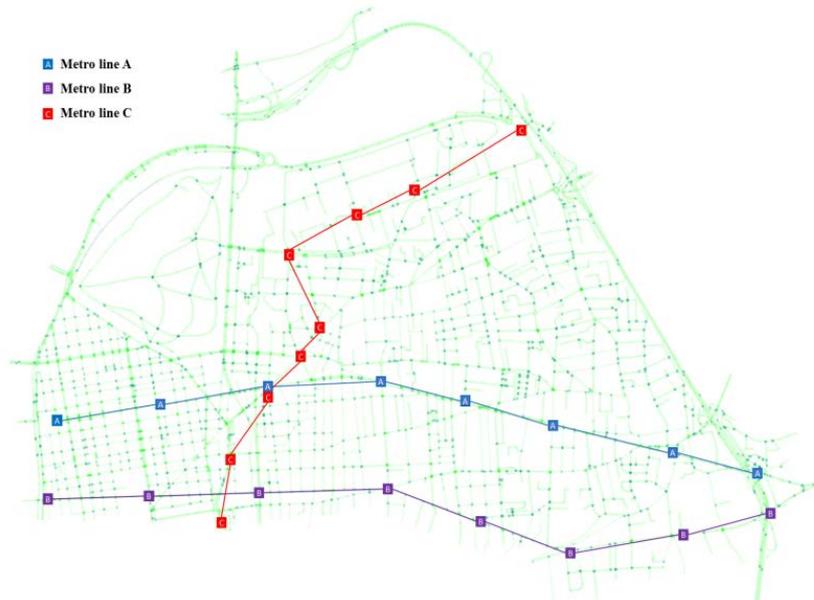


Figure 1. Lyon 6e + Villeurbanne

Each metro station has a parking. Parkings are the connector between the metro grid and traffic network. Therefore, the traveler can start their trip with the private car then use the parking to take the train. All mode changes during the trip have a walking time for connection and possibly a waiting time for the next bus or metro to arrive at the station. There are 7 scenarios to activate the metro lines: open 3 metro lines in the same time and 6 scenarios to open metro lines with the different order (figure 2). For each scenario, we run the day-to-day DTA for 300 days to represent the peak one and half an hour of the network. The experiments are designed for 3 phases. Each phase executes the simulation for 100 days to find the equilibrium (table 1). In the primary results, we allow all travelers to choose between a private car and public transportation and we converged to different UE solutions for 7 scenarios with homogenous travelers. Table 2 shows the number of travelers who use the metro lines in the UE solution for each scenario. Moreover, the total travel time is different for each scenario at the optimal solutions. The results show that not only do we have non-unicity but also, that we can save at least around 150 hours on total travel time by opening the metro lines in the optimal order (ACB). We are going to run the same experiments for BRUE with mono-class and UE with multi-class users in order to analyze the unicity of solutions by evaluating assignment patterns, links flow and paths flow.

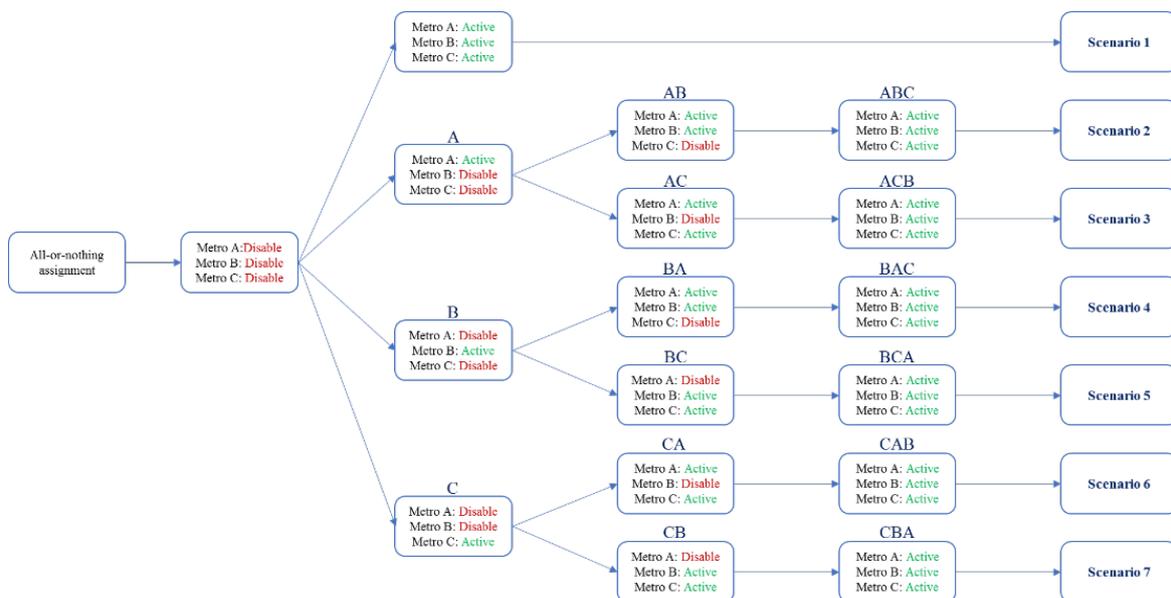


Figure 2. Chart of experiments

Table 1. Experiments description

|                      | # active metro | Sequence | Scenario     | Initial assignment pattern is obtained by (Description) | Optimal assignment path code |
|----------------------|----------------|----------|--------------|---------------------------------------------------------|------------------------------|
| Phase 1<br>(day 0)   | 0              | -        | -            | All-or-nothing assignment                               | P1.1                         |
|                      | 3              | A&B&C    | 1st Scenario | P1.1 (All metro lines are available for the users)      | P1.2                         |
|                      | 1              | A        | -            | P1.1 (Just metro A is available)                        | P1.3                         |
|                      | 1              | B        | -            | P1.1 (Just metro B is available)                        | P1.4                         |
|                      | 1              | C        | -            | P1.1 (Just metro C is available)                        | P1.5                         |
| Phase 2<br>(day 100) | 2              | AB       | -            | Phase 1 simulation code P1.3                            | P2.1                         |
|                      | 2              | AC       | -            | Phase 1 simulation code P1.3                            | P2.2                         |
|                      | 2              | BA       | -            | Phase 1 simulation code P1.4                            | P2.3                         |
|                      | 2              | BC       | -            | Phase 1 simulation code P1.4                            | P2.4                         |
|                      | 2              | CA       | -            | Phase 1 simulation code P1.5                            | P2.5                         |
|                      | 2              | CB       | -            | Phase 1 simulation code P1.5                            | P2.6                         |
| Phase 3<br>(day 200) | 3              | ABC      | 2nd Scenario | Phase 2 simulation code P2.1                            | P3.1                         |
|                      | 3              | ACB      | 3rd Scenario | Phase 2 simulation code P2.2                            | P3.2                         |
|                      | 3              | BAC      | 4th Scenario | Phase 2 simulation code P2.3                            | P3.3                         |
|                      | 3              | BCA      | 5th Scenario | Phase 2 simulation code P2.4                            | P3.4                         |
|                      | 3              | CAB      | 6th Scenario | Phase 2 simulation code P2.5                            | P3.5                         |
|                      | 3              | CBA      | 7th Scenario | Phase 2 simulation code P2.6                            | P3.6                         |

Table 2. Primary results

| Scenario | Sequence | Number of users used Metro line |     |    | Total travel time (hour) |
|----------|----------|---------------------------------|-----|----|--------------------------|
|          |          | A                               | B   | C  |                          |
| 1        | A&B&C    | 480                             | 138 | 9  | 6723.64                  |
| 2        | ABC      | 844                             | 282 | 6  | 6821.81                  |
| 3        | ACB      | 590                             | 132 | 15 | 6574.00                  |
| 4        | BAC      | 560                             | 210 | 84 | 6969.03                  |
| 5        | BCA      | 391                             | 225 | 27 | 6761.89                  |
| 6        | CAB      | 697                             | 168 | 15 | 6729.28                  |
| 7        | CBA      | 633                             | 360 | 9  | 6798.39                  |

## ACKNOWLEDGEMENT

This work is supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program. (Grant agreement No 646592 -- MAGnUM project)

## REFERENCES

- Beckmann, M., McGuire, C. B. and Winsten, C. B. (1956), *Studies in the Economics of Transportation*, No. 226 pp.
- Di, X. and Liu, H. X. (2016), Boundedly rational route choice behavior: A review of models and methodologies, *Transportation Research Part B: Methodological*, 85, 142-179.

- Iryo, T. (2011), Multiple equilibria in a dynamic traffic network, *Transportation Research Part B*, 45 867-879.
- Iryo, T. (2013), Properties of dynamic user equilibrium solution: existence, uniqueness, stability, and robust solution methodology, *Transportmetrica B: Transport Dynamics*, 1(1), 52-67.
- Iryo, T. (2015), Investigating factors for existence of multiple equilibria in dynamic traffic network, *Networks and Spatial Economics*, 15(3), 599- 616.
- Iryo, T. and Smith, M. J. (2017), On the uniqueness of equilibrated dynamic traffic flow pattern in unidirectional networks, *Transportation research procedia*, 23, 283-302.
- Leclercq, L., Laval, J. A. and Chevallier, E. (2007), The Lagrangian coordinates and what it means for first order traffic flow models, In *Transportation and Traffic Theory 2007, Papers Selected for Presentation at ISTTT17*.
- Marcotte, P. and Wynter, L. (2004). A new look at the multiclass network equilibrium problem. *Transportation Science*, 38(3), 282-292.
- Mounce, R., (2001), Non-monotonicity in dynamic traffic assignment networks, In: *Proceedings of the 33rd Universities Transport Study Group Annual Conference. University of Oxford, UK*.
- Mounce, R. (2007), Existence of equilibrium in a continuous dynamic queueing model for traffic networks. In: Heydecker, B.G. (Ed.), *Mathematics in Transport*. Elsevier, Oxford, pp. 219–229.
- Mounce, R. and Smith, M.J. (2007), Uniqueness of equilibrium in steady state and dynamic traffic networks, *Transportation and Traffic Theory. Oxford, Elsevier*, pp. 281–299.
- Peeta, S. and Ziliaskopoulos, A.K. (2001), Foundations of dynamic traffic assignment: The past, the present and the future, *Networks and Spatial Economics*, 1(3), pp.233-265.
- Smith, M. J. and Wisten, M. B. (1995), A continuous day-to-day traffic assignment model and the existence of a continuous dynamic user equilibrium, *Annals of Operations Research*, 60(1), 59-79.
- Smith, M., Hazelton, M. L., Lo, H. K., Cantarella, G. E. and Watling, D. P. (2014). The long term behaviour of day-to-day traffic assignment models. *Transportmetrica A: Transport Science*, 10(7), 647-660.
- Wardrop, J. G. (1952), Road paper. Some theoretical aspect of road traffic research, *Proceedings of the institution of civil engineers*, 1(3), 325-362.

# EFFECTS OF TRAFFIC INFORMATION ON DRIVERS' DAY-TO-DAY ROUTE CHOICES

Genaro PEQUE, Jr<sup>a</sup>, Toshihiko MIYAGI<sup>b</sup> and Fumitaka KURAUCHI<sup>c</sup>

<sup>a</sup> *Department of Civil Engineering,  
Kobe University, Japan*

*Email: [gpequejr@panda.kobe-u.ac.jp](mailto:gpequejr@panda.kobe-u.ac.jp)*

<sup>b, c</sup> *Department of Civil Engineering, Gifu University, Japan*

*Email: [t\\_miyagi@gifu-u.ac.jp](mailto:t_miyagi@gifu-u.ac.jp)*

*Email: [kurauchi@gifu-u.ac.jp](mailto:kurauchi@gifu-u.ac.jp)*

A multi-agent route choice learning model for the day-to-day dynamic traffic assignment is used to investigate the effects of information latency on drivers' route choice decisions. Using the total relative gap convergence metric with a fixed time assignment interval, preliminary results show that shorter information update cycles from a guidance generator produces better convergence compared to longer information update cycles. Moreover, when the periodic information update cycle is very long, it produces a worse result compared to a scenario where drivers rely only on their own travel experiences based on the routes they have chosen.

**Keywords:** ATIS, Day-To-Day Route Choice, Multi-Agent

## 1. INTRODUCTION

Each driver makes a decision based on his knowledge of the available alternative routes and their attributes, subject to time and cognitive capacity constraints. However, a driver's decision relies on the reliability of the traffic information he receives from an Advanced Traveler Information Systems (ATIS) which are intended to assist him to make better travel decisions.

In this paper, we are interested in the effects of information latency on drivers' route choice decisions. Specifically, we want to investigate how driver's route choice learning decisions are affected by different periodic information update cycles from a guidance generator and also compare the worst case (i.e. very long periodic information update cycle) with a scenario where drivers use only their day-to-day experiences.

Following Selten et al.'s experiments (2007) on drivers' route choice behaviors, we also categorize drivers into a.) partially-informed users (PIU) where drivers can acquire (noisy) traffic information for all possible routes to their destination at the end of each day and, b.) naïve users (NU) where drivers can only acquire the traffic information of the route they have chosen on that day.

## 2. THE MODEL

In SUE models, drivers' payoff for alternative  $a \in A$  are normally written as,

$$U(a) = u(a) + \varepsilon(a), \quad (1)$$

where  $A$  is a choice set,  $u(a)$  is a deterministic term and  $\varepsilon(a)$  is a random term assumed to follow some known distribution (e.g. Weibull, Normal, etc.). In this paper, driver  $i$ 's realized payoff for an alternative  $a^i \in A^i, i \in N$  is written as,

$$U^i(a^i) = u^i(a^i) + \varepsilon^i(a^i), \quad (2)$$

Where  $N$  is the set of agents (drivers) and the random term  $\varepsilon^i(a^i)$  in equation (2) is only assumed to have an unknown distribution with zero mean and bounded variance. Since a closed-form expression

to solve equation (2) cannot be derived, it necessitates the use of estimation.

Although the method of successive averages (MSA), based on the classical stochastic approximation theory, was successfully applied to solve the SUE problem through link flow estimation and simulation (Sheffi and Powell, 1982), it is potentially problematic when used as a convergence criterion in a simulation-based DTA as it may only indicate that drivers are satisfied with their routes but may also be an outcome imposed by the algorithm itself and, thus, have nothing to do with travel time. Hence, a more sophisticated approach to solve this problem should be used (Peque et al., 2018).

We assume that drivers use Q-learning to learn the true route travel times. Q-learning is a model-free reinforcement learning technique which ultimately gives the expected payoff of taking a given route in a given state without requiring a model of the environment. Specifically, Q-learning can identify an optimal-selection policy that drivers can follow for route selection, however, Q-learning alone can learn the optimal policy even when actions are selected randomly. Since the drivers might take a lot of random actions over the course of the simulation, the feedback used to update the policy will have a very high variance which means that it will take a lot of policy updates for it to converge. Therefore, there should be mechanism that should be followed to update this policy that will reduce this variance.

In RL, an algorithm with a coupled policy learning and payoff estimation is called an actor-critic algorithm (Sutton and Barto, 1988) where each driver has both an actor component (the policy) and a critic component (payoff estimates) which is used to inform the actor (i.e. update the policy). More formally, an actor-critic algorithm is a process of the form  $\{\pi_t, Q_t\}$  where,

$$\begin{cases} \pi_{t+1}^i(a^i) = \pi_t^i(a^i) + \alpha_{t+1}(\beta_t^i(a^i, Q_t^i)) \\ Q_{t+1}^i(a^i) = Q_t^i(a^i) + \lambda_{t+1}(U_t^i - Q_t^i(a^i)) \end{cases} \quad (3)$$

where  $\pi_{t+1}^i(a^i)$  is the probability of selecting route  $a^i$  on day  $t+1$  by driver  $i$  (i.e. the actor),  $\alpha_{t+1} = \frac{1}{t+1}$  is a learning rate and the parameter  $\beta_t^i$  is the Boltzmann learning policy given by,

$$\beta(a^i) = \frac{\exp\{Q_t^i(a^i) / \mu_t^i\}}{\sum_{\tilde{a}^i} \exp\{Q_t^i(\tilde{a}^i) / \mu_t^i\}}. \quad (4)$$

The  $Q$  value given in equation (3),  $Q_{t+1}^i(a^i)$  is the estimated payoff for route  $a^i$  on day  $t+1$  by driver  $i$  (i.e. the critic) and equation (3) is called the Boltzmann actor-critic algorithm. Equation (3) is used by PIUs where  $\lambda_{t+1} = \frac{1}{t+1}$  since all drivers can acquire (noisy) traffic information at each day for all the routes to their destination and therefore, can always obtain payoff estimates that are asymptotically close to the true route travel times. For NUs however, the actor component is the same but the critic component is changed into,

$$Q_{t+1}^i(a^i) = Q_t^i(a^i) + \mathbb{I}_{\{a_t^i = a^i\}} \tilde{\lambda}_{t+1}^i (U_t^i - Q_t^i(a^i)), \quad (5)$$

where  $\mathbb{I}_{\{a_t^i = a^i\}}$  is the indicator function which is equal to 1 if  $a_t^i = a^i$  and 0, otherwise, and  $\tilde{\lambda}_{t+1}^i = J + \#(a^i)$  where  $J > 0$  and  $\#(a^i)$  is the number of times route  $a^i$  has been selected by driver  $i$  since the driver can only update the  $Q$  value of the route he has chosen on the day he traversed the network. This leaves the problem on how to obtain payoff estimates that are asymptotically close to the true route travel times for NUs. Following Singh et al. (2000), this can be achieved through

the parameter  $\mu_t^i$  in equation (4) by allowing all routes to be selected often enough that estimates remain close. A method that has been studied by Leslie and Collins (2006) called the generalised weakened fictitious play (GWFP) process, which generally converges to Nash equilibrium whenever classical fictitious play does, can be used bounding the probability of selecting any route below by a suitable decreasing sequence, i.e.  $\mu_t \rightarrow 0$  as  $t \rightarrow \infty$ . We use,

$$\mu_t^i = |\bar{Q}_t^i - \bar{U}_t^i|, \quad (6)$$

where  $\bar{Q}_t^i = \bar{Q}_{t-1}^i + \frac{1}{t}(Q_t^i(a^i) - \bar{Q}_{t-1}^i)$  and  $\bar{U}_t^i = \frac{1}{t} \sum_{i=1}^t U_i^i(a^i)$  proposed by Peque et al. (2008), in the route selection model of each driver.

In this model, the  $U^i(a^i)$  is assumed to be the negative value of the travel cost,  $C(a^i)$ , for route  $a^i$ .

### 3. THE TOTAL RELATIVE GAP

For each case (i.e. different periodic information update cycles), we measure the total relative gap to quantify the effect of the scenario setting on drivers' route choice decisions.

The total relative gap is a convergence measure that quantifies how close the solution is to equilibrium. Its measure reflects the summation of the differences between the average route costs and the minimum route cost (route with the highest payoff) in each assignment time interval. More formally, the total relative gap is given by,

$$gap_{rel} = 1 - \frac{\sum_{t_k} \sum_{\omega \in \Omega} F_{\omega}^{t_k} \min_{r \in R_{\omega}} C^{t_k}(r)}{\sum_{t_k} \sum_{\omega \in \Omega} \left( \sum_{r \in R_{\omega}} f^{t_k}(r) C^{t_k}(r) \right)}, \quad (7)$$

where  $t_k$  is the  $k$ th assignment time interval,  $\omega$  is the origin-destination (OD) pair in the OD pair set  $\Omega$ ,  $r$  is a route in the route set,  $R = A^i, \forall i \in I$ , in OD pair  $\omega$ .  $f^{t_k}(r)$  represents the flow on route  $r$  departing at assignment interval  $t_k$ ,  $C^{t_k}(r)$  is the cost (travel time) on used route  $r$  for assignment interval  $t_k$ .  $F_{\omega}^{t_k}$  denotes the total flow for the OD pair at time interval  $t_k$  and  $\min_{r \in R_{\omega}} C^{t_k}(r)$  is the shortest route travel time for OD pair  $\omega$  and assignment time interval  $t_k$ . The intuition of the total relative gap is that if all used routes have travel time very close to the shortest route travel time, then the total relative gap will be close to zero. In most DTA applications, the solution is assumed to have converged to an equilibrium solution when the total relative gap is less than a pre-specified tolerance level.

### 4. SIMULATION SETTING AND PRELIMINARY RESULTS

The simulation-based DTA is carried out by the Simulation of Urban MObility (SUMO) traffic simulator using the Boltzmann actor-critic algorithm given by equation (3) for drivers' route choices in the Sioux Falls network. In most DTA applications, a driver's route set is updated either on a predefined time interval or the previous, current or forecasted network conditions. However, in this simulation it is just assumed time-invariant for simplicity.

In the PIU case, drivers rely on the information sent through an ATIS. This traffic information may be unreliable due to traffic information latency.

For example in figure 1, driver  $i$  entered the network on the first update cycle and exited the network on the third update cycle. Traffic information for all possible routes to the driver's destination during

these three update cycles (green) are averaged, sent to the driver at the end of the day and used by the driver to make a route choice decision for the next day. From figure 1, it is clear that the shorter the update cycle, the more accurate the traffic information becomes (note, that we assume the assignment time interval,  $t_k$ , of the total relative gap is shorter than the update cycle).

In the NU case, it is assumed that drivers cannot get traffic information through an ATIS and, thus, rely on their experiences based on the route they have chosen on the day they traversed the network (e.g. vehicle's total time in the network using the selected route).

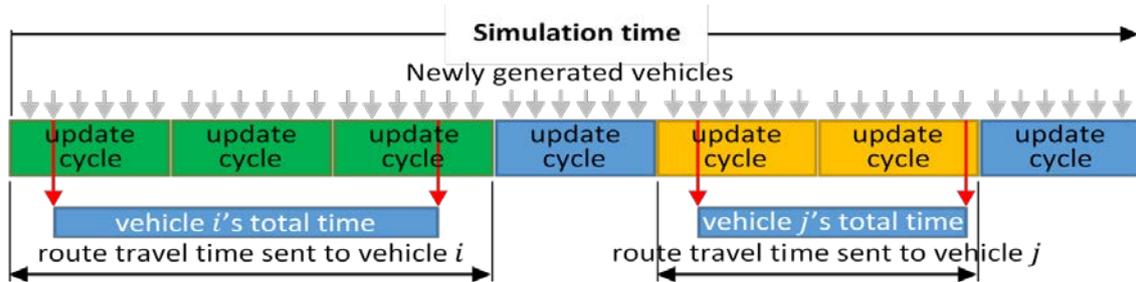


Figure 1. Partially-informed users using previous day's traffic information

Another model we present is where drivers get traffic information from the previous update cycle and use it to make a route choice decision when he/she departs from the current update cycle as shown in the figure below.

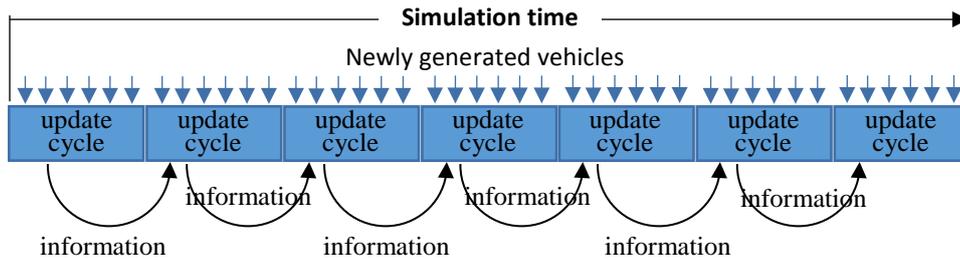


Figure 2. Partially-informed users using previous update cycle's traffic information

Preliminary results show that shorter information update cycles from a guidance generator produces better convergence compared to longer information update cycles. Moreover, when the periodic information update cycle is very long, it produces a worse result compared to a scenario where drivers rely only on their own travel experiences based on the routes they have chosen.

## REFERENCES

- Leslie, D. and Collins, E. (2006) Generalized weakened fictitious play. *Games Econom. Behav.*, 56, pp. 285–298.
- Peque, G., Miyagi, T. & Kurauchi, F. (2018) Adaptive Learning Algorithms for Simulation-Based Dynamic Traffic User Equilibrium, *International Journal of Intelligent Transportation Systems Research*.
- Singh, S.P., Jaakola, T., Littman, M.L., and Szepesvari, C. (2000) Convergence results for single-step on-policy reinforcement-learning algorithm, *Machine Learning*, 38(3), 287-308.
- Selten, R., Chmura, T., Pitz, T., Kube, S., Schreckenberg, M. (2007) Commuters route choice behaviour. *Games Econ. Behav.* 58, 394–406.
- Sheffi, Y, and Powell, W. (1982) An Algorithm for the Equilibrium Assignment Problem with Random Link Times, *Networks*, 12, 191-207.
- Sutton, R. and Barto, A. (1998) *Reinforcement Learning: An Introduction*. MIT Press.

# EFFECT OF INFORMATION COLLECTION BEHAVIOUR ON STABILITY OF A SIMPLIFIED DEPARTURE TIME CHOICE PROBLEM

T. IRYO <sup>a</sup>

<sup>a</sup> *Graduate School of Engineering,  
Kobe University, Japan  
Email: [iryo@kobe-u.ac.jp](mailto:iryo@kobe-u.ac.jp)*

The effect of information collection behaviour on stability of a simplified departure time choice problem is investigated. The instability of the departure time choice problem arises another problem regarding the evolutionary dynamics of this problem – i.e. how the difference in the model of evolutionary dynamics affects the stability of the system? Studies of the instability are very limited in transport literature and consequently there exists no clear answer to this question so far. The present study tries to show how the number of information collections affects the stability of a departure time choice problem by an analytical approach. To maintain mathematical tractability, this study first proposes a simplified departure time choice problem. The degrees of freedom of the proposed model is two and consequently the evolution dynamics can be easily tracked on a two-dimensional plane. It was shown that the increase of the number of information collections lets the system be more stable. This result implies that encouraging information exchange between users is beneficial to mitigate the instability of the system.

**Keywords:** Day-To-Day Dynamics, Evolutionary Dynamic, Departure Time Choice, Information Collection

## 1. INTRODUCTION

The instability of the departure time choice problem, which has been shown by Iryo (2008, 2016a) arises another problem regarding the evolutionary dynamics of this problem – i.e. how the difference in the model of evolutionary dynamics affects the stability of the system? Studies of the instability are very limited in transport literature and consequently there exists no clear answer to this question so far.

Starting from Smith (1984), a number of evolutionary dynamics have been invented to investigate users' adjustment behaviour in a congested transport system. A classification of these models is provided by Sandholm (2010). Further, Iryo (2016b) proposed an evolutionary dynamics that explicitly incorporates a simple form of users' information collection behaviour. In this model, the number of information collections performed at each occasion of revising an alternative was set as a parameter. Then, it was numerically shown that the greater number of information collections lets the system be close to an equilibrium point in a departure time choice problem. While this result implies an important property of the stability (i.e. encouraging information exchanges improve stability of equilibrium), it was only provided as a result of the numerical test.

The present study tries to show how the number of information collections affects the stability of a departure time choice problem by an analytical approach. To maintain mathematical tractability, this study first proposes a 'simplified' departure time choice problem, in which users in two groups select one of two time slots (i.e. users make binary choices). Queuing congestion is approximately formulated considering the temporal externality of the bottleneck model. The degrees of freedom of the proposed model is two and consequently the evolution dynamics can be easily tracked on a two-dimensional plane. In addition, limiting the number of choices to two is beneficial when the information collection behaviour is explicitly incorporated (as shown by Iryo (2016b), all available alternatives must be sorted in the order of their utility, which incurs mathematical difficulty when three or more alternatives exist). To evaluate the dynamics, a gap function that is similar to a Lyapunov function used in Hofbauer and Sandholm (2009) is formulated. Then, behaviour of the gap function is investigated with respect to the number of information collections.

## 2. MODEL

A simplified formulation of the departure time choice problem is proposed. We consider three time slots,

denoted by  $t=1, 2$ , and  $3$ . We also consider two user groups, denoted by  $A$  and  $B$ . The total number of users in each group is the same and set to 1 without loss of generality. Users in group  $A$  prefer time 2 to time 1 and never select time 3 because it is too late. Users in group  $B$  prefer time 3 to time 2 and never selects time 1 because it is too early. The number of users selecting time slot  $t$  is denoted by  $x_{At}$  and  $x_{Bt}$  and the numbers of users selecting time slot  $t$ , denoted by  $n_t$ , are calculated as

$$n_1 = x_{A1}, n_2 = x_{A2} + x_{B2}, n_3 = x_{B3}. \quad (1)$$

Travel utility of time slot  $t$  for group  $i$ , denoted by  $u_{it}$  is defined by

$$u_{A1} = -(w_1 + a), u_{A2} = -w_2, u_{B2} = -(w_2 + b), u_{B3} = -w_3 \quad (2)$$

where

$$w_1 = \gamma n_1, w_2 = n_1 + \gamma n_2, w_3 = n_1 + n_2 + \gamma n_3 \quad (3)$$

is a common part of travel cost depending on  $n_t$ , and  $a$ ,  $b$ , and  $\gamma$  are positive constants. Considering the temporal externality of the queueing, we set  $\gamma < 1$  so that the increase of the number of users in a time slot has more impact onto the successive time slots than onto the current time slot. Equations (1) – (3) derives

$$\begin{aligned} u_A &= u_{A1} - u_{A2} = (1-2\gamma)x_A + \gamma x_B + \gamma - a \\ u_B &= u_{B2} - u_{B3} = (\gamma-1)x_A + (1-2\gamma)x_B + 1 - b \end{aligned} \quad (4)$$

where  $u_A$  and  $u_B$  denote the relative utility of the earlier time slots and  $x_A = x_{A1}$  and  $x_B = x_{B2}$ . We also let  $\mathbf{x} = (x_A, x_B)$ . Note that  $x_{A2} = 1 - x_A$  and  $x_{B3} = 1 - x_B$ . Jacobian of  $\mathbf{u} = (u_A, u_B)$ , denoted by  $J$ , is

$$J = \begin{pmatrix} 1-2\gamma & \gamma \\ \gamma-1 & 1-2\gamma \end{pmatrix}. \quad (5)$$

If  $\gamma < 0.5$  (i.e. the temporal externality of the queueing is twice or more greater than its internal effect in the same time slot), the eigenvalues of  $J$  has the positive real part. We will use such  $\gamma$  afterwards.

The evolutionary dynamics proposed by Iryo (2016b) is used with minor modifications for binary-choice problems. The model incorporates users' information collection behaviour from other users. In the model, each user occasionally collect information from other users to revise his/her current choice. At each occasion of an information collection, he/she communicate  $n$  other users (referred to as *friends* here), where  $n$  is a positive constant. Any friend can only provide information on the utility of the alternative that he/she is currently selecting. If, at least, one of  $n$  friends currently selects the alternative whose utility is highest, the user finds it and may start selecting it. The likelihood of moving to the new alternative is proportional to the difference of utility between two alternatives. Otherwise, the user continue selecting the alternative currently selected. When the proportion of the users selecting the highest utility is  $x$ , the probability of *not* finding such a user is  $(1-x)^n$ . The evolutionary dynamics in this setting can be written as

$$\dot{x}_i = p_i^+ \pi_i^+ + p_i^- \pi_i^- \quad (6)$$

where the dot above  $x_i$  indicates the derivative with respect to date (which is described by a continuous number),

$$\pi_i^+ = \begin{cases} u_i & \text{if } u_i \geq 0 \\ 0 & \text{otherwise} \end{cases}, \quad \pi_i^- = \begin{cases} u_i & \text{if } u_i \leq 0 \\ 0 & \text{otherwise} \end{cases}, \quad (7)$$

$$p_i^+ = (1-x_i)\{1-(1-x_i)^n\}, p_i^- = x_i(1-x_i^n) \quad (8)$$

and  $i = \{A, B\}$ . We also use their derivatives such as

$$q_i^+ = \frac{\partial p_i^+}{\partial x_i}, q_i^- = \frac{\partial p_i^-}{\partial x_i}. \quad (9)$$

### 3. EVALUATING BEHAVIOUR OF EVOLUTION DYNAMICS

We use the gap function for the interior equilibrium solution (i.e. the equilibrium point satisfying  $0 < x_i < 1$  for  $i = \{A, B\}$ ), which is defined by

$$D = \sum_{i \in \{A,B\}} p_i^+ (\pi_i^+)^2 + p_i^- (\pi_i^-)^2. \quad (10)$$

$D$  becomes zero if and only if  $\mathbf{x}$  is an equilibrium solution when  $0 < x_i < 1$  for  $i = \{A, B\}$ . To investigate the derivative with respect to the date, the partial derivatives of  $D$  are calculated as follows:

$$\frac{\partial D}{\partial x_i} = \left\{ q_i^+ (\pi_i^+)^2 + q_i^- (\pi_i^-)^2 \right\} + 2 \sum_{j \in \{A,B\}} \frac{\partial u_j}{\partial x_i} \left\{ p_j^+ \pi_j^+ + p_j^- \pi_j^- \right\} = \left\{ q_i^+ (\pi_i^+)^2 + q_i^- (\pi_i^-)^2 \right\} + 2 \sum_{j \in \{A,B\}} \frac{\partial u_j}{\partial x_i} \dot{x}_j. \quad (11)$$

In addition,

$$\left\{ q_i^+ (\pi_i^+)^2 + q_i^- (\pi_i^-)^2 \right\} \dot{x}_i = p_i^+ q_i^+ (\pi_i^+)^3 + p_i^- q_i^- (\pi_i^-)^3. \quad (12)$$

Note that the cross term  $\pi_i^+ \pi_i^-$  is always zero. Finally, we obtain

$$\dot{D} = \sum_{i \in \{A,B\}} \frac{\partial D}{\partial x_i} \dot{x}_i = \sum_{i \in \{A,B\}} \left\{ p_i^+ q_i^+ (\pi_i^+)^3 + p_i^- q_i^- (\pi_i^-)^3 \right\} + {}^t \dot{\mathbf{x}} J \dot{\mathbf{x}}. \quad (13)$$

The second term can be calculated as

$${}^t \dot{\mathbf{x}} J \dot{\mathbf{x}} = {}^t \dot{\mathbf{x}} \begin{pmatrix} 1-2\gamma & \gamma \\ \gamma-1 & 1-2\gamma \end{pmatrix} \dot{\mathbf{x}} = {}^t \dot{\mathbf{x}} \begin{pmatrix} 1-2\gamma & \gamma-0.5 \\ \gamma-0.5 & 1-2\gamma \end{pmatrix} \dot{\mathbf{x}}. \quad (14)$$

Owing to  $\gamma < 0.5$ , the second matrix is positive definitive at any time, and hence this term is always positive. This implies that the system cannot converge to the equilibrium point. Note that the first term of Equation (13) is proportional to  $(\pi_i^+)^3$  or  $(\pi_i^-)^3$ , while its second term is proportional to  $(\pi_i^+)^2$  or  $(\pi_i^-)^2$ , implying that the first term is negligible when the system is close to the equilibrium point.

To evaluate the behaviour of  $D$  with respect to  $n$ , we evaluate the first term of Equation (13). This can be rewritten as

$$\sum_{i \in \{A,B\}} \left\{ p_i^+ q_i^+ (\pi_i^+)^3 + p_i^- q_i^- (\pi_i^-)^3 \right\} = \frac{1}{2} \sum_{i \in \{A,B\}} \left\{ (\pi_i^+)^3 \frac{\partial}{\partial x_i} p_i^{+2} + (\pi_i^-)^3 \frac{\partial}{\partial x_i} p_i^{-2} \right\}. \quad (15)$$

Note that  $(\pi_i^+)^3 \geq 0$  and  $(\pi_i^-)^3 \leq 0$ . Figure 1 shows the graph of  $(p_i^+)^2$  for  $n = 1, 50$  ( $(p_i^-)^2$  can be obtained by horizontally flipping the graph). When  $n$  is larger, unless  $x_i$  is close to 0 or 1,  $(p_i^+)^2$  is monotone decreasing and  $(p_i^-)^2$  is monotone increasing with respect to  $x_i$  and hence Equation (15) is negative. On the other hand, if  $n$  is smaller,  $(p_i^+)^2$  and  $(p_i^-)^2$  are not monotonic and consequently Equation (15) can have a positive number. Therefore, we can conclude that the greater number of  $n$  tends to let  $D$  be smaller and consequently let the system be closer to the equilibrium point. Figure 2 shows numerical examples of two cases ( $n = 2, 10$ ), where  $\gamma = 0.45$ ,  $a = 0.725$ , and  $b = 0.775$ . The trajectory in each case diverges from the initial point, which is near to the equilibrium point ( $x_A = 0.5$  and  $x_B = 0.5$ ), and converges to an orbit. Comparing two cases, we can see that the size of the orbit is smaller when  $n$  is greater.

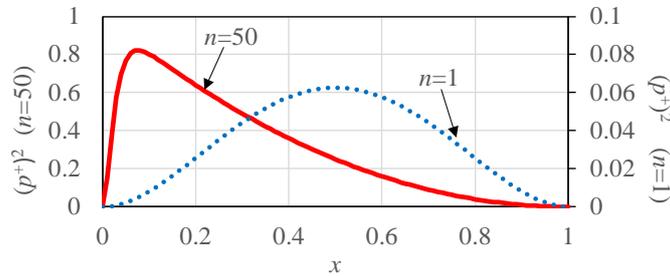


Figure 1. Graph of  $(p_i^+)^2$

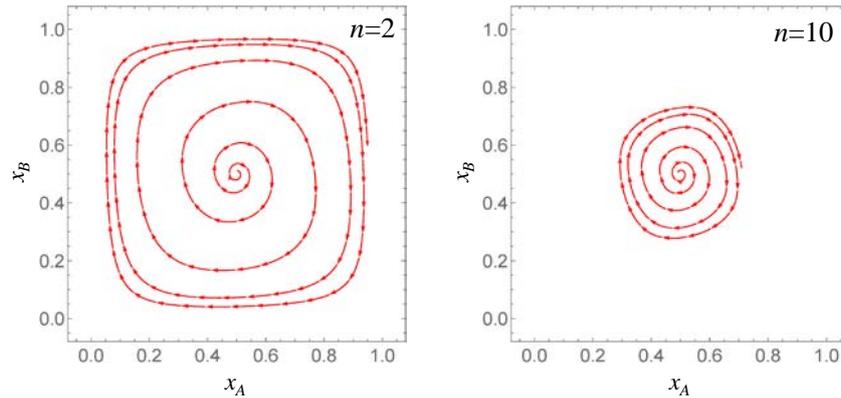


Figure 2. Numerically calculated solution trajectory (initial point:  $(x_A, x_B) = (0.45, 0.45)$ )

#### 4. CONCLUSIONS AND FUTURE TASKS

This study investigated the stability of the simplified departure time choice problem with the evolutionary dynamics explicitly considering the users' information collection behaviour. It was shown that the increase of the number of information collections lets the system be more stable, i.e. lets the vector  $\mathbf{x}$  be more close to the equilibrium point. This result implies that encouraging information exchange between users is beneficial to mitigate the instability of the system. Although this result has been indicated by a numerical test (Iryo, 2016), a mathematical analysis has been provided in the proposed study thanks to the simplified departure time choice problem. The simplified departure time choice problem proposed in this study consists of simple linear equations, which is easy to be analysed. The model successfully described the unstable feature of the departure time choice problem with only two variables. It would be also useful to other analytical investigations of this problem.

A few tasks are left. Equation (13) should be further investigated by substituting Equation (6) to it. It will provide more quantitative properties of the dynamics. While the proposed linear model was easy to handle, a non-linear version of the simplified departure time choice problem would be useful to describe properties of this problem more precisely. Analysing a model with higher degrees of freedom is more attractive because it should prove more precise result, while it is still a challenging task when the information collection behaviour is explicitly considered.

#### ACKNOWLEDGEMENTS

This study was financially supported by JSPS Grant-in-Aid #16H02368.

#### REFERENCES

- Hofbauer, J., and Sandholm, W. H. (2009) Stable games and their dynamics. *Journal of Economic Theory*, 144(4), 1665–1693.
- Iryo, T. (2008) An analysis of instability in a departure time choice problem. *Journal of Advanced Transportation*, 42(3), 333–356.
- Iryo, T. (2016a) Instability of departure time choice problem: a case with replicator dynamics, *6th DTA symposium*, Sydney, Australia.
- Iryo, T. (2016b) Day-to-day dynamical model incorporating an explicit description of individuals' information collection behaviour. *Transportation Research Part B*, 92, 88–103.
- Sandholm, W. H. (2010). *Population Games and Evolutionary Dynamics*. Cambridge: MIT Press.
- Smith, M. J. (1984) The stability of a dynamic model of traffic assignment - an application of a method of Lyapunov. *Transportation Science*, 18(3), 245–252.

**Session B3**  
**Behavior Modeling**

# MODELING THE USERS DEPARTURE TIME CHOICE ADAPTION FOR A TOLLING SYSTEM

B. BRACHER<sup>a</sup> and K. BOGENBERGER<sup>b</sup>  
*Institute for Intelligent Transportation Systems,  
Bundeswehr University Munich, Germany*  
<sup>a</sup> Email: [benedikt.bracher@unibw.de](mailto:benedikt.bracher@unibw.de)  
<sup>b</sup> Email: [klaus.bogenberger@unibw.de](mailto:klaus.bogenberger@unibw.de)

This work aims to show a way to simulate the departure-time choice behavior incorporating changed mobility costs, caused, for example, by a dynamic toll. This reflects the learned behavior of users reacting to a toll known from a previous day. One of the core demands to the whole method is that the original O-D-matrices are not aggregated and distributed during the process, thus preserving the valuable information stored within. A logit procedure is shown that compares varied parameters, which results in a realistic representation of the effects of a toll onto the departure-time choice behavior. In addition, the different sensitivities of the users for earlier and later departure times can be incorporated in the calculation and be reflected in the results.

**Keywords:** Departure Time Choice, Tolling System, Logit, User Adaption

## 1. INTRODUCTION

For traffic simulation, knowledge of the traffic relations within the assessed network is fundamental as they build the foundation for every simulation and reflect the traffic load that has to be managed by the road network. These traffic relations are represented by origin-destination-(O-D)-matrices, representing the number of trips starting at one origin and leading to one destination over the course of an average day. The realized trips on the road network, and therefore the O-D-matrices, are the result of several decisions people make, who have a desire to change their location. It is infeasible to measure this desire directly. Several estimation techniques have been developed in order to enable traffic planners to estimate the demand in the network and its modal and temporal distribution. This is still a field of widespread research. In addition, there are several different procedures to forecast traffic demand.

In order to produce realistic O-D-matrices, especially in presence of a tolling system, understanding the choice behavior of road users is helpful, and therefore a wide field of research. Modal, temporal and route choice behavior have to be considered.

To correctly understand and simulate the temporal choice behavior, several studies were conducted that examined behavioral data and traffic patterns of citizens. (Jackson & Jucker, 1981) created a survey study to identify the risk adversity and time variability of road users. Based on this research (Senna, 1994) implemented an extension to this model, adding more calibration parameters. (Iida, et al., 1992) expanded the field of knowledge regarding the influence of information onto choice behavior of road users in another study.

Several other papers pursue the idea of building decision models, mostly using the logit-formulation, and calibrate several parameters in order to recreate the temporal distribution of traffic observed in reality. (Abkowitz, 1981) designed a model representing the behavior of 425 test people by optimizing the utility of their modal and temporal choices. Another significant survey study was done by (Gunn, et al., 1999), using 15 different parameters to recreate the choice behavior of over 5000 questioned subjects. (Mahassani & Chang, 1986) developed a learning algorithm, modifying the departure time of users depending on the previous traffic occurrence. (Noland & Small, 1995) formulated a mathematical environment to recreate the temporal choice behavior in traffic adding stochasticity to the parameters. All these models are capable of recreating a realistic departure time distribution by spreading a given amount of traffic. However, they are not able to analyze an existing set of O-D-matrices and modify them considering a varying and temporally limited change of utility, as it is caused by a dynamic toll.

The subject of this paper is to develop a technique that can be used to quantify the effects of such a toll,

taking into account existing O-D-matrices and a known temporal progress of the toll height. This can be used for example to estimate the effects of a dynamic, traffic-dependent toll. As the modal and temporal choice decisions reflect long-time and learned behavior, it is necessary to calculate these effects in advance of each simulation run, forming an iterative method. The approach pursued in this paper will be described shortly on the next pages. The developed method is designed to work in conjunction with a microsimulation environment and a modal choice simulation in order to simulate the control of a dynamic area toll and determine the traffic effects caused by such a toll. We discussed our framework thoroughly in (Bracher & Bogenberger, 2018).

## 2. METHODOLOGY

As mentioned in the introduction, the subject of this paper is the development of a new way to estimate the temporal shift caused by a dynamic or static toll. In addition, it is possible to address various other effects that change the utility of travel. Since the information stored in the original O-D-matrices for a given simulation network is seen as the optimal solution for the initial temporal distribution, it shall not be changed if there is no need to do so. This requires a methodology that recreates the temporal choice behavior of the users and identifies the amount of changes induced by a variable time-dependent toll. The toll is seen as a given set of known, quasi-historic data, helping to adjust the O-D-matrices for the next simulation run, taking into account the changed conditions, and emulating the learning behavior of the users.

As this paper deals with modifying an existing O-D-demand, there is a given amount of drivers  $F_{O,D,z}$  for every time slice  $z$  of the demand matrix. The time interval for assessing the departure-time choice behavior is set to 5 minutes. A shorter interval is seen as irrelevant, as users will not schedule their trips more precisely, whereas a longer interval, e.g. 15 minutes, reduces the accuracy of the calculation and therefore the reliability of the results disproportionately. It is therefore necessary that the O-D-matrices are given in the same interval, or are split evenly into such time slices.

We assume that users choose their departure time in a time window of one hour. This time window is divided in time intervals, denoted by  $t \in [-6; 6]$ , with values  $S_t = [-30, -25, -20, -15, -10, -5, 0, 5, 10, 15, 20, 25, 30]$ .

As the change of time choice induced by a toll shall be assessed, a theoretical time choice distribution has to be calculated for all  $F_{O,D,z}$  for the whole range of  $t$ , serving as a reference value. This initial distribution is subsequently compared to the toll-influenced choice distribution. Assuming that the given temporal distribution already depicts the optimal departure-time solution for every road user, going earlier or later is directly connected with an additional cost, usually time-, money-, or comfort-wise. As the sensitivity for departing earlier is different from departing later, setting two different cost factors  $C_{early}$  and  $C_{late}$  is common practice (Mandir, 2012; Noland & Small, 1995). The different kinds of cost are hereby joined into one global cost variable, representing the time sensibility of the users measured in pecuniary units. The abstract utility of going earlier or later is calculated by:

$$U_t^{ref} = |S_t| * C_{Time,t} \quad \text{with } C_{Time,t} = \begin{cases} C_{early} & \text{if } t < 0 \\ -1 & \text{if } t = 0 \\ C_{late} & \text{if } t > 0 \end{cases} \quad (1)$$

The cost factors  $C_{early}$  and  $C_{late}$  are negative, as shifting the departure time reduces the utility for each user, whereas keeping departure time ( $t = 0$ ) gives an unchanged utility value. Based on this utility, the reference probability distribution  $P_t^{ref}$  for choosing a departure time slot  $t$  can be calculated using the logit-function

$$P_t^{ref} = \frac{e^{U_t^{ref}}}{\sum_t e^{U_t^{ref}}} \quad (2)$$

This distribution is valid for each O-D-combination and time-slice  $z$ . To calculate the utilities  $U_{O,D,z,t}$  with a toll present, the overall cost increase due to the toll has to be calculated. The used method

is explained in full detail in (Bracher & Bogenberger, 2018). The cost increase  $C_{inc,O,D,z}$  caused by the toll for each time interval is given as a percent value, normalized to the average mobility cost for a given O-D-connection. This value is calculated by analyzing the realized trajectories in a microsimulative environment, incorporating the experienced toll of the respective trajectory. The proposed technique can also be performed using real historic toll data. The utility is influenced by the toll as follows:

$$U_{O,D,z,t} = U_t^{ref} + U_{O,D,z,t}^{toll} \quad (3)$$

As the given cost increase  $C_{inc,O,D,z}$  denotes the average increase of overall transportation cost, it will also affect the utilities for the respective time interval  $t$  by further reducing the utility of the interval. Using a cost factor  $C_{toll}$  the sensitivity of users reacting to the toll can be calibrated, incorporating the share of users unable or unwilling to change their departure time and the influence of the amount of non-pecuniary utility that is not affected by the cost increase  $C_{inc,O,D,z}$ . As the sensitivity for the cost increase is dependent on the direction and the extent of the time shift, the cost increase is also multiplied by  $C_{time}$  to scale it appropriately. It is equally important that for  $t \neq 0$  the cost increase  $C_{inc,O,D,z+t}$  of the relevant time slice the departure-time shift takes place in, is considered. With  $U_{O,D,z,t}^{toll} = - C_{time} * C_{toll} * C_{inc,O,D,z+t}$  it follows:

$$U_{O,D,z,t} = |S_t| * C_{time} - C_{inc,O,D,z+t} * C_{toll} * C_{time} \quad (4)$$

Again, the choice probability is calculated using the logit formulation:

$$P_{O,D,z,t} = \frac{e^{U_{O,D,z,t}}}{\sum_t e^{U_{O,D,z,t}}} \quad (5)$$

As the departure-time change induced by the toll is the value of interest in this paper, it is calculated by

$$P_{O,D,z,t}^{new} = P_{O,D,z,t} - P_t^{ref} \quad (6)$$

Using this new departure-time probability, the number of users of  $F_{O,D,z,t}$  changing their departure-time by interval  $S_t$  can be calculated. It is worth mentioning that the calculation for the users for  $t = 0$  differs from the other intervals, since the amount of users staying in the interval is relevant.

$$F_{O,D,z,t} = \begin{cases} F_{O,D,z,t} * P_{O,D,z,t}^{new} & \text{for } t \neq 0 \\ F_{O,D,z,t} * (1 + P_{O,D,z,t}^{new}) & \text{for } t = 0 \end{cases} \quad (7)$$

Finally, the users have to be aggregated, allocating the users in the time intervals  $t$  to the correct time slices  $z$

$$F_{O,D,z}^{new} = \sum_t F_{O,D,z+t} \quad (8)$$

If applicable, these new numbers of users now can be re-aggregated into the original time slices.

As mentioned above, the developed departure-time choice model is meant to be integrated into a higher-level simulative environment. The layout and functionality is based on the layout described in (Bracher & Bogenberger, 2018).

### 3. RESULTS

To assess the functionality of the method depicted, the departure time shifts induced by different toll cost factors were calculated. A selection of the results is shown in Figure 1. The percentual reduction of the departures at each time slice is depicted by the solid line referring to the primary axis, whereas the increase of the average transportation cost caused by the toll in percent is depicted by the dotted line, referring to the secondary axes. Both are printed against time on the horizontal axis. Looking first at Figure 1a), it can be seen that a rise of the overall transportation cost by 10 % for 15 minutes results in a reduction of 4 % in departures at this time, equivalent to a demand elasticity of -0.4. This is in the range of demand elasticities found in literature (e.g. Jong & Hugh, 2001, Oum, et al., 1992). In addition, the demand shifts to earlier times, as intended and supported by literature. Figure 1b) shows that in case

of an incremental rise of the toll cost, the users in times of lower toll do not adjust their departure times to a later time slice with a higher toll as this does not prove to have benefits. The users of the time slices with the higher cost increase, however, switch to times with lower toll, as this provides benefits, although the change is not as high as without a toll present. This shows the correct functionality of the developed method calculating the shift at time slices with deviating neighboring tolls, which is a crucial point. Figure 1c) shows the time shift behavior with a very high cost  $C_{late}$ , synonymous to no possibility of the users to be late. As expected, the whole demand shifts to an earlier time, with almost no shift to later departure times visible. This again shows the correct functionality of the proposed method incorporating different cost weights for earliness or lateness.

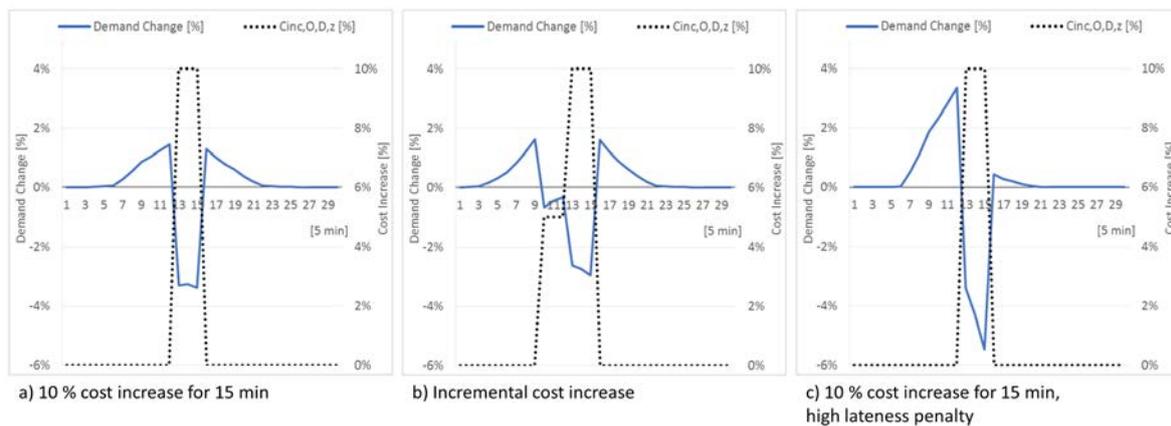


Figure 8: Coherence of toll cost and departure time choice

Furthermore, the initial departure-time information stored in the O-D-matrices is not changed through the process in case where there is no toll present, but the quantity of users at each time slice is redistributed according to the changed conditions under presence of a toll, as it was one of the main requirements of the whole method.

#### 4. CONCLUSION

This work aims to show a way to simulate the departure-time choice behavior incorporating changed mobility costs, caused, for example, by a dynamic toll. One of the core demands to the whole method was that the original O-D-matrices are not aggregated and distributed during the process, thus preserving the valuable information stored within. Developing a logit procedure comparing varied parameters was the goal, which was achieved with a realistic representation of the effects of a toll onto the departure-time choice behavior. Also the different sensitivities of the users for earlier and later departure times can be incorporated in the calculation and be reflected in the results.

Future work consists of applying the developed method in a simulation environment with real network data to assess the overall traffic effects of a dynamic area toll.

#### REFERENCES

- Abkowitz, M. D., 1981. An Analysis of the Commuter Departure Time Decision. *Transportation*, Volume 10, pp. 283 - 297.
- Bracher, B. & Bogenberger, K., 2018. *Modelling the long-term modal effects of a dynamic congestion charging zone*. Washington, accepted at TRB 2018.
- Gunn, H., Tuinenga, J., Cheung, Y. & Kleijn, H., 1999. Value of Dutch Travel Time Savings in 1997. *Planning, operation, management and control*, pp. 513 - 526.
- Iida, Y., Takamasa, A. & Takashi, U., 1992. Experimental Analysis of dynamic Route Choice Behaviour. *Transportation Research Part B*, Volume Vol. 26B, No. 1, pp. 17 - 32.
- Jackson, B. W. & Jucker, J. V., 1981. *An Empirical Study of Travel Time Variability and Travel Choice behaviour*. s.l., Operations Research Society of America, pp. 460 - 475.

- Jong, G. d. & Hugh, G., 2001. Recent Evidence on Car Cost and Time Elasticities of Travel Demand in Europe. *Journal of Transport Economics and Policy*, May, 35(2), pp. 137-160.
- Mahassani, H. S. & Chang, G.-L., 1986. Experiments with Departure Time Choice Dynamics of Urban Commuters. *Transportation Research Part B*, Volume Vol. 20B, No. 4, pp. 297 - 320.
- Mandir, E., 2012. *Potential of Traffic Information to Optimize Route and Departure Time Choice*, Stuttgart: Institut für Straßen- und Verkehrswesen, Universität Stuttgart.
- Noland, R. B. & Small, K. A., 1995. *Travel-Time Uncertainty, Departure Time Choice, and the Cost of the Morning Commute*, University of California, Irvine: Department of Economics and Institute of Transportation Studies.
- Oum, T. H., Waters, W. & Yong, J.-S., 1992. Concepts of Price Elasticities of Transport Demand and Recent Empirical Estimates. *Journal of Transport Economics and Policy*, May, 26(2), pp. 139-154.
- Senna, L. A., 1994. *The influence of travel time variability on the value of time*, Netherlands: Kluwer Academic Publishers.

# EMPIRICAL ANALYSIS OF DETOURS ON FREEWAYS IN UK, FRANCE AND GERMANY

M. AUER<sup>a</sup>, H. REHBORN<sup>a</sup> and K. BOGENBERGER<sup>b</sup>

<sup>a</sup> Daimler AG, Research & Development, Germany

Email: [markus.m.auer@daimler.com](mailto:markus.m.auer@daimler.com), [hubert.rehborn@daimler.com](mailto:hubert.rehborn@daimler.com)

<sup>b</sup> Chair of Traffic Engineering, University of Federal Armed Forces Munich, Germany

Email: [klaus.bogenberger@unibw.de](mailto:klaus.bogenberger@unibw.de)

Congestion on freeways is a widespread problem in traffic networks and results in lots of wasted time. While congestion is not avoidable overall, an optimization of individual travel times is possible by intelligent route choice. In case of congestion on the main route the driver can either stay on the freeway or take an alternate route. In this paper the route choice behavior of drivers in Germany, United Kingdom and France is evaluated for the freeways A8, M1 and A1, respectively. The trajectory data for the analysis is obtained by a floating car data system which is used for the generation of live traffic information services. The travel times are measured for main and alternate route and a corresponding predicted travel time is computed based on logged live traffic information of the respective days. The empirical analysis shows that on average drivers can achieve slightly shorter travel times by bypassing congested freeways. The frequency and travel time advantage of the alternate routes in Germany and United Kingdom are similar on the analyzed freeways. In contrast, the route choice behavior is different on the freeway in France due to the toll on the A1.

**Keywords:** Dynamic Vehicle Routing, Empirical Traffic Data Analysis, Floating Car Data, Connected Vehicles, Route Choice Decision

## 1. INTRODUCTION

Driving on a freeway is trouble free without severe congestion. But often drivers suffer from delays due to accidents or congestion on the freeway. With live traffic information available in navigation devices drivers are able to adopt their route according to the current traffic network state. The influence of traffic information on the route choice behavior of travelers has been analyzed intensively in the past years (see overview by Chorus et al., 2006) and the research can be grouped into stated preference (SP) and revealed preference (RP) experiments (Polydoropoulou et al., 1994). Moreover, uncertainty in transportation systems heavily influences the route choices of travelers and the resulting travel times as depicted by Bonsall (2004) and Han et al. (2008). Iida et al. (1992) and Selten et al. (2007) investigated the route choice behavior of drivers in iteration-based, SP laboratory experiments using a simple traffic network with two parallel alternative routes. They showed that the travel times are fluctuating over a long time period and do not converge to equilibrium. Mahmassani and Liu (1999) include a dynamic interactive traffic simulation into the SP experiment and so the route choices of the participants influence the simulated travel conditions in real time.

In the past RP studies were based only on diary surveys due to the lack of other available data. More recently Schlaich (2010) generated trajectory data from mobile phone data and evaluated the route choice behavior in a subpart of the German freeway network spanned up by the cities Stuttgart Heilbronn, Mannheim and Karlsruhe. Tiratanapakhom et al. (2014) used traffic data from an electronic toll collection system to analyze the route choice behavior. In contrast to earlier research works, Auer et al. (2017) observed the route choices of drivers on a German freeway by analyzing real world vehicle trajectory data from a floating car data (FCD) production system. The focus of the research is the observation and evaluation of travel time benefits by route switching in case of congestion and not the development of route choice models based on explanatory variables. In addition to the previous work, traffic information available in the vehicle navigation system is taken into account in this paper and the observation of route switching behavior is expanded to road networks in France and United Kingdom.

## 2. EXPERIMENTAL SETUP AND DATA BASIS

The empirical observations in this paper are based on trajectory data from a large vehicle fleet in Europe. The outstanding feature of this data set is the accuracy and completeness of the trajectories compared to routes obtained by travel diaries. All vehicles are equipped with navigation devices which receive live travel information. In case of a longer delay on the congested freeway and a predicted faster alternate route, the navigation system will suggest and guide to an expected faster alternate route in the secondary network. The predicted and real travel time differences by route switching between main and alternate routes are determined and analyzed in this paper. In order to maintain the privacy of the drivers, all trajectories are anonymized and thus no conclusions about vehicle, driver or purpose of the trip are possible. Hence, it is also not possible to complement the study by a survey as there is neither a technical nor a legal opportunity to contact the drivers. The data basis consists on the one side of the vehicle trajectories obtained from the FCD system and on the other side of the corresponding traffic service messages available in the vehicles. The anonymized trajectory data contains timestamp, latitude and longitude and has a sampling rate of 5 to 10 seconds in this specific case study. With this sampling rate and precision of the GPS signal the exact routes of the drivers can be reconstructed road by road.

Table 1. Recorded trajectories per day and direction depending on region and time period

|                                    | Germany (A8) | United Kingdom (M1) | France (A1) |
|------------------------------------|--------------|---------------------|-------------|
| 1st/2nd/5th – 19th October 2015    | ~5000        | -                   | -           |
| 25th November – 22th December 2015 | ~6000        | -                   | -           |
| 1st – 11th April 2016              | ~7000        | ~4500               | -           |
| 25th October – 22th November 2016  | ~12500       | ~10500              | ~3000       |

The FCD system generates million trajectories per day consisting of billions of GPS points in Europe. The analyzed regions and time periods are given in Table 1. After matching the trajectory data on the road network, empirical travel times and average velocities can be computed for arbitrary road stretches. The live traffic information send to the vehicles is recorded and so the travel time predictions available in the vehicles can be reconstructed. Throughout our analysis we focus on three freeways in Germany, United Kingdom and France, and their corresponding alternate routes as given in Table 2.

Table 2. Available trajectory data depending on region and time period

|                   | Start   | End       | Length (km)     | On- & off-ramps |
|-------------------|---------|-----------|-----------------|-----------------|
| Germany A8        | München | Karlsruhe | 266             | 44              |
| United Kingdom M1 | London  | Leeds     | 296             | 65              |
| France A1         | Paris   | Lille     | 208 (toll: 129) | 31              |

The freeways are selected such that the results are in principle comparable to each other. All freeways connect major metropolitan areas in the respective countries and are often very congested. In case of congestion drivers can decide to take an alternative route in the secondary road network as there are no other freeways in the proximity suitable for detours. The freeway A8 in Germany connects the city of Munich with the metropolitan area of Stuttgart. Furthermore, the road passes several cities like Ulm and Augsburg. In the European route network the A8 is part of the route E45, which leads from Strasbourg in France to Salzburg in Austria. Due to the high traffic flows, the freeway is very often severely congested especially in the area of Stuttgart. There is no toll payable on any part of the road. As with every German highway, there are officially marked detours between subsequent on- and off-ramps. These detours can be used in case of large delays on the freeway. The permanent detour sign is sign number 460 in the German StVO and consists of a white text (“U” for German “Umleitung” followed by a number) on a blue background. Similarly, the M1 motorway in the United Kingdom connects the metropolitan area of London with Sheffield and Leeds. The M1 coincides with the route E13 of the European road network. Also there is no toll payable and the freeway is often heavily congested. In addition, in United Kingdom there exist official emergency diversion routes marked by simple geometric shapes like open and filled circles, squares, and diamonds. The freeway A1 in France is in contrast to the other chosen freeways on large segments a toll road. Due to the toll additional time is required for

access and exit of the freeway. In case of an accident on the freeway and resulting time delays, the probability of a useful detour is lower than on non-toll roads. The freeway A1 connects Paris with Lille and is part of the European routes E17 and E19. The average length between two subsequent interchanges varies between 4.5 km for the M1, 6.0 km for the A8 and 6.7 km for the A1. So for the toll road in France the average distance between interchanges is greater than for the other two freeways, but still very close to the others. As the possible detours are selected by the drivers and can vary from driver to driver, it is difficult to determine average lengths or travel time for the detours. Especially the travel times on the detours vary heavily from day to day.

### 3. RESULTS

The analysis of alternate routes regarding travel time advantages summarized in Table 3 reveals that there is only little reduction of travel time possible in all cases. On all freeways except on the M1 from Leeds to London drivers reduce the average travel time by detours. However, the majority of drivers increase their travel time by taking detours except on the freeway A1 in France from Lille to Paris as the median shows. The average number of vehicles on alternate routes varies for the analyzed freeways and directions and is not directly comparable as it depends on the penetration rate with connected vehicles as well as on the length of the analyzed freeway. By computing the ratio of vehicles on alternate routes to main route, the dependency on the penetration rate is avoided. Although the freeways in United Kingdom and Germany have almost the same length, the relative usage of alternate routes is more than double as high in Germany as in United Kingdom. The relative usage of alternate routes also depends on the direction of the freeway. This is caused by the network topology, e.g., in France the freeway heading into the city is often congested with reasonable and often used alternate routes at the same time, while in the other direction almost no detours occur.

The active route of the vehicle navigation system is not transmitted to the backend and is in a consequence not available for this empirical analysis. Some drivers probably decide to switch onto an alternate route although it is unreasonable with the current congestion on the main route and not suggested by the navigation device. By comparing travel time on the alternate route with the travel time on the freeway and taking delays through congestion into account, detours with intermediate stop can be determined and removed from the analysis. However, this heuristic can still include detouring vehicles that are nevertheless driving to an intermediate stop when where is congestion at the freeway at the same time. This influences median and average of the travel time differences in Table 3.

Table 3. Overview of travel time differences between main and alternate route (positive value: faster on alternate route; negative value: faster on main routes)

| Country (freeway)   | direction           | Average number of vehicles on alternative routes per day |          | Median in [s] | Average in [s] |
|---------------------|---------------------|----------------------------------------------------------|----------|---------------|----------------|
|                     |                     | absolute                                                 | relative |               |                |
| Germany (A8)        | München → Karlsruhe | 50.6                                                     | 0.63%    | -10           | 213            |
|                     | Karlsruhe → München | 33.9                                                     | 0.40%    | -113          | 152            |
| United Kingdom (M1) | London → Leeds      | 21.3                                                     | 0.24%    | -90           | 203            |
|                     | Leeds → London      | 17.0                                                     | 0.19%    | -100          | -38            |
| France (A1)         | Paris → Lille       | 3.1                                                      | 0.11%    | -50           | 129            |
|                     | Lille → Paris       | 18.9                                                     | 0.66%    | 52            | 68             |

The possible travel time increase or reduction by an alternate route is influenced by the delay on the main route, the uncongested travel time difference between main and alternate route as well as on the congestion on the alternate route. The biggest time factor is in most cases the delay on the main route which can reach in case of severe accidents several hours. In case of the daily congestion during rush hour the delays on the main route are often comparable small and there are also delays on the alternate route. It seems that the distribution of travel time deviations is similar for freeways without distant dependent toll like in Germany or United Kingdom. On distant dependent toll freeways like the A1 in France detours are seldom necessary as there is also almost no congestion except due to accidents. The

negative median values in Tab. 3 imply that half of the drivers taken a detour have a higher travel time on the alternate route instead of remaining on the main route. Only on the toll road from Lille to Paris the majority of the vehicles have a positive time gain choosing the alternate route. Additionally to the analysis of real travel times also the predicted travel times in the navigation device in the vehicle should be considered. As it is not possible to log and transmit the routes to the backend due to the technical implementation of the navigation devices used for the analysis in this paper, the routes and predicted travel times are reconstructed in a post processing step based on the dynamic traffic information available in the vehicles. This enables the categorization into recommended and not recommended detours. A detour is recommended, if the predicted travel time is lower on the alternate route than on the main route.

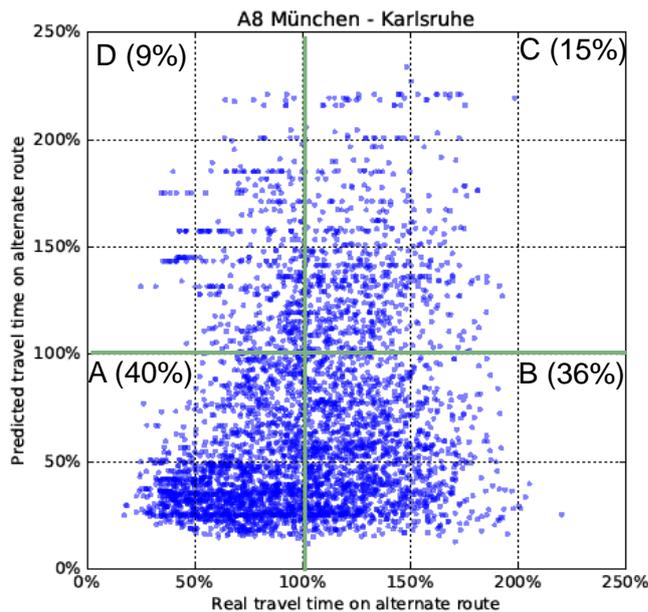


Fig. 9. Comparison of predicted and real travel time on alternate routes compared with main route

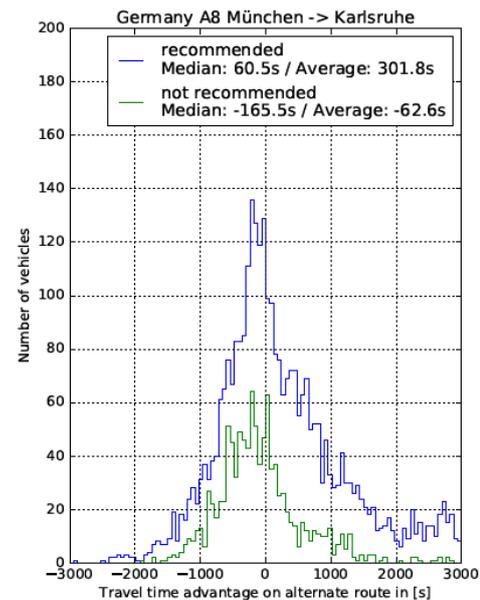


Fig. 10. Distribution of predicted and real travel time differences on alternate routes compared with main route

In Fig. 1 the real and predicted travel times of the vehicles on alternate routes are analyzed. The points in the diagram can be divided into 4 separate groups:

- Group A: Both real and predicted travel time is shorter for the alternate route compared to the main route
- Group B: Only the predicted travel time is shorter for the alternate route compared to the main route
- Group C: Neither real nor predicted travel time is shorter for the alternate route compared to the main route
- Group D: Only real travel time is shorter for the alternate route compared to the main route

On this specific freeway part the predicted travel time is lower for 76% of the vehicles on alternate routes compared to the main route. However, the ratio for the real travel time on the alternate route is roughly 50% lower and 50% higher. The ratio between the 4 groups is similar for the other direction on the A8 from Karlsruhe to München (not shown). The deviation between predicted and real travel time is partly caused by missing or incorrect traffic information on the freeway and also on the corresponding alternatives. Finally, the prediction of the travel times can be only accurate within a certain range as the traffic situation might change. This explains why there are still drivers on alternate routes reducing their travel time despite of opposite prediction and other way round. Those drivers maybe profit from their personal experience, can better predict the development of the travel times or obtain traffic information from another source. In Fig. 2 the distributions of travel time advantages for both recommend and not recommended detours are given and both exhibit a similar symmetric form. The maximum of both distributions is located close to 0, but in the range with increased travel times by alternate routes. The

median and average travel times are positive for the recommended alternate routes, i.e., that travel time could be reduced by following to the routing suggestion of the navigation device. In contrast, drivers deciding for detours without recommendation increase their travel time on median and average. But there are still some drivers reducing their travel time on not recommended detours, although the predicted travel time on the alternate route is greater than on the main route and in consequence not suggested by the navigation system. Using recommended alternate routes enables 18% of the drivers to reduce their travel time by over 1000 seconds in contrast to 6% on not recommended routes.

#### 4. CONCLUSION

The result of the empirical analysis shows that on average the drivers can slightly reduce their travel time in the range of some minutes by detouring from freeways in case of congestion. Some of the drivers can significantly reduce their travel time, while others spent much more time on the alternate route compared to the main route. However, the majority of the drivers even increase their travel time by detours by up to 2 minutes. The travel time differences are very similar for the toll free freeways A8 Germany and M1 in United Kingdom. The freeway A1 in France is partly a toll road resulting in almost no detours on this part.

The use of live traffic information for the routing decision leads as expected to better results. The recommended detours are on average 6 minutes better than driven detours without recommendation. So in most cases it is advisable to follow the route recommendation of the navigation device. However, there are still a few drivers which perform better than the recommendation by incorporating their own knowledge of the traffic network. This shows not only the importance of precision and coverage of live traffic information, but also the necessity and advantage of traffic predictions and risk-averse routing.

#### REFERENCES

- Auer, M., Rehborn, H. and Bogenberger, K. (2017). Empirical Analysis of Dynamic Route Choice Behavior on German Freeway A8 Based on Large Scale Vehicle Fleet Data. *6th Symposium of the European Association for Research in Transportation*
- Bonsall, P. (2004). Traveller behavior: Decision-making in an unpredictable world. In *Intelligent Transportation Systems*, 8 (1), 45-60
- Chorus, C. G., Molin, E. J., & Van Wee, B. (2006). Travel information as an instrument to change cardrivers' travel choices: a literature review. *European Journal of Transport and Infrastructure Research*, 6 (4), 335-364
- Han, Q., Timmermans, H., Dellaert, B., and van Raaij, F. (2008). Route choice under uncertainty: effects of recommendations. *Transportation Research Record: Journal of the Transportation Research Board*, (2082), 72-80.
- Iida, Y., Akiyama, T. and Uchida, T. (1992). Experimental analysis of dynamic route choice behavior. *Transportation Research Part B: Methodological*, 26 (1), 17-32.
- Mahmassani, H.S. and Liu, Y. H. (1999). Dynamics of commuting decision behaviour under advanced traveller information systems. *Transportation Research Part C: Emerging Technologies*, Vol. 7, No. 2, 1999, 91-107.
- Schlaich, J. (2010). Analyzing route choice behavior with mobile phone trajectories. *Transportation Research Record*, 2157, 78-85.
- Selten, R., Chmura, T., Pitz, T., Kube, S. and Schreckenberg, M. (2007). Commuters route choice behaviour. *Games and Economic Behavior*, 58 (2), 394-406.
- Tiratanapakhom, T., Oguchi, T., Tanaka, S., Hong, S., and Warita, H. (2014). Analyses of route choice and route switching behavior using ETC data from Tokyo metropolitan expressway. *SEISAN KENKYU*, 66(2), 205-209.
- Polydoropoulou, A., M. Ben-Akiva, and I. Kaysi. (1994). Influence of Traffic Information on Drivers' Route Choice Behavior. *Transportation Research Record*, 1453

# ASSESS THE EFFECTIVENESS OF EN-ROUTE INFORMATION PROVISION IN NO-NOTICE EVACUATION USING AGENT-BASED SIMULATION

Yi WANG<sup>a</sup>, Miltos KYRIAKIDIS<sup>b</sup> and Vinh N. DANG<sup>c</sup>

<sup>a</sup> *ETH Zurich, Future Resilient Systems, Singapore – ETH Centre, Singapore*  
Email: [yi.wang@frs.ethz.ch](mailto:yi.wang@frs.ethz.ch)

<sup>b</sup> *ETH Zurich, Future Resilient Systems, Singapore – ETH Centre, Singapore*  
Email: [miltos.kyriakidis@frs.ethz.ch](mailto:miltos.kyriakidis@frs.ethz.ch)

<sup>c</sup> *Laboratory for Energy Systems Analysis, Paul Scherrer Institute, Switzerland*  
*ETH Zurich, Future Resilient Systems, Singapore – ETH Centre, Singapore*  
Email: [vinh.dang@psi.ch](mailto:vinh.dang@psi.ch)

In a large number of transportation evacuation studies, optimal operating plans and pre-trip routing schedules are generated through iterations to approximate user equilibrium or system optimum traffic assignment. However, these assumption on travellers' learning and habit formation are most likely not an appropriate for the case of no-notice evacuation route choice behaviour. The evacuation time estimates of a no-notice evacuation are closely associated with the level of information perceived as well as evacuees' reactive traveller behaviour. Different from previous studies, the impact of providing real-time traffic information to evacuees and en-route drivers' adaptive routing behaviour are explicitly modelled. Given the complexity of the underlying processes and the multitude of factors influencing these procedures, agent-based traffic simulation approach is adopted. This study aims to model real-time traveller responses to real-time changes in the road network conditions due to the hazard's evolution in space and dynamic traffic conditions. A no-notice evacuation induced by HazMat spill from a train derailment accident is set up. Experimental results show that the more agents have access to real-time information and use adaptive routing, the faster concentrated traffic flows are dispersed, which helps reduce network clearance time.

**Keywords:** Evacuation Modelling, Agent-Based Simulation, Adaptive Routing

## 1. EXTENDED ABSTRACT

Emergency evacuation refers to the provision of transferring evacuees from disaster-affected areas to safer and protected shelters. It is a crucial operation often deployed in situations where the community or infrastructure is hit or impacted by natural (e.g., earthquakes, floods, bushfires) or man-made (e.g., terrorist attack) hazards. In evacuation planning, two primary categories are defined according to the amount of advanced notice available: no-notice events and pre-warned event. No-notice events are usually caused by natural and unintentional as well as intentional man-made hazards, where the warning time is zero. Pre-warned events mostly include natural hazards that give hours to days of advanced notice. The success of an evacuation strongly depends on many factors, such as warning time, response time, information and instructions dissemination procedure, evacuation routes, traffic flow conditions, dynamic traffic control measures, etc. (Dash and Gladwin, 2007). However, the evacuation planning varies significantly for pre-warned and no-notice evacuation, each of which necessitates different operational responses. Pre-warned evacuation planning mostly aims to develop traffic management strategies (e.g., contraflow and staged evacuation) by emergency service agencies, which becomes difficult to deploy in a no-notice evacuation due to the time constraint.

In a large number of transportation evacuation studies, optimal operating plans and pre-trip routing schedules are generated through iterations to approximate user equilibrium or system optimum traffic assignment. Such plans are built upon a number of desired assumptions, such as perfect traffic information or full compliance (or cooperative behavior). However, these assumption on travelers' learning and habit formation are most likely not an appropriate for the case of evacuation route choice behavior (Pel et al., 2012). During the occurrence of no-notice evacuations, a large share of travelers is inclined to switch routes based on the prevailing traffic information (Knoop et al., 2010) rather than to adhere to the pre-trip routing choices. The evacuation time estimates of such no-notice evacuation are closely associated with the level of information perceived as well as evacuees' reactive traveler behavior.

In a no-notice evacuation, evacuees respond to both the real-time traffic conditions and the hazard's evolution in space and time. It is therefore essential that transportation evacuation simulation takes this into consideration, therewith incorporating the important role of time-varying traffic information, disaster conditions, and warnings.

This study aims to model real-time traveler responses to real-time changes in the road network conditions due to the hazard's evolution in space and dynamic traffic conditions. The within-day re-planning technique using agent-based simulation framework proposed by Dobler et al. (2012) is adopted and implemented in the multi-agent transport simulation framework, allowing simulated agents to re-plan the routes between their activities while they are traveling. Compared to the user equilibrium condition wherein all travelers are fully aware of network conditions during pre-trip routing, varied levels of information provision as well as en-route route choice behavior with imperfect or limited information (e.g., myopic, aggressive, etc.) are explicitly modeled.

To demonstrate and investigate the impact of providing real-time traffic information to evacuees, a simulation case study is conducted based on real-world traffic and demographic data collected from, Sioux Falls, South Dakota. A no-notice evacuation induced by HazMat spill from a train derailment accident is set up. The user equilibrium assignment procedure in which evacuees are assumed to choose the shortest path from their origins to destinations is considered as the base case. Various levels of real-time traffic and hazard information dissemination are assessed, in order to illustrate the sensitivity of evacuation time estimates in relation to the en-route information provision, while incorporating various evacuees' en-route route choice behavior.

## REFERENCES

- Dash, N., & Gladwin, H. (2007). Evacuation decision making and behavioral responses: Individual and household. *Natural Hazards Review*, 8(3), 69-77.
- Dobler, C., Kowald, M., Rieser-Schüssler, N., & Axhausen, K. (2012). Within-day replanning of exceptional events. *Transportation Research Record: Journal of the Transportation Research Board*, (2302), 138-147.
- Knoop, V.L., Hoogendoorn, S.P., Van Zuylen, H.J. (2010). Rerouting behaviour of travellers under exceptional traffic conditions—an empirical analysis of route choice. In: Hoogendoorn, S.P., Pel, A.J., Taylor, M.A.P., Mahmassani, H. (eds.) *Proceedings of the 1st International Conference on Evacuation Modeling and Management*, pp. 113–128. Elsevier, Amsterdam.
- Pel, A. J., Bliemer, M. C., & Hoogendoorn, S. P. (2012). A review on travel behaviour modelling in dynamic traffic simulation models for evacuations. *Transportation*, 39(1), 97-123.
- Wolshon, B., & Murray-Tuite, P. (2013). The Role of OR in Emergency Evacuation from Hazmat Incidents. In *Handbook of OR/MS Models in Hazardous Materials Transportation* (pp. 269-292). Springer New York.

## USER'S BEHAVIOR AND DYNAMIC NETWORK LOADING

S. F. A. BATISTA <sup>a</sup> and L. LECLERCQ <sup>b</sup>

<sup>a</sup> Univ. Lyon, IFSTTAR, ENTPE, LICIT,  
F-69675, Lyon, France

Email: [sergio.batista@ifsttar.fr](mailto:sergio.batista@ifsttar.fr)

<sup>b</sup> Univ. Lyon, IFSTTAR, ENTPE, LICIT,  
F-69675, Lyon, France

Email: [ludovic.leclercq@entpe.fr](mailto:ludovic.leclercq@entpe.fr)

The main goal of this work is to investigate the influence of different types of users' behavior on: (i) the individual routes flows; and (ii) network performance in terms of its internal, inflow and outflow capacities. For this, we consider the risk-seeking and risk-aversion behavior (Prospect Theory), the bounded rational and the regret-aversion (Regret Theory) users' behavior. To determine time-dependent cost paths that account for congestion, shock-waves and spillback effects, we consider a Lighthill-Whitham-Richards (LWR) mesoscopic traffic model. We show that different types of users' behavior have a direct impact on the network performance. Depending on the model setting users spend more time to complete their trips, increasing the internal accumulation of the network, spreading the congestion backwards and thus decreasing the network inflow capacity.

**Keywords:** Dynamic Traffic Assignment, User's Behavior, Network Equilibrium, Network Loading

### 1. INTRODUCTION

The most commonly used dynamic traffic assignment (DTA) model is based on the Wardrop first principle. It considers that users are perfectly rational, have perfect information about the transportation system and aim at minimizing their own travel time. The related equilibrium network is known as the Deterministic User Equilibrium (DUE) (Sheffi, 1985). But, information about route travel times is not necessarily perfect. The Stochastic User Equilibrium (SUE) introduces error terms to cope up with imperfect information. In DTA, the most commonly used models to solve the SUE are the Multinomial Logit and the C-Logit. There are also other Random Utility (RU) models (see e.g., Prato, 2009) with different properties that can be used to solve the SUE. Of particular interest for this work, we highlight the Probit model, which introduces error terms at the link level rather than at the path level making possible to deal with path correlations. It is usually not considered in DTA problems because it is computationally expensive and it does not have a closed form. This model can also be solved using Monte Carlo simulations (Sheffi, 1985) and gamma distributed link travel times (Nielsen, 1997). But, RU models together with the DUE have been criticized in the literature due to its lack of behavioral realism. In fact, users tend to deviate from perfect rationality and thus from optimal route choices. In the literature, there are alternative frameworks discussed that account for different kinds of users' behavior. Based on the original ideas of *satisficing* behavior introduced by Simon (1957), Mahmassani and Chang (1987) introduced the Bounded Rational User Equilibrium (BR-UE) to model the departure time choice. Users choices are driven by aspiration levels, representing a set of goal variables that act as a maximum threshold to define the users' satisfaction with their choices. Batista et al. (in prep.) revisited these ideas of bounded rationality and tested a route choice framework that considers different definitions for the users' search order for the *satisficing* routes and different definitions of the aspiration levels, including the concept of indifference band. Other authors consider the application of regret theory (RT) applied to static traffic assignment (see e.g., Chorus 2014; Li and Huang, 2016). In RT, users aim to minimize their regret with respect to other unselected routes. If there is one route that has a lower travel time than the selected one, users will feel regret. Or, if not, users will feel joy. Other authors also consider the application of Prospect Theory (PT) to model users risk-aversion and risk-seeking behavior for their route choice (Avineri, 2006, Gao et al., 2010). In PT, users make their route choices in terms of time prospects that are evaluated through gains and losses against a reference point. According to this theory, users are: risk-averse when confronted with the prospect of gains; risk-seeking when confronted with the prospect of losses; and more sensible to losses than gains (loss effect). In this work, we propose to revisit these route choice frameworks and investigate how different kinds of users' behavior influences the network performance in terms of level of congestion and travel times. We also analyze

the differences between the individual route flows. For this, we consider time-dependent route costs that account for congestion and spillback effects. Traffic dynamics is reproduced by a LWR mesoscopic simulator (Leclercq and Becarie, 2012) and the case study is a Manhattan network.

## 2. NETWORK EQUILIBRIUM AND METHODOLOGY

To solve for the different network equilibrium, we consider the classical Method of Successive Averages (MSA) as described in Sheffi (1985) and a descent step size of  $1/k$  to ensure the algorithm is converging. As the convergence criteria, we consider the Gap function and the number of violations  $N(\Phi)$  (Sbayti et al., 2007), as well as a maximum number of descent step iterations.

### 2.1 Deterministic User Equilibrium (DUE) and Stochastic User Equilibrium (SUE)

The DUE is based on the Wardrop first principle, where users are assigned to the route with the lowest travel time for each origin-destination (od) pair and based on an all-or-nothing principle. To solve the SUE, that account for the variance of the route travel time distributions, we consider Monte Carlo simulations (Sheffi, 1985) and gamma distributed link travel times (Nielsen, 1997). The idea is to sample link travel times according to a gamma distribution and locally solve a DUE problem for each sample.

### 2.2 Prospect Theory Stochastic User Equilibrium (PT-SUE)

In PT, route travel times are evaluated in terms of time prospects that are framed as gains and losses against a reference point  $T_0^{od}$ . This means that users are prospect maximizers instead of utility minimizers. Each prospect is evaluated through a value function  $v(x)$  and a weighting function  $w(p)$  (Kahnemann and Tversky, 1979).  $v(x)$  depends on a set of parameters  $(\alpha, \beta, \lambda)$  that capture the users sensitivity for gains and losses. While,  $w(p)$  depends on a set of parameters  $(\gamma, \delta)$  that capture distortion in the users perception of the probabilities of gains and losses. In this work, we consider three different  $T_0^{od}$  defined as the mean, median and mode of the average route travel times for each od pair; and two sets of  $(\alpha, \beta, \lambda, \gamma, \delta)$  as defined in Tversky and Kahnemann (1992) (referred as KT) and Xu et al. (2011) (referred as Xu). We highlight that the parameters defined in Tversky and Kahnemann (1992) are not calibrated in a route choice context. However, we also consider this set for comparison purposes. Our implementation of PT considers route travel time distributions. To do so, we make use of Monte Carlo simulations and sample the link travel times according to a gamma distribution. Then, we calculate the samples of route travel times, for each od pair. The route samples of travel times are then evaluated in terms of time prospects and framed as gains and losses against  $T_0^{od}$ . To calculate the probabilities of gains and losses, we discretize the route travel time distributions into small bins, to define a set for each route and od pair. These probabilities are then considered for gains or losses, depending on the sample of the route travel time evaluated against  $T_0^{od}$ .

### 2.3 Bounded Rational Stochastic User Equilibrium (BR-SUE)

We consider the bounded rational framework as discussed in Batista et al. (in prep.). To consider the route travel time distributions, the authors also make use of Monte Carlo simulations. For each route travel time sample, a set of *satisficing* routes are identified for each od pair. That is, routes samples that are inferior to the aspiration level defined at the od level. Then, users are assigned based on an all-or-nothing assignment according to the defined search order. The authors consider a: stochastic search order, where users are randomly assigned to one of the *satisficing* routes; and a strict preferences search order (Zhao and Huang, 2016), where users are assigned to the first most preferred route that is *satisficing*. If there are no *satisficing* alternatives, users choose the least of the worst routes, i.e. the route with the lowest travel time sample. We adopt the concept of the indifference band  $\Delta^{od}$  for the definition of the aspiration level, for which we consider three values: 0; 100; and 500.

## 2.4 Regret Theory Stochastic User Equilibrium (RT-SUE)

In the implementation of RT, we follow the framework of Li and Huang (2016). But, we consider that link travel times are gamma distributed and make use of Monte Carlo simulations to calculate the equilibrium. In this work, we set three values of the regret aversion parameter  $\delta^{od}$ : 0; 0.1; and 10.

## 3. TEST SCENARIO AND PRELIMINARY RESULTS

For the dynamic tests, we consider a Manhattan network composed by 134 links (Figure 1 (i)) and a dynamic LWR mesoscopic simulator (Leclercq and Becarie, 2012). All links have a length of 100 meters and traffic lights regulate all internal intersections, i.e. excluding the entry and exit nodes. The traffic lights of the vertical links have green light duration of 45 s with a cycle of 60 s. Green times are set to be on the West-East and on the North-South directions. Two offsets of 10 and 20 s are introduced. We consider a fundamental diagram for each lane, with the following set of parameters:  $u=15$  m/s;  $w=5$  m/s; and  $k_{jam}=0.2$  veh/m/lane. The entry links (O1 to O6) have two lanes, while all the remaining links have 1 lane. The network has six entries (O1 to O6) and six exits (D1 to D6) yielding a total of 36 possible od pairs. To define the route choice set, we consider a maximum of 3 paths per od pair, that are calculated using a K-shortest path algorithm in distance. This gives a total of 108 routes. For the convergence criterion we set: 0 violations for  $N(\Phi)$ ; and a maximum of 250 descent steps. We consider a simulation period  $T=3000$  s.

We present the preliminary results in Figure 1. We first analyze the individual flows for od pair 3-4. We observe that PT does not give different route flows compared to SUE, that is also verified for the others od pairs. This is because PT is very sensible to the definition of  $T_0^{od}$  and the low number of prospects that are evaluated per od pair. In the case of BR, as  $\Delta^{od}$  increases, the users' indifference for the route choice increases. Then route flows will be equally distributed across the three routes as all routes are satisfying for  $\Delta^{od}=500$ . This shows the users' indifference for the route choice when all alternatives are *satisficing*. In the case of RT, as  $\delta^{od}$  increases, users switch to the route with the lowest travel time. We analyze now the MFD function of the network, considering the different models. Our results show a good agreement with Leclercq and Geroliminis (2015). For PT and RT, the network shows a better or worse performance depending on the total travel time. This can be observed in Figure 1 (iii) comparing PT-KT ( $T_0^{od}$  median) and SUE, where for the same total travel distance users need a higher total travel time. The BR also decreases the network capacity (Figure 1 (iv)), leading to an increasing on the waiting time for users to enter the network. This is related with the way that spillback effects propagate on the network and influences the entries. In the extended version of this work, we analyze a morning demand peak on the Sioux Falls network.

## ACKNOWLEDGEMENTS

This project is supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No 646592 - MAGnUM project). S. F. A. Batista also acknowledges funding support by the ARC7 Research Program.

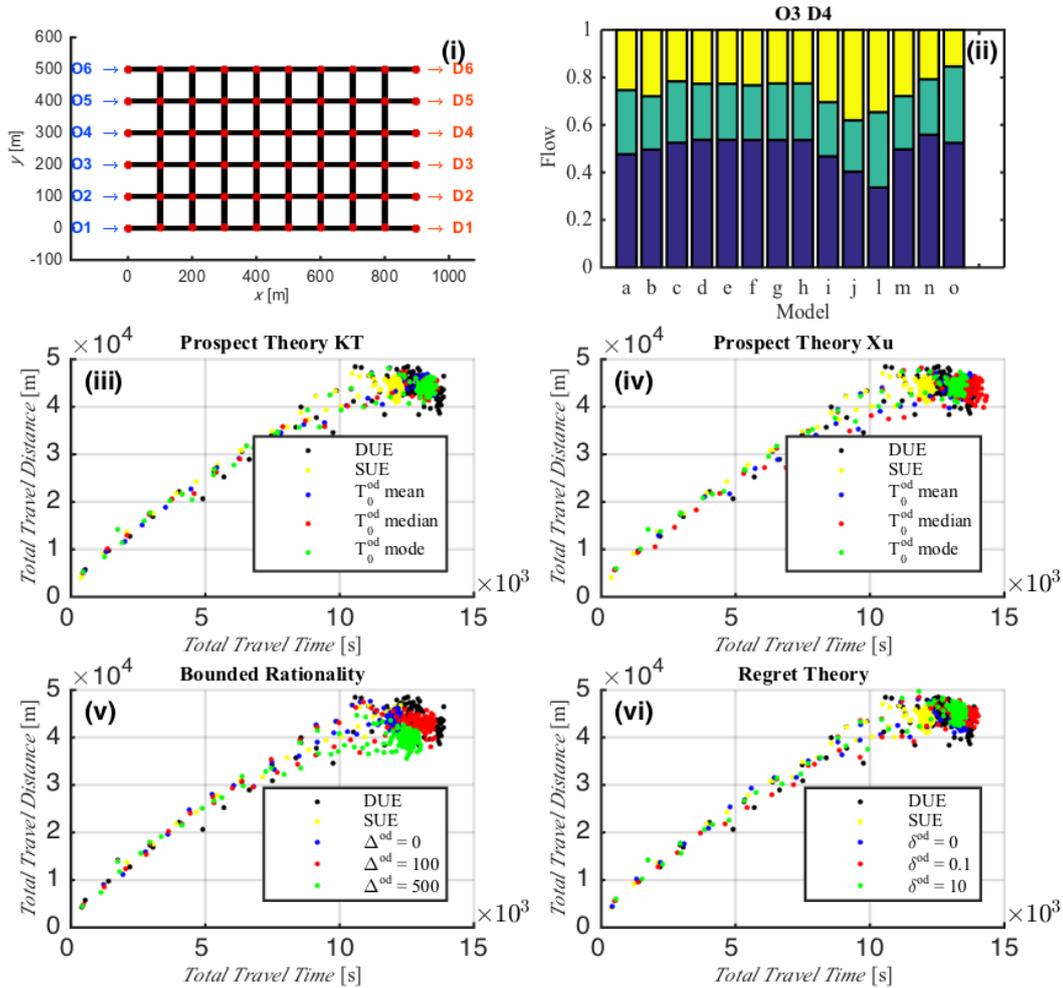


Figure 1. (i) Manhattan network. (ii) Three individual route flows for the od pair 3-4 and the different model settings (a to o). Macroscopic Fundamental diagram for: (iii) Prospect Theory and considering the KT parameters and  $T_0^{od}$ ; (iv) Prospect Theory and considering Xu parameters and  $T_0^{od}$ ; (v) Bounded rationality and different  $\Delta^{od}$ ; and (vi) Regret Theory and different  $\delta^{od}$ . Legend for models in (ii): a – DUE; b – SUE; c – PT-KT ( $T_0^{od}$  mean); d – PT-KT ( $T_0^{od}$  median); e – PT-KT ( $T_0^{od}$  mode); f – PT-Xu ( $T_0^{od}$  mean); g – PT-Xu ( $T_0^{od}$  median); h – PT-Xu ( $T_0^{od}$  mode); i to l –  $\Delta^{od} = 0, 100, 500$ ; and m to o –  $\delta^{od} = 0, 0.1, 10$ .

## REFERENCES

- Avineri, E. (2006) The effect of the reference point on stochastic network equilibrium. *Transportation Science*, 4, pp. 409-420.
- Batista, S.F.A., Zhao, C.-L. and Leclercq, L. (in prep.)
- Chorus, C. (2014) A generalized random regret minimization model. *Transportation Research Part B*, 68, pp. 224-238.
- Gao, S., Frejinger, E. and Bierlaire, M. (2010) Adaptive route choices in risky traffic networks. *Transportation Research Part C: Emerging Technologies*, 18, pp. 727-740.
- Leclercq, L. and Geroliminis, N. (2013) Estimating MFDs in Simple Networks with Route Choice. *Procedia – Social and Behavioral Sciences*, 80, pp. 99-118.
- Li, M. and Huang, H.-J. (2016) A regret theory-based route choice model. *Transportmetrica A: Transportation Science*, 13, pp. 250-272.
- Kahnemann, D. and Tversky, A. (1979) Prospect Theory: An analysis of decision under risk. *Econometrica*, 47, pp. 263-292.
- Mahmassani, H.S. and Chang, G.L. (1987) On boundedly rational user equilibrium in transportation systems. *Transportation Science*, 21, pp. 89-99.

- Nielsen, O.A. (1997) On the distributions of the stochastic components in SUE (Stochastic User Equilibrium) traffic assignment models. *Proceedings of seminar E held at the European Transport Forum Annual Meeting*, Brunel, England, pp. 77-93.
- Prato, G.C. (2009) Route choice modeling: past, present and future research directions. *Journal of Choice Modeling*, 2, pp. 65-100.
- Sbayti, H., Lu, C.-C. and Mahmassani, H.S. (2007) Efficient implementation of method of successive averages in simulation-based dynamic traffic assignment models for large-scale network applications. *Transportation Research Record*, 2029, pp. 22-30.
- Sheffi, Y. (1985) Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods. eds G.K. Chinenko, R. Herion and N. DeWolfe, pp. 261-344. Prentice Hall Inc., United States.
- Simon, H.A. (1957) A behavioral model of rational choice. Wiley, New York.
- Tversky, A. and Kahnemann, D. (1992) A theory of anticipated utility. *Journal of Risk and Uncertainty*, 5, pp. 297-323.
- Xu, H., Lou, Y., Yin, Y. and Zhou, J. (2011) A prospect-based user equilibrium model with endogenous reference points and its application to congestion pricing. *Transportation Research Part B: Methodological*, 45, pp. 311-328.
- Zhao, C.-L. and Huang, H.-J. (2016) Experiment of boundedly rational route choice behavior and the model under satisficing rule. *Transportation Research Part C*, 68, pp. 22-37.

**Session B4**  
**Sharing Economy**

# ON THE MORNING COMMUTE PROBLEM IN A Y-SHAPED NETWORK WITH INDIVIDUAL AND HOUSEHOLD TRAVELERS

Dongdong HE<sup>a,b</sup>, Yang LIU<sup>b\*</sup>, Qiuyan ZHONG<sup>a</sup> and David Z. W. WANG<sup>c</sup>

<sup>a</sup> Faculty of Management and Economics, Dalian University of Technology, China

Email: [dut\\_hedd@163.com](mailto:dut_hedd@163.com), Email: [zhongqiuy@dlut.edu.cn](mailto:zhongqiuy@dlut.edu.cn)

<sup>b</sup> Department of Industrial System Engineering and Management,

National University of Singapore, Singapore

Email: [e0144152@u.nus.edu](mailto:e0144152@u.nus.edu), Email: [iseliuy@nus.edu.sg](mailto:iseliuy@nus.edu.sg)

<sup>c</sup> School of Civil and Environmental Engineering, College of Engineering,

Nanyang Technological University, Singapore

Email: [wangzhiwei@ntu.edu.sg](mailto:wangzhiwei@ntu.edu.sg)

This paper examines the morning commute problem when both household travelers and individual travelers are considered in a Y-shaped network with two upstream links and a single downstream link. The household parents daily pass through an upstream bottleneck with a limited capacity before a school and drop off their children. Then, they traverse the downstream bottleneck common to both household and individual travelers and arrive at the workplace. We derive necessary conditions for the existence of equilibrium. We find that staggering bottleneck congestion, i.e., increasing the schedule gap between the work and the school start times, may not always improve social welfare. When the schedule gap is large enough to separate two groups, increasing the schedule gap will increase the total system cost. When the schedule gap is small, increasing the schedule gap may increase or decrease the total system cost depending on the group demand and the bottleneck capacities. Furthermore, we show that, when the schedule gap is large, expanding the capacity of the downstream bottleneck will always reduce the total system cost, and thereby the paradoxical phenomenon will not arise. When the schedule gap is small, the paradox may emerge in certain subcases.

**Keywords:** Bottleneck Congestion, Merging, Household Activity, Household Joint Decision, Capacity Expansion Paradox, Y-Shaped Network

## 1. INTRODUCTION

The analysis of dynamic congestion patterns is first proposed by (Vickrey 1969). The bottleneck model is introduced to explain the traffic dynamics of the morning commute. A large number of studies have extended Vickrey's model in various ways. The existence and the uniqueness of user equilibrium of a single bottleneck are examined by e.g., (Smith 1984, Daganzo 1985, Lindsey 2004). The morning commute with heterogeneous travel preferences is investigated, e.g., (Arnott *et al.* 1994, Liu *et al.* 2011, Liu *et al.* 2015). However, most studies in the bottleneck literature consider individual traveler as a decision-making unit. It is assumed that a commuter chooses departure time so as to minimize his or her travel cost. (Bhat *et al.* 2005) present some instances to demonstrate that individual decision-making is biased or counterintuitive in reality. "Household" joint decisions are considered when household members' decision-making are mutually affected by each other's preferences and activities (Ho *et al.* 2015). It is common that a family with a child chooses to live in a neighborhood with a school. In the morning, a parent drops off the child at school and then goes to workplace. Queues form on roads connected to the school and the workplace. On the other hand, an individual commuter without child chooses to live in a neighborhood with no school congestion so that he or she only needs to traverse the congested road linked to the workplace. This paper examines the trip scheduling problem when both household decision and individual decision are considered in the morning commute. We study a Y-shaped road network with two origins (i.e., a neighborhood of households and a neighborhood of individuals) and two bottlenecks (one before the school and one before the workplace). (Jia *et al.* 2016) study the equilibrium trip scheduling of a group of households in a single-bottleneck network with common origin and destination. Each household consists of an adult traveler and a child. The adult first passes a bottleneck before school, drops off the child at school, and then goes to the destination, i.e., workplace. (Liu *et al.* 2016) examine the equilibrium patterns using the same road network but considering a mixed flow of both household travelers and individual travelers. (Zhang *et al.* 2017) re-examine the problem in the case where a bottleneck is located before workplace instead of before school.

Along the line of (Liu *et al.* 2016, Zhang *et al.* 2017), we consider a group of household commuters and a group of individual commuters. However, our model differentiates from previous studies by considering a Y-shaped network with two bottlenecks (Figure 1) in which (a) the household commuters and the individual commuters live in two different neighborhoods; and (b) household parent may first pass through an upstream bottleneck before school due to the standard school start time. The Y-shaped networks have been studied in the literature. (Kuwahara 1990) study the equilibrium of a two-tandem bottleneck in the morning rush hours. (Arnott *et al.* 1993) investigate the properties of dynamic traffic equilibrium of a Y-shaped corridor with one bottleneck on each arm and a third bottleneck at downstream. (Lago *et al.* 2007) adopt a similar Y-shaped highway corridor to study the spillovers and merges. (Daniel *et al.* 2009) conducted a behavioral experiment in a controlled environment which is similar to that in (Arnott *et al.* 1993) and confirmed the theoretical bottleneck paradox. (Xiao *et al.* 2014) investigate the impacts of stochastic capacity and merging behavior at equilibrium. All these papers examine the equilibrium at Y-shaped network with individual travelers as decision-making units. None of them consider the joint household activities and decisions.

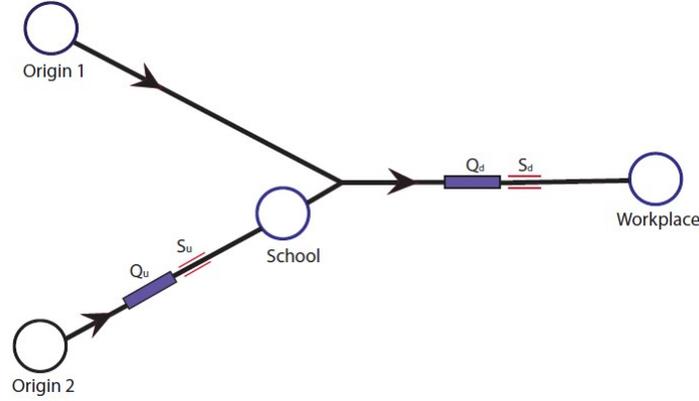


Figure 1. A Y-shaped road network with two bottlenecks

## 2. MAIN RESULTS

### 2.1 Model settings

We consider a Y-shaped road network that consists of two upstream links and a single downstream link as illustrated in Figure 1. There are two groups of daily commuters travel along the Y-shape corridor from home to a common destination, i.e., workplace at the central business district (CBD). A group of  $N_1$  individual travelers without child live in the neighborhood of Origin 1. They pass through an upstream link with an infinite service capacity (no congestion) and a downstream bottleneck with a service capacity  $s_d$  to the workplace. A group of  $N_2$  households with children live in the neighborhood of Origin 2. The household parents daily pass through an upstream bottleneck with a service capacity  $s_u$  before a school and drop off their children. Then, they traverse the downstream bottleneck common to both groups and arrive at the workplace. When the demand, i.e., the arrival rate of vehicles, exceeds the capacity of a bottleneck, a queue forms behind the bottleneck and consequently commuters experience queuing delays. Therefore, the travel time along a link consists of two parts: (1) a fixed free-flow travel time, and (2) a time-varying queuing delay. Following (Jia *et al.* 2016, Liu *et al.* 2016, Zhang *et al.* 2017), we assume that the free-flow travel time of each link is zero and each “household” consists of an adult and a child. The school starts at  $t_s^*$  and the work starts at  $t_w^*$ . Let  $\Delta t$  denote the schedule gap  $t_w^* - t_s^*$ . When commuters arrive at their school or workplace, their schedule delay will be the difference between the actual arrival times and desired arrival times, i.e.,  $t_s^*$  at the

school and  $t_w^*$  at the workplace. Therefore, the travel cost consists of the costs of queuing delay and schedule delay as in the classic bottleneck analysis. Commuters choose a departure time from home so as to minimize their own travel cost. At equilibrium, no individual or household can unilaterally change departure time from home to reduce the individual travel cost or the household travel cost. Let  $\alpha$  denote the value of travel time. Let  $\beta$  denote the unit cost of early schedule delay and  $\gamma$  denote the unit cost of late schedule delay. We assume that  $\alpha > \beta$  which is a necessary condition for the existence of equilibrium (Lindsey 2004). Now, we can formulate the individual travel cost and household travel cost, respectively.

*Individual travel cost* at departure time  $t$  from Origin 1 is:

$$C_I(t) = \alpha \cdot T_d(t) + \beta \cdot \max\{0, t_w^* - t - T_d(t)\} + \gamma \cdot \max\{0, t + T_d(t) - t_w^*\} \quad (1)$$

where  $T_d(t)$  is the queuing delay at the downstream bottleneck located before the workplace.

*Household joint travel cost* at departure time  $t$  from Origin 2 is:

$$C_H(t) = \alpha \cdot T_u(t) + \beta \cdot \max\{0, t_s^* - t - T_u(t)\} + \gamma \cdot \max\{0, t + T_u(t) - t_s^*\} + \alpha \cdot (T_u(t) + T_d(t')) \quad (2)$$

$$+ \beta \cdot \max\{0, t_w^* - t - T_u(t) - T_d(t')\} + \gamma \cdot \max\{0, t + T_u(t) + T_d(t') - t_w^*\}$$

where  $t'$  is the time that household parents arrives at downstream bottleneck if he/she departs at time  $t$  from Origin 2 ( $t' = t + T_u(t)$ ), and  $T_u(t)$  is the queuing delay at the upstream bottleneck located before the school. The household joint travel cost includes the travel costs of a work trip and a school trip within a family experienced by an adult and a child, respectively.

## 2.2 Findings

We derive necessary conditions for the existence of equilibrium. When the upstream capacity is smaller than the downstream capacity  $0 < s_u < s_d$ , equilibrium always exists. When the upstream capacity is larger than the downstream capacity  $s_u > s_d$ , there are two subcases: (a) if the schedule gap is large enough, i.e.,  $\Delta t \geq \overline{\Delta t}_1 = (2\beta + \gamma - 2\alpha)/(\beta + \gamma) \cdot (N_2/s_d) - (\gamma - 2\alpha)/(\beta + \gamma) \cdot (N_2/s_u) + \gamma/(\beta + \gamma) \cdot (N_1/s_d)$ , the individual travelers pass the downstream bottleneck after all the household travelers pass the downstream bottleneck. The two groups do not merge at downstream link and equilibrium always exists; and (b) if the schedule gap is small  $\Delta t < \overline{\Delta t}_1$ , equilibrium exists if and only if the condition  $s_u \leq \alpha/(\alpha - \beta) \cdot s_d$  is met.

Our Y-shaped network is different from the network topology in (Jia *et al.* 2016, Liu *et al.* 2016, Zhang *et al.* 2017) in which two sequential links with a single bottleneck are considered. Only when there is no merging interaction between two groups at the downstream link and one of the bottleneck does not emerge, our equilibrium pattern is identical to the cases in which two groups are separated in (Liu *et al.* 2016, Zhang *et al.* 2017). For instance, when the upstream capacity  $s_u$  is infinite, and the schedule gap is large, i.e.,  $\Delta t \geq \overline{\Delta t}_1$ , that is no merging interaction between two groups at the downstream link. The equilibrium pattern of our model is identical to the case where two groups are separated in (Zhang *et al.* 2017). The merging effects in the Y-shaped network significantly increase the number of possible equilibrium queuing patterns.

Increasing the schedule gap  $\Delta t$ , may not always improve social welfare. When the schedule gap is large, increasing the schedule gap will increase the total system cost. When the schedule gap is small, several subcases are found in which increasing the schedule gap will not reduce the total system cost as well. Therefore, we should design congestion staggering policy by examining its impacts on both the total

system cost and the overall congestion.

When the schedule gap  $\Delta t$  is large enough to separate the two groups, there are two subcases: (a) if the ratio of the number of household travelers to the number of individual travelers is larger than the relative cost of late arrival and early arrival time  $2N_2 / N_1 \geq \gamma / \beta$ , expanding the capacity of the upstream bottleneck or the downstream bottleneck will reduce total system cost. The paradoxical phenomenon (Arnott *et al.* 1993, Xiao *et al.* 2016, Daniel *et al.* 2009) will not arise; and (b) if  $2N_2 / N_1 < \gamma / \beta$ , expanding the capacity of the downstream bottleneck will reduce the total system cost, but expanding the upstream bottleneck capacity will increase the total system cost. On the other hand, when the schedule gap  $\Delta t$  is small, expanding the downstream bottleneck will always reduce the total system cost. However, expanding the upstream bottleneck may increase or decrease the total system cost depending on the group demand and the bottleneck capacities.

## ACKNOWLEDGEMENTS

The work was supported by Singapore Ministry of Education Academic Research Fund Tier 1 (WBS NO: R-266-000-0940112), the National Nature Science Foundation of China (NSFC,71533001), and the China Scholarship Council which supports Dongdong He's visiting study at the National University of Singapore for two years.

## REFERENCES

- Arnott, R., De Palma, A. and Lindsey, R. (1993) Properties of dynamic traffic equilibrium involving bottlenecks, including a paradox and metering. *Transportation Science*, 27, pp. 148-160.
- Arnott, R., De Palma, A. and Lindsey, R. (1994) The welfare effects of congestion tolls with heterogeneous commuters. *Journal of Transport Economics and Policy*, pp. 139-161.
- Bhat, C.R. and Pendyala, R.M. (2005) Modeling intra-household interactions and group decision making. *Transportation*, 32, pp. 443-448.
- Daganzo, C.F. (1985) The uniqueness of a time-dependent equilibrium distribution of arrivals at a single bottleneck. *Transportation Science*, 19, pp. 29-37.
- Daniel, T.E., Gisches, E.J. and Rapoport, A. (2009) Departure times in y-shaped traffic networks with multiple bottlenecks. *The American Economic Review*, 99, pp. 2149-2176.
- Ho, C. and Mulley, C. (2015) Intra-household interactions in transport research: a review. *Transport Reviews*, 35, pp. 33-55.
- Jia, Z., Wang, D. Z. and Cai, X. (2016) Traffic managements for household travels in congested morning commute. *Transportation research part E: logistics and transportation review*, 91, pp. 173-189.
- Kuwahara, M. (1990) Equilibrium queueing patterns at a two-tandem bottleneck during the morning peak. *Transportation Science*, 24, pp. 217-229.
- Lago, A. and Daganzo, C.F. (2007) Spillovers, merging traffic and the morning commute. *Transportation Research Part B: Methodological*, 41, pp. 670-683.
- Lindsey, R. (2004) Existence, uniqueness, and trip cost function properties of user equilibrium in the bottleneck model with multiple user classes. *Transportation Science*, 38, pp. 293-314.
- Liu, W., Zhang, F. and Yang, H. (2016) Modeling and managing morning commute with both household and individual travels. *Transportation Research Part B: Methodological*, 103, pp. 227-247.
- Liu, Y. and Nie, Y.M. (2011) Morning commute problem considering route choice, user heterogeneity and alternative system optima. *Transportation Research Part B: Methodological*, 45, pp. 619-642.
- Liu, Y., Nie, Y.M. and Hall, J. (2015) A semi-analytical approach for solving the bottleneck model with general user heterogeneity. *Transportation research part B: methodological*, 71, pp. 56-70.
- Smith, M.J. (1984) The existence of a time-dependent equilibrium distribution of arrivals at a single bottleneck. *Transportation Science*, 18, pp. 385-394.
- Vickrey, W.S. (1969) Congestion theory and transport investment. *The American Economic Review*, 59, pp. 251-260.

- Xiao, L.L, Liu, R. and Huang, H.J. (2014) Stochastic bottleneck capacity, merging traffic and morning commute. *Transportation Research Part E: Logistics and Transportation Review*, 64, pp. 48-70.
- Zhang, F., Liu, W., Wang, X. and Yang, H., (2017) A new look at the morning commute with household shared-ride: How does school location play a role?. *Transportation Research Part E: Logistics and Transportation Review*, 103, pp. 198-217.

**Session B5**  
**Travel Demand Management and Estimation**

# AN EFFICIENT ALGORITHM FOR HIGH-DIMENSIONAL DYNAMIC ORIGIN-DESTINATION DEMAND ESTIMATION OF INEFFICIENT LARGE-SCALE STOCHASTIC SIMULATION-BASED TRAFFIC MODELS

Carolina OSORIO <sup>a</sup>

<sup>a</sup> *Department of Civil and Environmental Engineering,  
Massachusetts Institute of Technology, USA*

*Email: [osorioc@mit.edu](mailto:osorioc@mit.edu)*

This paper focuses on the offline calibration of demand, as defined by dynamic origin-destination (OD) matrices, for simulation-based traffic models. We propose a metamodel approach. Recently, a metamodel approach for the calibration of a single-time interval has been proposed. We extend the approach to consider the joint calibration of OD's for multiple time intervals. For a network with  $n$  links,  $r$  endogenous OD entries for each of  $k$  time intervals, the dynamic OD estimation problem is defined as a simulation-based optimization problems with a decision vector of dimension  $k*r$ . Our proposed metamodel approach solves, at every iteration of the algorithm, a series of  $k$  decoupled analytical subproblems with a decision vector of dimension  $r$  and constrained by a system of  $n$  nonlinear equations. Hence, the complexity of the metamodel optimization problem scales linearly with the size of the network and independently of the size of the route choice set. The approach is validated with a toy network. It is benchmarked versus a traditional derivative-free algorithm and versus SPSA for a Singapore case study. The talk discusses the potential of these ideas for real-time OD calibration.

**Keywords:** Dynamic OD Estimation, Simulation-Based Optimization, Metamodeling

## 1. INTRODUCTION

High-resolution urban traffic and mobility data is becoming increasingly available worldwide. This has sparked an increased interest in the development of traffic models to inform the design and the operations of urban mobility networks. Additionally, both the supply and the demand of our transportation systems are becoming more intricate (e.g., with vehicle-to-vehicle and vehicle-to-infrastructure communications). This is leading to more sophisticated and more intricate traffic models. Nonetheless, in order to enable the use of this new generation of higher resolution traffic models to inform practice, there is a pressing need to provide practitioners with systematic tools that enable the adequate calibration and validation of these models. The problem of model calibration has been widely studied in the literature. Nonetheless, there is a lack of algorithms that can efficiently address the difficult (e.g., high-dimensional, simulation-based, non-convex) calibration problems faced in practice.

This paper focuses on the offline calibration of demand, as defined by dynamic origin-destination (OD) matrices, for simulation-based traffic models. In general, the demand calibration problem aims to identify demand inputs that minimize a distance function between network performance estimates (e.g., expected link flows, speeds, travel times) obtained from field measurement and those obtained via simulation. The goal of this paper is to design an OD calibration algorithm suitable for high-dimensional problems and large-scale networks. Moreover, the aim is to design a computationally efficient algorithm that can identify good quality solutions within few simulation runs. In practice, calibration algorithms are used within tight computational budgets (i.e., few simulation runs are carried out). Hence, the design of efficient algorithms contributes to current, and pressing, needs of practitioners.

A review of the recent OD calibration literature is provided in Zhang and Osorio (2017). The most common approach to simulation-based OD calibration has been the use of general-purpose algorithms, such as Stochastic Perturbation Simultaneous Approximation (SPSA) (Balakrishna et al.; 2007; Vaze et al.; 2009; Lee and Ozbay; 2009; Cipriani et al.; 2011; Ben-Akiva et al.; 2012; Lu et al.; 2015; Tympakianaki et al.; 2015) and the genetic algorithm (GA) (Kim et al.; 2001; Stathopoulos and Tsekeris; 2004; Kattan and Abdulhai; 2006; Vaze et al.; 2009). These commonly used general-purpose algorithms are guaranteed to achieve asymptotic convergence properties for a broad class of problems (e.g., non-

transportation problems). Nonetheless, this generality comes with a lack of computational efficiency. In other words, the algorithms are not designed to identify good quality solutions fast (i.e., within tight computational budgets or few simulation runs). Nonetheless, when used to address OD calibration problems they are typically used within tight computational budgets. There is a current need to design efficient calibration algorithms.

This paper proposes to achieve computational efficiency by designing algorithms specifically tailored for calibration problems. More specifically, we propose to embed within the algorithm analytical and differentiable problem-specific structural information that enables the algorithm to identify good quality solutions within few simulation runs. The main idea, which we have successfully used for other continuous transportation problems (e.g., signal control (Chong and Osorio; 2017), congestion pricing (Osorio and Atastoy; 2017)), is to formulate, and embed within the algorithm, an analytical network model that provides an approximation of the (simulation-based) mapping between the decision vector and the objective function. For a calibration problem, the mapping approximates the relationship between the calibration vector (e.g., OD matrix) and the simulation-based components of the objective function (e.g., expected link flows).

Recently, this idea has been formulated for a time-independent calibration problem (i.e., a single time interval was considered along with a dynamic stochastic traffic simulator) (Osorio; 2017). In this abstract, we refer to the latter approach as the static method. In this paper, we extend the static method to propose a method suitable for dynamic OD estimation. Let us first summarize the main ideas of the static method. We then discuss its extension and discuss preliminary results.

## 2. FORMULATION

The static method is a metamodel method. This means that at every iteration of the algorithm, the following two main steps are carried out: (i) all simulation observations collected so far are used to fit the parameters of an analytical model, known as the metamodel; (ii) an analytical (i.e., not simulation-based) optimization problem is solved that optimizes the metamodel function (in this abstract, we refer to this metamodel optimization problem as the subproblem). The advantage of a metamodel approach is that the difficult simulation-based optimization problem is solved by solving a series of analytical subproblems, for which traditional and efficient gradient-based algorithms can be used. The challenge of such an approach lies in the formulation of a suitable metamodel. More specifically, the metamodel should: (i) be analytical and differentiable (such that the gradient-based algorithms can be used to solve the subproblems), (ii) be computationally tractable or computationally efficient (because the subproblems are solved at every iteration of the calibration algorithm), (iii) be scalable (such that calibration problems for large-scale networks can be addressed), and it should also (iv) provide a good approximation of the simulation-based objective function in the entire feasible region.

The formulation of a tractable, efficient yet also accurate model is a major challenge. The static method proposed such a formulation that satisfied the above criteria. For a network with  $n$  links, the analytical model was formulated as a system of  $n$  nonlinear equations. The model provides an analytical approximation of the mapping between OD demand and link counts and speeds. It has endogenous route choice. Note in particular, that the dimension of the system of equations scales independently of the dimension of the route choice set and of the link lengths. This makes it a scalable model.

Nonetheless, the model provides a stationary (hence, time-independent) mapping of demand to link metrics (e.g., expected link counts or speeds). A natural extension of the static method for a dynamic OD estimation problem would be to formulate a dynamic analytical model. Nonetheless, this would not be sufficiently tractable and scalable. Hence, we propose to use a stationary model, yet to correct for the transients through a low-dimensional set of metamodel parameters that are estimated (or fitted) at every iteration of the algorithm.

Consider a problem with  $K$  time intervals. We wish to determine an OD matrix for each of the  $K$  time intervals. Let  $O$  denote the number of endogenous entries of the OD matrix for a given time interval.

The dynamic OD estimation problem is defined as a simulation-based optimization problems with a decision vector of dimension  $KO$ . Our proposed approach proceeds as follows. At every iteration of the algorithm, we solve a series of  $K$  analytical subproblems, one for each time interval. For each subproblem, a metamodel is optimized. The metamodel used is the stationary network model of (Osorio; 2017) (i.e., for a network of  $n$  links, it is defined as a system of  $n$  nonlinear equations). In other words, each subproblem is an analytical non-convex optimization problem with a decision vector of dimension  $O$  and a set of  $n$  nonlinear equality constraints. The coefficients of the metamodel are fitted such as to minimize the distance with estimates of the simulation-based objective function. The dependency of the subproblems (e.g., the fact that counts during time interval  $k$ , depend on counts and on demand of past time intervals) is captured through the estimation of the coefficients of the metamodel parameters.

This simple idea has been tested both on toy networks (9 ODs, 28 links) with excellent results. Ongoing work is applying it to a calibration problem of a Singapore network, with four 30-minute time intervals during the morning peak period and 4050 ODs per time interval. The network contains 1150 links and over 18,000 routes. The preliminary results obtained so far are promising. The algorithm is benchmarked versus a standard general-purpose derivative-free optimization algorithm suitable for high-dimensional problems. The talk will also discuss the potential of these ideas for real-time OD calibration.

## REFERENCES

- Balakrishna, R., Ben-Akiva, M. E. and Koutsopoulos, H. N. (2007). Offline calibration of dynamic traffic assignment: simultaneous demand-and-supply estimation, *Transportation Research Record* 2003: 50–58.
- Ben-Akiva, M., Gao, S., Wei, Z. and Wen, Y. (2012). A dynamic traffic assignment model for highly congested urban networks, *Transportation Research Part C* 24: 62–82.
- Chong, L. and Osorio, C. (2017). A simulation-based optimization algorithm for dynamic large-scale urban transportation problems, *Transportation Science* . Forthcoming. Available at: <http://web.mit.edu/osorioc/www/papers/osoChoDynSOsubmitted.pdf> .
- Cipriani, E., Florian, M., Mahut, M. and Nigro, M. (2011). A gradient approximation approach for adjusting temporal origin-destination matrices, *Transportation Research Part C* 19(2): 270–282.
- Kattan, L. and Abdulhai, B. (2006). Noniterative approach to dynamic traffic origin-destination estimation with parallel evolutionary algorithms, *Transportation Research Record* 1964: 201–210.
- Kim, H., Baek, S. and Lim, Y. (2001). Origin-destination matrices estimated with a genetic algorithm from link traffic counts, *Transportation Research Record* 1771: 156–163.
- Lee, J. B. and Ozbay, K. (2009). New calibration methodology for microscopic traffic simulation using enhanced simultaneous perturbation stochastic approximation approach, *Transportation Research Record* 2124: 233–240.
- Lu, L., Xu, Y., Antoniou, C. and Ben-Akiva, M. (2015). An enhanced SPSA algorithm for the calibration of dynamic traffic assignment models, *Transportation Research Part C* 51: 149–166.
- Osorio, C. (2017). High-dimensional offline od calibration for stochastic traffic simulators of large-scale urban networks, *Technical report*, Massachusetts Institute of Technology. Under review. Available at: <http://web.mit.edu/osorioc/www/papers/osoODCalib.pdf> .
- Osorio, C. and Atastoy, B. (2017). Efficient simulation-based toll optimization for large-scale networks, *Technical report*, Massachusetts Institute of Technology. Under review. Available at: <http://web.mit.edu/osorioc/www/papers/osoAtaTollSO.pdf> .
- Stathopoulos, A. and Tsekeris, T. (2004). Hybrid meta-heuristic algorithm for the simultaneous optimization of the O-D trip matrix estimation, *Computer-Aided Civil and Infrastructure Engineering* 19(6): 421–435.
- Tympakianaki, A., Koutsopoulos, H. and Jenelius, E. (2015). c-SPSA: Cluster-wise simultaneous perturbation stochastic approximation algorithm and its application to dynamic origindestination matrix estimation, *Transportation Research Part C* 55: 231–245.
- Vaze, V., Antoniou, C., Wen, Y. and Ben-Akiva, M. (2009). Calibration of dynamic traffic assignment models with point-to-point traffic surveillance, *Transportation Research Record* 2090: 1–9.

Zhang, C. and Osorio, C. (2017). Efficient offline calibration of origin-destination (demand) for large-scale stochastic traffic models, *Technical report*, Massachusetts Institute of Technology. Under review. Available at: <http://web.mit.edu/osorioc/www/papers/zhaOsoODcalib.pdf> .

# HOURLY ORIGIN-DESTINATION DEMAND ESTIMATION USING TIME COEFFICIENTS FROM TRAFFIC COUNTS BASED ON SEMI-DYNAMIC TRAFFIC ASSIGNMENT

Motohiro FUJITA <sup>a</sup> and Shinji YAMADA <sup>b</sup>

<sup>a</sup>*Department of Civil Engineering, Nagoya Institute of Technology, Japan*

*E-mail: fujita.motohiro@nitech.ac.jp*

<sup>b</sup>*Nagoya Expressway Public Corporation, Japan*

*E-mail: Shinji.Yamada@nagoya-expressway.or.jp*

Day-long origin-destination (OD) demand for transportation prediction is advantageous in terms of accuracy and reliability because it is not affected by hourly variation of OD distribution. However, hourly traffic prediction is important for transportation analysis. We proposed and examined the TCoE model that estimates the time coefficients for OD demand from observed link flows given a proven day-long OD demand, which is based on a bi-level formulation of the generalized least square and the semi-dynamic traffic assignment (OD-modification approach). In this paper, we propose an formulation for the TCoE model and apply it into a large scale road network considering many subareas and dual directions and examine the accuracy in difference of the number of subareas.

**Keywords:** OD Demand Estimation, Semi-Dynamic Traffic Assignment, Observed Link Flow

## 1. INTRODUCTION

Day-long origin-destination (OD) demand for transportation prediction is advantageous in terms of accuracy and reliability because it is not affected by hourly variation of OD distribution. However, hourly traffic prediction is important for transportation analysis. We proposed and examined the TCoE model [Fujita et al. (2017) ] that estimates the time coefficients for OD demand from observed link flows given a proven day-long OD demand, which is based on a bi-level formulation of the generalized least square and the semi-dynamic traffic assignment(OD-modification approach). In this paper, the time coefficients of OD demand are calculated as follows:

$$E_{rs}^n = G_{rs}^n / Q_{rs} \quad (1)$$

where  $E_{rs}^n$  is the time coefficient of the OD pair  $rs$  for the period  $n$  ( $n = 1, 2, \dots, 24$ ),  $G_{rs}^n$  is the hourly OD demand of OD pair  $rs$  for period  $n$ , and  $Q_{rs}$  is the day-long OD demand of OD pair  $rs$ .

Previous works have developed a bi-level formulation for OD demand estimation from observed link flow and dynamic OD demand estimates based on the time-dependent proportional and dynamic assignments operated over continuous discrete periods (for instance, Yang et al.(1992), Zhou et al.(2003), Cascetta et al.(2013)). In this paper, we propose another formulation for the TCoE model and apply it into a large scale road network considering many subareas and dual directions and examine the accuracy in difference of the number of subareas.

## 2. CONCEPT OF SEMI-DYNAMIC OD -MODIFICATION APPROACH AND TCoE MODEL

This section reviews the semi-dynamic concept in the OD-modification approach as comparing it with the OD modification in the TCoE model. When an hour is set for a study time period, the OD-modification approach semi-dynamically assigns hourly OD demand for 24 periods in the day by each hour based on the UE principle that drivers select a route with minimum time. Let  $T$  (=60 min) be the length of a period and  $G_{rs}^n$  be the hourly OD demand between OD pair  $rs$  ( $\in R, S$ ) in period  $n$  ( $\in N$ ). The  $G_{rs}^n$  is aggregated based on departure time (hourly OD-dep). Furthermore let  $\lambda_{rs}^n$  be the travel time for OD pair  $rs$  in period  $n$ . The OD-modification approach assumes that the maximum travel time between OD pair  $rs$  is less than  $T$  ( $\lambda_{rs}^n < T$ ) and that the hourly OD demand  $G_{rs}^n$  departs from the origin uniformly at the rate  $(G_{rs}^n/T)$  during period  $n$ .

Now we consider an OD pair  $rs$  with only a path and some links, the hourly OD demand  $G_{rs}^n$ , and the

travel time  $\lambda_{rs}^n$  for the path during period  $n$ . The OD-modification approach cannot describe the variation of link flows midway along a path because of static assignment, although it can uniformly assign the same traffic volume as the demand to all links along the path. Therefore, to minimize the error between observed and assigned link flows midway along the path, the OD-modification approach modifies  $G_{rs}^n$  to the level of demand along the path at

$$G_{rs}^n - \lambda_{rs}^n G_{rs}^n / (2T)$$

which subtracts half of the non-arrived traffic from the current OD demand. Furthermore, the subtracted OD demand  $q_{rs}^n = \lambda_{rs}^n G_{rs}^n / (2T)$ , which is a residual OD demand in the current period  $n$ , is assigned to the next period. By considering continuous periods, the residual OD demand of the previous period also flows in the current network. Therefore, the following equation expresses the hourly OD demand based on the midway ( $g_{rs}^n$ : hourly OD-mid) and that averages the error of link flows along a path as follows;

$$g_{rs}^n = q_{rs}^{n-1} + G_{rs}^n - \lambda_{rs}^n G_{rs}^n / (2T) \quad (2)$$

In the above equation,  $g_{rs}^n$  is the demand function for the OD-modification approach, the travel time  $\lambda_{rs}^n$  is a variable, and the previous residual demand  $q_{rs}^{n-1}$  is a constant in current period and is based on an estimate made in previous period.

When we set  $g_{rs}^n$  to the demand function for the OD modification, the OD-modification approach can be formulated as a nonlinear minimization problem; that is, a static UE assignment with elastic demand. Fujita et al.(2001) and Ujii et al.(2003) explored the basic model above with the extended OD modification to an urban road network including expressways with toll load. In this paper, we examine the characteristics of several time-of-day assignments of OD modification approach for a large-scale road network. We use this extended model (hereafter TUE) as the OD-modification approach. For comparison, we also adopt the TUE with fixed hourly demand (TUE-f) to eliminate the residual demand in TUE and assign hourly OD demand separately in each hour. As mentioned earlier, the OD modification is the semi-dynamic assignment method that modifies the hourly OD-dep ( $G_{rs}^n$ ) into the hourly OD-mid ( $g_{rs}^n$ ) to minimize the error of estimated link flow averaged midway over a path for considering residual demand. Note that the TCoE model proposed herein also estimates the hourly OD-mid ( $g_{rs}^n$ ) under given a day-long OD demand and operates the residual OD demand in almost the same way as the OD-modification approach because the TCoE model modifies the hourly OD-dep ( $G_{rs}^n$ ) into the hourly OD-mid ( $g_{rs}^n$ ) to minimize the error between the observed and estimated link flow midway along a path as a bi-level problem with elastic hourly demand given day-long OD demand.

### 3. FORMULATION OF TIME-COEFFICIENT ESTIMATION MODEL

Now we set a departure subarea to  $i$  in set  $K$  and an arrival subarea to  $j$  in set  $L$ . The upper problem of TCoE minimizes square errors between the observed and estimated hourly link flow under given the link flow proportion as follows [Fujita et al. (2017) ]:

$$\min_E \mathcal{Z} = \sum_n \sum_a \left( \sum_{kl} \sum_{r \in K} \sum_{s \in L} P_{a,rs}^n E_{kl}^n Q_{rs} - \hat{x}_a^n \right)^2 \quad (3)$$

$$s.t. \quad \sum_n E_{kl}^n = 1, \quad E_{kl}^n \geq 0 \quad \forall n, k, l \quad (4)$$

where  $Q_{rs}$  is the day-long OD demand for OD pair  $rs$ ,  $\hat{x}_a^n$  is the observed link flow for link  $a$  in period  $n$ ,  $P_{a,rs}^n$  is the flow proportion of link  $a$  for hourly OD demand for OD pair  $rs$  in period  $n$ , and  $E_{kl}^n$  is the time coefficient for departure area  $K$  and arrival area  $L$  in period  $n$ .

When the model is applied to many subareas, we propose another formulation by adding an additional equation to Eq. (3) in order to control an excess of fluctuation of time coefficients. That is, the additional equation is the sum of square errors between successive time coefficients with weight parameter;  $\sum_{kl} w_{kl} \sum_n (E_{kl}^n - E_{kl}^{n-1})^2$ . This model is a bi-level problem in which the upper problem is the above minimization and the lower problem is the TUE-f assignment. Therefore, the upper problem estimates the time coefficients under the given link flow proportions and the lower problem estimates link flow

proportions by using the TUE-f assignment with the hourly OD demand, which is calculated by multiplying the given day-long OD demand by the time coefficients of the upper problem. The time coefficients of the solution that minimizes the optimum function Z can be obtained as a convergence value after alternately calculating the upper and lower problems, and the hourly OD demand and link flows are also estimated simultaneously.

#### 4. CALCULATION METHOD OF HOURLY OD-dep FROM THE RESULT OF TCoE MODEL

Hourly OD demand obtained by TCoE is the hourly OD-mid that minimizes the error of estimated and observed link flows midway along paths between OD pairs. Therefore, the hourly OD demand by the TCoE model differs a little from OD demand aggregated based on departure time (hourly OD-dep) from survey data. When the hourly OD-dep is required for practical use, we explain the method to calculate the hourly OD-dep from the TCoE result.

When the hourly OD demand by TCoE is assumed to be the hourly OD-mid  $g_{rs}^n$ , the hourly OD-dep  $G_{rs}^n$  in period  $n$  is obtained by transforming Eq. 2 as follows:

$$G_{rs}^n = [g_{rs}^n - \lambda_{rs}^{n-1} G_{rs}^{n-1} / (2T)] / [1 - \lambda_{rs}^n / (2T)] \quad \forall n, r, s \quad (5)$$

where  $G_{rs}^n$  is the hourly OD demand based on departure time (hourly OD-dep) for OD pair  $rs$  in period  $n$ ,  $g_{rs}^n$  is the hourly OD demand based on the midway value (hourly OD-mid), taking into account the residual demand for OD pair  $rs$  in period  $n$ ,  $\lambda_{rs}^n$  is the minimum travel time for OD pair  $rs$  in period  $n$  but is set to  $\lambda_{rs}^n = T$  when  $\lambda_{rs}^n > T$  ( $T$  is the period). Here the residual demand in period  $n$  is expressed by  $\lambda_{rs}^n G_{rs}^n / (2T)$ .

### 5. APPLICATIONS AND CONSIDERATIONS

#### 5.1 Comparison in Estimation Accuracy for Basic Application

The hourly OD demand based on departure time aggregated by the road traffic census 2010 is hereafter called ‘‘initial OD.’’ We examined the characteristics of the hourly variation pattern of the initial OD by applying it to the TUE, TUE-f, and TCoE models. The study network is composed of 484 zones, 6683 links, and 4468 nodes, which is the Chukyo metropolitan network based on the road traffic census 2010. The observed link flow for examining assignment accuracy uses 292 links with 24 hours of data from the road traffic census 2010. From the result of the day-long UE assignment, the day-long UE predicts link flows with good accuracy and little bias in data variation.

Table 1 Comparison of RMS error for link flows for all vehicles

|                         | 7:00 | 8:00 | 9:00 | 22:00 | Total in a day |
|-------------------------|------|------|------|-------|----------------|
| TUE-f (with Initial OD) | 936  | 646  | 470  | 340   | 10177          |
| TUE (with Initial OD)   | 706  | 766  | 534  | 276   | 9450           |
| TCoE                    | 571  | 532  | 434  | 264   | 8695           |

Next, we examine the validity of the TCoE model by applying it to the road network. First, the hourly OD demand and link flow estimated by TCoE for a passenger car and a truck are examined. Table 1 shows the RMS errors of link flows by TUE-f, TUE, and TCoE. Comparing them for 24 hours and all vehicle types, indicates that the TUE-f assignment is much less accurate than the TCoE model. Conversely, the TCoE model can significantly improve its estimation accuracy by modifying the hourly OD demand. The accuracy at 7:00 for the TCoE model decreases by about 40% of RMS errors when compared with that of TUE-f with initial OD demand. The TCoE model can increase the accuracy for all periods uniformly because it modifies time coefficients for all periods simultaneously.

Comparing the scatter diagram of link flows at 7:00 and 22:00 for all assignments, these results indicated that the TCoE model could cancel the bias for overestimating at peak hour and underestimating at nighttime relative to the initial OD, and thereby greatly improve the estimation accuracy.

## 5.2 Examination for the Method to Calculate Hourly OD-dep from the Hourly OD-mid Estimated by TCoE

In Section 4 we discussed the method to calculate hourly OD demand based on departure time (hourly OD-dep) from the hourly OD demand based on the midway (hourly OD-mid) of the TCoE estimation. We applied the hourly OD-dep from the TCoE model to the TUE and analyze the hourly variation pattern to test the validity of the method. Figure 1 compares hourly variation patterns of the hourly OD-dep obtained by the calculation method in Eq. 5, the hourly OD-mid estimated by the TCoE model and the initial OD by survey. As seen in this figure, the variation pattern for hourly OD-dep has a higher value during the peak hour than the hourly OD-mid because the TCoE model estimates the hourly OD-mid that is justified to fit the link flow midway along each path, not fit the link flow near each origin. We compared the results of link flows assigned by the TUE with hourly OD-dep from TCoE model and the TCoE model directly. Based on this comparison, these results were almost the same.

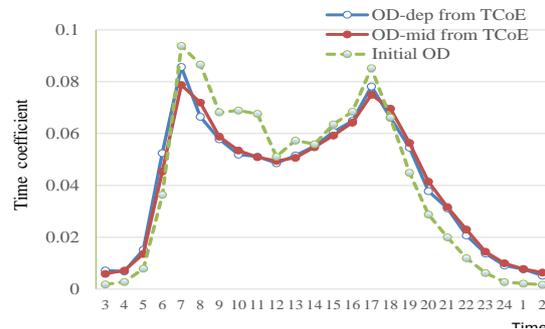


Figure 1 Comparison of hourly variation

## 5.3 Examination for the Number of Subareas for TCoE

Finally, the TCoE model was applied to several subareas and dual directions into which is divided the study area. In this study, we try to set several patterns of subareas and directions and examine which pattern is the highest in accuracy using RMS errors of estimated link flows of TCoE model. These results indicate that setting the time coefficient by subarea and direction can increase the estimation accuracy in which a pattern with 16 variables of time coefficients is the highest in estimation accuracy. Future research should examine the relationship between estimation accuracy and location of observed links should also be analyzed in a large-scale road network.

## ACKNOWLEDGMENT

This work was supported by JSPS KAKENHI Grant Number JP16K06534.

## REFERENCES

- Cascetta, E., Papola, A., Marzano, V., Simonelli, F., and Vitiello, I. (2013) Quasi-dynamic estimation of o-d flows from traffic counts: Formulation, statistical validation and performance analysis on real data. *Transportation Research* 55B, pp. 171-187.
- Fujita, M., Yamada, S. and Murakami, S. (2017) Time coefficient estimation for hourly origin-destination demand from observed link flow based on semi-dynamic traffic assignment. *Journal of Advanced Transportation*.
- Yang, H., Sasaki, T., Iida, Y., and Asakura, Y. (1992) Estimation of origin-destination matrices from link traffic counts on congested network. *Transportation research B*, 26B, pp.417-434.
- Zhou, X., Qin, X., and Mahmassani, H.S. (2003) Dynamic origin-destination demand estimation using multi-day link traffic counts for planning applications. *Transportation Research Record* 1831, 30-38.

# A LINEAR PROGRAMMING FORMULATION FOR INCENTIVE-BASED TRAVEL DEMAND MANAGEMENT

S. CHAKRABORTY <sup>a</sup>, D. REY <sup>a</sup> and S. T. WALLER <sup>a</sup>

<sup>a</sup> *School of Civil and Environmental Engineering,  
University of New South Wales, Sydney*

*Email: [shantanu.chakraborty@unsw.edu.au](mailto:shantanu.chakraborty@unsw.edu.au), [d.rey@unsw.edu.au](mailto:d.rey@unsw.edu.au), [s.waller@unsw.edu.au](mailto:s.waller@unsw.edu.au),*

In this study, we propose a linear programming (LP) framework for incentive-based travel demand management. The framework adopts the link transmission model as underlying traffic flow model for network loading due to its simpler traffic flow modelling approach and reduced computational complexity compared to the other solution methods such as link exit functions and cell transmission model. The proposed LP consists of an objective function that minimises total system travel time (TSTT) subject to a set of flow conservation and incentive budget constraints. We illustrate the formulation by solving an example network with six nodes and eight links. The solution to the LP includes the TSTT and percentage demand shift within timesteps with allocated incentives. Based on the available budget for incentives, desired level of improvement in TSTT and compliance rate of commuters, different proportion of demand shifts has been observed. We conclude the study with a sensitivity analysis and a proposed incentive scheme considering the base case as “no incentives” and other cases with different compliance rates.

**Keywords:** Incentive-Based Travel Demand Management, Linear Programming, Link Transmission Model, Dynamic Traffic Assignment

## 1. INTRODUCTION

All large cities and major highways throughout the world are facing severe traffic congestion resulting in an increased loss of man-hours for passengers as well as freight transportation. Congestion pricing is widely accepted as a powerful concept for managing traffic congestion and for providing effective travel demand management strategies (Nijkamp & Shefer, 1998) (Rouwendal & Verhoef, 2006) (Small, et al., 2007). However, the main challenge of congestion pricing remains as the details of the implementation combined with regional and political compatibility. An alternative approach to congestion pricing would be an incentive-based travel demand management (TDM) strategy that has the potential to create externalities equivalent to pricing. As per psychological theory, both pricing and incentive schemes bring out similar behavioural changes in the commuters towards an overall improvement of the performance of the system (Kahneman & Tversky, 1984) (Knockaert, et al., 2007) (Ettema, et al., 2010). Further, incentive based TDM approach has a better chance of gaining public acceptability compared to congestion pricing as per previous studies that resulted in increased frequency of bus usage (Fuji & Kitamura, 2003), adoption of safe driving (Dionne, 2011) and reduction in peak hour volumes (Currie & Walker, 2009) (Zhang, 2014) due to incentives. A lottery-based incentive mechanism to promote the usage of public transport during off-peak hours was found to shift driver behaviour significantly (Rey, et al., 2015). Another reward scheme to choose alternate departure times and routes was also found to gain a participation rate of up to 60% with an experienced travel time reduction of nearly 20% (Hu, et al., 2014). On the other hand, congestion pricing is visualized as an unfair strategy for the low-income segment of the population leading to low business activities and considerable resistance from the daily commuters and policy makers. Considering these favourable aspects, this paper proposes an incentive-based TDM approach in a linear programming framework.

Despite significant exploration, incentive-based TDM remains as an active area of research since modelling and development of optimal incentive schemes with significant compliance from the commuters is highly complex. These schemes may be devised to incentivise the travellers to either shift their departure times, routes to destinations or mode of transport. The conventional TDM strategies are mostly targeted towards addressing the peak period congestion associated with morning and afternoon commutes of daily travellers with an idea of spreading the peak throughout the day (Bliemer, et al., 2009) (Ben-Elia and Dick, 2009). This will lead to maximum utilization of the limited transport infrastructure in a network. The “peak avoidance” concept termed as “Spitsmijden” has been tested in several projects

in the Netherlands to motivate car users to shift to an alternate mode of transport or time of travel to avoid rush-hour traffic (Knockaert, et al., 2007) (Ettema, et al., 2010). A reduction of up to 60% in the peak hour traffic was observed in these studies due to the travellers shifting their car trips to before and after the peak period (Ettema, et al., 2010).

The research interest discussed in this paper aims to develop an incentive scheme to obtain an optimum demand split over the study period in a linear programming framework. The methodology of the study is presented in the following section.

## 2. METHODOLOGY

In this study a link transmission model (LTM) based linear programming (LP) formulation is adopted for system optimum dynamic traffic assignment (SODTA). The LTM is selected for network loading due to its simpler traffic flow modelling approach and reduced computational complexity compared to the other solution methods such as link exit functions and cell transmission model (Long, et al., 2016). The input parameters to the LTM are time-varying travel demand ( $d_t^0$ ), free-flow speed ( $v_f$ ), backward wave speed ( $w$ ), jam density ( $k_j$ ) and flow capacity ( $Q_{max}$ ) for each link. The proposed LP consists of an objective function that minimizes total system travel time (TSTT) and a set of flow conservation and incentive budget constraints. We illustrate the proposed LP formulation by solving an example network with six nodes and eight links.

The proposed LP for incentive-based TDM is presented as follows:

Objective function:

$$\text{Minimize } \delta \sum_{t \in T} \sum_{i \in A_C} x_{i,t}$$

subject to

$$\text{Cumulative vehicle numbers (u/s): } z_{i,t}^+ = \sum_{t' < t} \sum_{h \in r^-(i)} y_{hi,t'} \quad \forall i \in A_D, \forall t \in T \quad (2)$$

$$\text{Cumulative vehicle numbers (d/s): } z_{i,t}^- = \sum_{t' < t} \sum_{h \in r^+(i)} y_{ij,t'} \quad \forall i \in A_C, \forall t \in T \quad (3)$$

$$\text{Number of vehicles on each link: } x_{i,t} = z_{i,t}^+ - z_{i,t}^- \quad \forall i \in A_C, \forall t \in T \quad (4)$$

Flow propagation constraint on sending flow:

$$\sum_{j \in r^+(i)} y_{ij,t} \leq (z_{i,t_s}^+ - z_{i,t}^-) \quad \forall i \in A_C, \forall t \in T \quad (5)$$

where  $t_s = t + \delta - \frac{L_i}{v_{f,i}}$

$$\text{Capacity constraint on sending flow: } \sum_{j \in r^+(i)} y_{ij,t} \leq \delta(q_i) \quad \forall i \in A_C, \forall t \in T \quad (6)$$

Flow propagation constraint on receiving flow:

$$\sum_{j \in r^-(j)} y_{ij,t} \leq (k_j L_j - (z_{j,t}^+ - z_{j,t_r}^-)) \quad \forall j \in A_D, \forall t \in T \quad (7)$$

where  $t_r = t + \delta - \frac{L_j}{w_j}$

$$\text{Capacity constraint on receiving flow: } \sum_{j \in r^-(j)} y_{ij,t} \leq \delta(q_j) \quad \forall j \in A_D, \forall t \in T \quad (8)$$

$$\text{Demand constraint: } z_{i,t}^+ = \sum_{t' < t} d_{t'} \quad \forall i \in A_R, \forall t \in T \quad (9)$$

$$\text{Incentive budget constraint: } \sum_{t \in T} |d_t^0 - d_t| \leq \frac{B}{f} \quad (10)$$

Demand shift impedance definition:  $p_{tt'} = |t - t'| * f$  (11)

Demand shift impedance constraint:  $\sum_{t,t' \in T} u_{tt'} * p_{tt'} \leq B$  (12)

Optimum demand definitional constraint:  
 $d_t = \sum_{t' \in T} u_{tt'} * d_t^0 \quad \forall t \in T$  (13)

Initial flow and non-negativity constraints:

$$y_{ij,0} = 0 \quad \forall i \in A, \forall j \in A \quad (14)$$

$$y_{ij,t} \geq 0 \quad \forall i \in A_C, \forall t \in T \quad (15)$$

$$x_{i,t} \geq 0 \quad \forall i \in A_C, \forall t \in T \quad (16)$$

$$z_{i,t}^+ \geq 0 \quad \forall i \in A_C, \forall t \in T \quad (17)$$

$$z_{i,t}^- \geq 0 \quad \forall i \in A_C, \forall t \in T \quad (18)$$

$$u_{tt'} \in [0,1] \quad (19)$$

The network consists of a set of source and sink centroid connectors and physical links. The demand is loaded into the network through the source centroid connector. The objective function of the LP minimises the total system travel time (TSTT) considering the waiting time of the vehicles at the source connector due to any bottleneck in the network. Constraints (2) - (4) are definitional constraints where constraints (2) and (3) define the cumulative vehicle numbers at upstream and downstream end of a link at each time step as the total transfer flow from the predecessor links and total transfer flow to the successor links respectively until the previous time step. Constraint (4) defines the number of vehicles in a link as a difference of these cumulative vehicle numbers at upstream and downstream end of the link at each time step.

Constraints (5) and (6) refer to the sending flow constraints of the original LTM model. According to the LTM theory, if a free flow state of traffic is observed at the downstream end of a link at time  $t$ , this state must have emerged from the upstream end and travelled at free flow speed of the link. The constraint (6) presents the upper limit of this sending flow as the capacity of the incoming link. Constraints (7) and (8) refer to the receiving flow capacity of each link accounting for congestion effects. As per the LTM theory, if a congested traffic state is observed at the upstream end of a link, this state must have emerged from the downstream end, traveling backwards at a speed equal to the congestion wave speed. Constraint (8) presents the upper limit of the receiving flow as the capacity of the receiving link improvement function. Constraint (9) expresses the CVN at the upstream of the source centroid connector as the sum of demand expected to enter the network over time. The Equation (10) is related to the incentive budget constraint. A variable  $d_t$  is introduced that represents the optimum demand split at each timestep keeping the total demand fixed. This optimum demand split leads to the minimum value of TSTT for a fixed total demand. The original time varying demand at each time step ( $d_t^0$ ) is incentivized to match this optimum demand split ( $d_t$ ) where a variable  $u_{tt'}$  represents the percentage of demand shift from timestep  $t$  to  $t'$ . In this demand shift process, we involve a compliance factor ( $f$ ) which is a function of value of time (VOT) of different class of commuters. We also introduce a demand shift impedance parameter ( $p_{tt'}$ ) which assigns a larger incentive to be allocated for demand shift to a further timestep. Equation (11) defines this parameter and (12) expresses the budget constraint based on this impedance. Equation (13) defines the optimum demand at each timestep after incentivising the original demand. Constraint (14) specifies the initial flow conditions. Constraints (15) – (19) state the non-negativity conditions.

The solution to the proposed LP includes the TSTT and percentage demand shift within timesteps with allocated incentives. Based on the available budget for incentives, desired level of improvement in TSTT and compliance rate of commuters, different proportion of demand shifts has been observed. We conclude the study with a sensitivity analysis and a proposed incentive scheme considering the base case

as “no incentives” and other cases with different compliance rates.

## **BIBLIOGRAPHY**

- Charness, G. & Gneezy, U., 2009. Incentives to exercise. *Econometrica*, 77(3), pp. 909-931.
- Currie, J. & Walker, W., 2009. Traffic congestion and infant health: Evidence from E-ZPass. *American Economic Journal: Applied Economics*, 3(1), pp. 65-90.
- Dionne, G., Pinquet, J., Maurice, M. & Vanasse, C., 2011. Incentive mechanisms for safe driving: a comparative analysis with dynamic data. *The Review of Economics and Statistics*, 93(1), pp. 218-227.
- Ettema, D., Knockaert, J. & Verhoef, E., 2010. Using incentives as traffic management tool: empirical results of the "peak avoidance" experiment. *Transportation Letters*, Volume 2, pp. 39-51.
- Fuji, S. & Kitamura, R., 2003. What does a one-month free bus ticket do to habitual drivers. *Transportation*, Volume 30, pp. 81-95.
- Hu, X. et al., 2014. *Behavior insights for an incentive-based active demand management platform*. Washington, DC., In Proceedings of the 93rd Annual Meeting of the Transportation Research Board.
- Kahneman, D. & Tversky, A., 1984. Choices values and frames. *American Psychologist*, 39(4), pp. 341-350.
- Knockaert, J. et al., 2007. Experimental design and modelling Spitsmijden. *Utrecht, Consortium Spitsmijden*.
- Merugu, D. P. B. R. N., 2009. *An incentive mechanism for decongesting the roads: A pilot program in Bangalore*. s.l., In Proceedings of ACM NetEcon Workshop.
- Nijkamp, P. & Shefer, D., 1998. Urban Transport Externalities and Pigouvian Taxes: A Network Approach. *Road Pricing, Traffic Congestion and the Environment*, pp. 171--189.
- Rey, D., Dixit, V. V., Ygnace, J.-L. & Waller, S. T., 2016. An endogenous lottery-based incentive mechanism to promote off-peak usage in congested transit systems. *Transport Policy*, Volume 46, pp. 46-55.
- Rouwendal, J. & Verhoef, E. T., 2006. Basic economic principles of road pricing: From theory to applications. *Transport policy*, 13(2), pp. 106--114.
- Small, K. A., Verhoef, E. T. & Lindsey, R., 2007. *The Economics of Urban Transportation*, Routledge, London. s.l.:Routledge.
- Zhang, Z., Fujii, H. & Managi, S., 2014. How does commuting behavior change due to incentives? An empirical study of the Beijing Subway System. *Transportation Research Part F*, Volume 24, pp. 17-26.

**Session B6**  
**Public Transport I**

# MIXED BUS FLEET ROUTING AND SCHEDULING CONSIDERING DRIVING RANGE AND LOCATIONS OF CHARGING STATIONS

Lu LI<sup>a</sup>, Hong K. LO<sup>b</sup> and Feng XIAO<sup>c</sup>

<sup>a</sup> *Department of Civil and Environmental Engineering,  
The Hong Kong University of Science and Technology, Hong Kong, PR China  
Email: [lliaw@connect.ust.hk](mailto:lliaw@connect.ust.hk)*

<sup>b</sup> *Department of Civil and Environmental Engineering,  
The Hong Kong University of Science and Technology, Hong Kong, PR China  
Email: [cehklo@ust.hk](mailto:cehklo@ust.hk)*

<sup>c</sup> *School of Business Administration,  
Southwestern University of Finance and Economics, PR China  
Email: [evan.fxiao@gmail.com](mailto:evan.fxiao@gmail.com)*

This paper develops a formulation for the multiple depot vehicle routing and scheduling problem with multiple vehicle types, which can explicitly deal with the range and charging issues, as well as locating the charging stations. Unlike the conventional electric vehicle scheduling problem, this problem is formulated as an integer linear program to find the global optimal solution. We first develop a novel approach to generate the feasible time-space-energy network for bus flow and time-space network for passenger flow. Then we introduce the external cost associated with emissions into the system, and investigate the minimum total system cost to operators and passengers by scheduling the bus fleet and locating the refueling stations. Furthermore, for computational efficiency, we develop another alternative formulation based on time-space bus flow network to handle the large-scale problems for approximate solutions. We apply the methods to bus services in Hong Kong to systematically analyze the number of buses needed, the operational cost, the passenger cost, and the bus emissions generated by buses with multiple energy sources.

**Keywords:** Bus Scheduling, Mixed Fleet, Electric Bus, Driving Range, Refueling/Charging

## 1. INTRODUCTION

Roadside pollution has attracted the attention of more and more people recently. It is reported by the European Commission that poor air quality causes more premature deaths than road accidents, responsible for 0.31 million premature deaths in Europe every year. In England in 2013, 11,490 deaths were caused by heavily NO<sub>2</sub> pollution (European Environmental Protection Agency, 2016). Obviously, the serious consequence of roadside emissions is underestimated. To improve air quality at roadside, over 220 cities and towns in fourteen countries around Europe are operating or preparing for Low Emission Zones. This has led to subsidy policies to promote alternative energy vehicles, such as hybrid and electric buses, to reduce or remove tail-pipe emissions. For public transportation, the majority of the bus fleet are heavy-duty diesel vehicles, which produce a substantial amount of air pollutants (US EPA, 2008). Converting the bus fleet from diesel to alternative energy sources will hence bring in sizeable reductions in emissions. During the process, range constraints are one critical issue to be tackled. Buses travel high daily mileages, so the range concern arising from electric or alternative energy buses becomes an important issue. In any case, in the deployment of alternative energy buses, particularly electric buses, the range constraints must be duly incorporated in bus route assignment and scheduling.

The traditional bus-scheduling problem is to cover all the trips in the timetable with fixed travel times and start and end locations. The objectives are to minimize the bus fleet size and the operating cost. Bunte and Kliwer (2009) gave an overview on vehicle scheduling problems and discussed several modeling approaches. Earlier bus-scheduling studies seldom considered energy and emissions. The main idea was to minimize the fleet size. The conventional bus service, which has fixed service schedules run by a single bus type, is not cost-effective due to the variable demand densities. Hence, some studies were conducted to investigate the multi-vehicle-type bus scheduling problem (MVT-VSP), involving different vehicle types of diverse capacities for timetabled trips (Hassold and Ceder, 2012; Ceder et al., 2013; Kim and Schonfeld, 2014; Hassold

and Ceder, 2014). It is worth noting that the majority of studies on VSP are trip-based. Only a few considered time-dependent passenger demand by using passenger waiting time to reschedule the service trips. Hassold and Ceder (2012) and Ceder et al. (2013) investigated the problem that how to make public bus services more attractive. Both of them aimed to minimize the passenger waiting time to improve public transport reliability. An and Lo (2015) and Lo et al. (2013) considered the passenger cost when they formulated the network design problem on transit and ferry services. In addition to fulfilling all the timetabled trips, this study seeks to find the most cost-effective design to balance the costs between the operator and the passengers via the waiting penalty, which will improve both the utilization of buses and increase the attractiveness of the services.

In recent years, more and more VSP studies considered clean-energy buses due to environmental and energy concerns. Due to range constraints and long recharging time, the approach to schedule the bus fleet is quite different from the conventional VSP. Li (2013) proposed a vehicle scheduling model for electric buses, as well as compressed natural gas (CNG), diesel, and hybrid buses, respectively, with the maximum route distance constraints. Zhu and Chen (2013) established a single depot vehicle scheduling model (SDVS) for electric buses, in which the charging time has a positive linear relationship with the corresponding period of bus mileage. Both of these studies considered a single fleet VSP and set the timetable for bus charging according to its maximum driving distance. Notice that using the travel distance to determine the driving range is not appropriate, since energy consumption is not only determined by distance, but also travel speed, passenger loading, gradient of the terrain, battery temperature, and current battery life (Thein and Chang, 2014; Goeke and Schneider, 2015), etc. In this paper, we incorporate energy consumption into the bus flow network generation by considering the impacts of travel speed and bus loading.

Locating the refueling stations is another indispensable issue when considering the range constraints in VSP. He et al. (2013) proposed a macroscopic planning model to design the optimal number of charging stations allocated to each metropolitan area, whereas the exact locations and capacities were not optimized. Lim and Kuby (2010) developed three heuristic algorithms to locate the refueling stations for alternative-fuels using path-based demands. Schneider et al. (2014) presented an electric vehicle routing problem, which considered setting a set of available recharging stations beforehand. In this paper, we generate the bus flow network by considering a set of feasible candidate refueling stations to tackle the station location problem.

For the mixed fleet VSP concerning the alternative energy sources, Li and Head (2009) established a bus-scheduling model to minimize the operating cost and excess vehicle emissions involving CNG and Hybrid buses, while range constraint and charging time were not taken into account in this study. Beltran et al. (2009) and Pternea et al. (2015) focused on developing an efficient model to solve a sustainability-oriented variant of the transit route network design problem. Both of them assigned the electric vehicles to pre-determined routes without considering the charging time issue. By using a set of given weights, the objective functions sought to minimize user, operator and external costs. A model of transit design for a mixed bus fleet was developed by Fusco et al. (2013) to compute the internal and external costs during the bus lifetime. It introduced an electric bus fleet to operate some lines of a transit corridor in an urban area. Charging facilities were considered in this paper whereas the assignment of travel routes for each electric bus was not addressed. Goeke and Schneider (2015) optimized the routes of a mixed fleet of electric and diesel commercial vehicles to fulfill the customer demand. Routing model was developed to design the travel routes for each electric vehicle. Yet the emission problem is not taken into consideration in their paper. Actually, few studies have been conducted on the mixed bus fleet scheduling problem with regards to the range and refueling constraints, and emissions reduction simultaneously. This paper considers not only the daily travel routes of two types of buses with different energy sources, but also different bus capacities to handle the demand.

In our previous work, we considered the mixed bus fleet management and government subsidy

scheme for early-retiring, purchasing, and routing problems based on frequency (Li et al., 2015; Li et al., 2016). In this paper, we will consider the problem based on schedule, which can explicitly deal with the range and refueling issues, as well as some additional problems, such as locating the refueling stations, considering the interaction with passengers and so on. We develop an integer linear program (ILP) to formulate the multiple depot (MD) vehicle scheduling problem with multiple vehicle types (MVT), referred to here as MD-MVT-VSP. The proposed MD-MVT-VSP, which involves the range and refueling (charging time) constraints, is different from the conventional VSP. A novel approach is developed to generate the feasible time-space-energy network for bus flow and time-space network for passenger flow, where the range and refueling issues can be precisely addressed. Then we introduce the external cost associated with emissions into the system, and investigate the minimum total system cost to operators and passengers by scheduling the bus fleet and locating the refueling stations. Furthermore, for computational efficiency, we propose another approach based on time-space bus flow network to solve the large-scale problem for approximate solutions. It is worth noting that the second approach have the identical solutions to the first approach when 1) the planning horizon is within the proposed feasible range; 2) the time interval in the network is larger than or equal to the full refueling time. Compared with the conventional VSP, the proposed MD-MVT-VSP has distinctive advantages. The main contributions of this paper are as follows:

- (1) We consider a mixed bus fleet scheduling problem with bus service coordination. By integrating bus services, passenger movements, and bus emissions into one model, the optimal bus scheduling scheme is determined;
- (2) We propose a novel approach to generate feasible time-space-energy network for bus flow by capturing energy consumption. The range and refueling issues are precisely tackled based on these networks, as well as the refueling station location problem. We then utilize the approach to study some practical problems, such as how to subsidize the electric buses, how to schedule the mixed fleet under Low Emission Zone;
- (3) We formulate the VSP as an integer linear program to find the global optimal solution instead of using the heuristic algorithm adopted by most studies. Furthermore, for computational efficiency, we develop an alternative formulation based on the time-space network to handle larger scale problems.

We first conduct a small network to verify these two approaches. Two experiments are conducted with different OD patterns and demands, and different driving range constraints for electric buses. Both of these approaches produce the same solutions for all the experiments as expected. The results show that with a larger number of OD demands and larger range constraints, the computational time for the first approach increases rapidly, almost 60 times in some cases. We then apply the second approach to the case of bus services in Hong Kong to systematically analyze the number of buses needed, the operational cost, the passenger cost, and the bus emissions generated by buses with multiple energy sources. Moreover, some practical problems are studied by using this method, such as government subsidy on electric buses, and bus scheduling under Low Emission Zone.

## REFERENCES

- An, K., & Lo, H. K., 2015. Robust transit network design with stochastic demand considering development density. *Transportation Research Part B: Methodological* 81, 737-754.
- Beltran, B., Carrese, S., Cipriani, E., & Petrelli, M., 2009. Transit network design with allocation of green vehicles: A genetic algorithm approach. *Transportation Research Part C: Emerging Technologies* 17(5), 475-483.
- Bunte, S., & Kliewer, N., 2009. An overview on vehicle scheduling models. *Public Transport* 1(4), 299-317.
- Ceder, A., Hassold, S., Dunlop, C., & Chen, I., 2013. Improving urban public transport service using new timetabling strategies with different vehicle sizes. *International Journal of Urban Sciences* 17(2), 239-258.
- Chao, Z., & Xiaohong, C., 2013. Optimizing battery electric bus transit vehicle scheduling with battery exchanging: Model and case study. *Procedia-Social and Behavioral Sciences* 96,

2725-2736.

- European Environmental Protection Agency, 2016. <<https://www.eea.europa.eu/>>
- Fusco, G., Alessandrini, A., Colombaroni, C., & Valentini, M. P., 2013. A model for transit design with choice of electric charging system. *Procedia-Social and Behavioral Sciences* 87, 234-249.
- Goeke, D., & Schneider, M., 2015. Routing a mixed fleet of electric and conventional vehicles. *European Journal of Operational Research* 245(1), 81-99.
- Hassold, S., & Ceder, A., 2012. Multiobjective approach to creating bus timetables with multiple vehicle types. *Transportation Research Record: Journal of the Transportation Research Board* (2276), 56-62.
- Hassold, S., & Ceder, A., 2014. Public transport vehicle scheduling featuring multiple vehicle types. *Transportation Research Part B: Methodological* 67, 129-143.
- He, F., Wu, D., Yin, Y., & Guan, Y., 2013. Optimal deployment of public charging stations for plug-in hybrid electric vehicles. *Transportation Research Part B: Methodological* 47, 87-101.
- Kim, M. E., & Schonfeld, P., 2014. Integration of conventional and flexible bus services with timed transfers. *Transportation Research Part B: Methodological* 68, 76-97.
- Kliwer, N., Gintner, V., & Suhl, L., 2008. Line change considerations within a time-space network based multi-depot bus scheduling model. In *Computer-aided Systems in Public Transport*, 57-70. Springer Berlin Heidelberg.
- Kliwer, N., Mellouli, T., & Suhl, L., 2006. A time-space network based exact optimization model for multi-depot bus scheduling. *European journal of operational research* 175(3), 1616-1627.
- Li, J. Q., & Head, K. L., 2009. Sustainability provisions in the bus-scheduling problem. *Transportation Research Part D: Transport and Environment* 14(1), 50-60.
- Li, J. Q., 2013. Transit bus scheduling with limited energy. *Transportation Science* 48(4), 521-539.
- Li, L., Lo, H. K., & Cen, X., 2015. Optimal bus fleet management strategy for emissions reduction. *Transportation Research Part D: Transport and Environment* 41, 330-347.
- Li, L., Lo, H. K., Xiao, F., & Cen, X., 2016. Mixed bus fleet management strategy for minimizing overall and emissions external costs. *Transportation Research Part D: Transport and Environment*. In press.
- Lim, S., & Kuby, M., 2010. Heuristic algorithms for siting alternative-fuel stations using the flow-refueling location model. *European Journal of Operational Research* 204(1), 51-61.
- Lo, H. K., An, K., & Lin, W. H., 2013. Ferry service network design under demand uncertainty. *Transportation Research Part E: Logistics and Transportation Review* 59, 48-70.
- Pternea, M., Kepaptsoglou, K., & Karlaftis, M. G., 2015. Sustainable urban transit network design. *Transportation Research Part A: Policy and Practice* 77, 276-291.
- Schneider, M., Stenger, A., & Goeke, D., 2014. The electric vehicle-routing problem with time windows and recharging stations. *Transportation Science* 48(4), 500-520.
- Thein, S., & Chang, Y. S., 2014. Decision making model for lifecycle assessment of lithium-ion battery for electric vehicle—A case study for smart electric bus project in Korea. *Journal of Power Sources* 249, 142-147.
- US Environmental Protection Agency, 2008. Toxic air pollutants. <http://www.epa.gov/oar/toxicair/newtoxics.html>.

**Session B7**  
**Network Loading and Simulation II**

# A PROBABILISTIC ANALYTICAL TRAFFIC-THEORETIC NETWORK LOADING MODEL FOR LARGE-SCALE NETWORK OPTIMIZATION

Jing LU<sup>a</sup> and Carolina OSORIO<sup>b</sup>

<sup>a</sup> *Operation Research Center, Massachusetts Institute of Technology, USA*

*Email: [jl3724@mit.edu](mailto:jl3724@mit.edu)*

<sup>b</sup> *Department of Civil and Environmental Engineering,*

*Massachusetts Institute of Technology, USA*

*Email: [osorioc@mit.edu](mailto:osorioc@mit.edu)*

We formulate an analytical stochastic network loading model. It extends the model of Lu and Osorio (2017) by making it more scalable and computationally efficient. The proposed model has a complexity of 2, a constant regardless of the link's space capacity. This makes it suitable for large-scale network optimization where repeated evaluations of the network performances through the model are required. The model is validated versus a simulation-based implementation of the stochastic link transmission model (LTM). The experiment results indicate that the proposed model captures the dynamics of the link correctly. We then use the proposed model to address a city-wide signal control problem and benchmark it versus the model of Lu and Osorio (2017). The experiment shows, on average, an improvement of computational efficiency by 2 orders of magnitude, while preserving similar quality of the derived solutions.

**Keywords:** Stochastic Network Loading, Probabilistic Traffic Flow Modeling, Computational Efficiency

## 1. INTRODUCTION

There has been a growing interest in the formulation of traffic models that are both probabilistic and consistent with traditional deterministic traffic flow theory. This has been mostly triggered by both an interest from major transportation agencies around the world to estimate and improve the robustness and reliability of their networks, and by an increased availability of higher resolution traffic data, which enables the validation of the more probabilistic detailed models. Nonetheless, there is a lack of models that are both: (i) consistent with mainstream traditional deterministic traffic flow theory, and (ii) tractable enough to enable the efficient analysis and optimization of large-scale networks. This paper addresses this challenge. It formulates an analytical and probabilistic model, that both stems from deterministic traffic flow theory and is sufficiently efficient to be suitable to address large-scale network optimization problems.

Osorio and Flötteröd (2015) proposed a link model that is a stochastic formulation of the deterministic link-transmission model of Yperman et al. (2007), which itself is an operational formulation of Newell's simplified theory of kinematic waves (Newell; 1993). The main ideas of that model are summarized in Figure 1. Upon entrance to the link, vehicular traffic flow is delayed for  $k^{\text{fwd}}$  time steps in the lagged inflow queue LI. It then enters the downstream queue DQ. Flows in DQ are those ready to leave the link. Upon departure from the link, the newly available space is delayed for another  $k^{\text{bwd}}$  time steps in the lagged outflow queue LO before it is available at the upstream end of the link. Figure 1 also denotes the links expected inflow (resp. outflow) during time interval  $k$  as  $q^{\text{in}}(k)$  (resp.  $q^{\text{out}}(k)$ ). The paper also defines the upstream queue UQ as the sum of LI, DQ and LO. The model describes the link's upstream and downstream boundary conditions through UQ and DQ, respectively. It tracks the joint distribution of LI, DQ and LO. Hence, the joint distribution of the link's upstream and downstream conditions can be derived (i.e., the joint of UQ and DQ). The model complexity is in the order of  $\mathcal{O}(\ell^3)$  where  $\ell$  is the space capacity of the link.

The recent work of Lu and Osorio (2017) extends the model of Osorio and Flötteröd (2015) by making it more computationally efficient. More specifically, it no longer tracks the joint distribution of the various queues, instead it only tracks their marginals. Hence, it yields the marginal distribution of the link's upstream boundary conditions (UQ) and that of the link's downstream boundary conditions (DQ). It provides a simplified description of the dependencies between the upstream and the downstream

boundary conditions. It is therefore a more computationally efficient model, with a complexity of  $\mathcal{O}(\ell)$ . This reduction in model complexity facilitates large-scale network analysis.

In this paper, we formulate a model with further enhanced computational efficiency. The goal is to enable large-scale network optimization problems to be addressed efficiently. The proposed model has a state space of dimension 2. In other words, its complexity is now constant and no longer depends on the links space capacity of the link,  $\ell$ . This makes the model scalable, i.e., it is suitable for the optimization of large-scale networks.

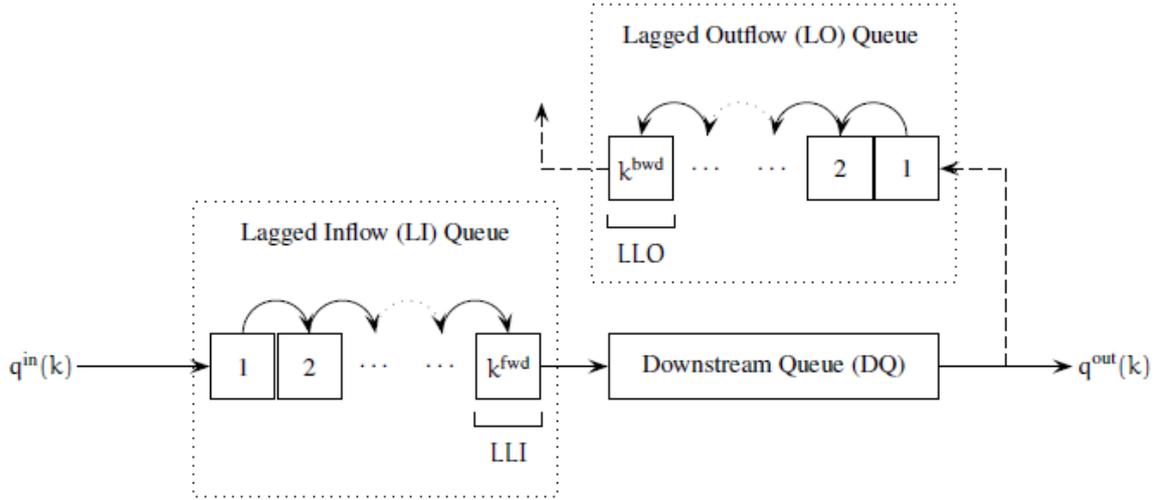


Figure 1. Link dynamics of the model of Osorio and Flötteröd (2015)

## 2. MODEL FORMULATION

The proposed formulation is based on the following observations. The upstream boundary conditions are mainly governed by whether there is space available in the link to enable the entry of a new vehicle. This is represented by the probability  $P(UQ < \ell)$ . The probability  $1 - P(UQ < \ell)$  is known as the spillback probability. Similarly, the downstream boundary conditions are mainly governed by whether there is a vehicle at the downstream end of the link ready to leave. This is represented by the probability  $P(DQ > 0)$ . The goal of this paper is to approximate each of these two probabilities, and to yield a time-dependent description of the link's boundary conditions by using only these two probabilities. This contrasts with our past works, which describe the boundary conditions by using full joint distributions (Osorio and Flötteröd; 2015) or full marginal distributions (Lu and Osorio; 2017). In other words, the proposed model only tracks these two key probabilities over time. For each discrete time interval indexed by  $k$ , the model yields approximations of the corresponding probabilities:  $P(UQ(k) < \ell)$  (which is the probability that there is space available at the upstream end of the link during time interval  $k$ ) and  $P(DQ(k) > 0)$  (which is the probability that there are vehicles ready to leave the downstream end of the link during time interval  $k$ ).

Consider a link with space capacity  $\ell$ , exogenous time-dependent (time indexed by  $k$ ) arrival rate  $q^{\text{in}}(k)$ , and exogenous downstream flow capacity  $q^{\text{out}}(k)$ . We first describe the main ideas underlying the derivation of  $P(DQ(k) > 0)$ , and then those of  $P(UQ(k) < \ell)$ . Hereafter,  $k$  refers to the discrete time interval index.

### 2.1 Downstream link boundary conditions

We model  $DQ(k)$  as an  $M/M/1/\ell$  queue. Recall that newly entered vehicles are delayed  $k^{\text{fwd}}$  time steps before entering  $DQ$ . Hence, at time  $k$  the arrival rate to  $DQ(k)$  is approximated as the expected inflow to the link at time  $k - k^{\text{fwd}}$ , which is denoted  $q^{\text{in}}(k - k^{\text{fwd}})$ . The service times of  $DQ(k)$  have an exponential distribution with exogenous rate  $\mu(k)$ .

In order to derive a closed-form expression for  $P(DQ(k) > 0)$ , we combine ideas from the works of Morse (1958, Chap. 6) and of Chong and Osorio (2017). Morse (1958, Chap. 6) yields an analytical expression for the transient queue-length distribution of an M/M/1/ℓ. Nonetheless, this expression requires tracking of entire marginal queue-length distribution over time. Since we would like to only track the scalar  $P(DQ(k) > 0)$ , we use the idea proposed in Chong and Osorio (2017) to approximate the transient evolution of a single queue-length probability. More specifically, we track  $P(DQ(k) = 0)$ , and then obtain  $P(DQ(k) > 0) = 1 - P(DQ(k) = 0)$ . The combination of these two ideas yields the following expression:

$$P(DQ(k) = 0) = P_k(DQ = 0) + [P(DQ(k-1) = 0) - P_k(DQ = 0)]e^{-\tau_{DQ}(k)\delta}, \quad (1)$$

where  $\tau_{DQ}(k)$  is the inverse of the relaxation time at time interval  $k$ ,  $\delta$  is the duration of time interval  $k$ ,  $P(DQ(k) = 0)$  is the probability of  $DQ(k)$  being empty at the end of time interval  $k$ ,  $P_k(DQ = 0)$  is the time-interval specific stationary probability of  $DQ = 0$ ,  $P(DQ(k-1) = 0)$  is the probability of  $DQ$  being empty at the end of time interval  $k-1$  (which is also the beginning of time interval  $k$ ). This equation states that the transient probability  $P(DQ(k) = 0)$  at the end of time interval  $k$  is defined as the sum of a stationary probability (term  $P_k(DQ = 0)$ ) and a term that decays exponentially with time. The latter term is the difference between the initial condition of time interval  $k$  (term  $P(DQ(k-1) = 0)$ ) and the corresponding stationary probability  $P_k(DQ = 0)$ .

The expression for the stationary probability at time interval  $k$ ,  $P_k(DQ = 0)$ , is obtained by approximating  $DQ(k)$  as an M/M/1/ℓ queue with arrival rate  $q^{\text{in}}(k - k^{\text{fwd}})$  and service rate  $\mu(k)$ . For an M/M/1/ℓ queueing system, there is a closed form expression for the stationary queue-length distribution (e.g., Morse (1958, Equation (3.5), Chap. 3)):

$$P_k(DQ = 0) = \frac{1 - \rho(k)}{1 - \rho(k)^{\ell+1}} \quad (2)$$

$$\rho(k) = q^{\text{in}}(k - k^{\text{fwd}})/\mu(k). \quad (3)$$

In order to derive an analytical expression for the function  $\tau_{DQ}(k)$ , we have carried out numerical simulation-based experiments, which have led to the following formulation:

$$\tau_{DQ}(k) = \tau_{DQ} \cdot \left( \frac{12|P(DQ(k-1) = 0) - P_k(DQ = 0)|^{\rho(k)}}{1 + e^{-1.5\rho(k)}} \right) \quad (4)$$

$$\tau_{DQ} = \frac{\mu(k)(1 - \rho(k))^2 + 1.5q^{\text{in}}(k - k^{\text{fwd}})\rho(k)\sqrt{\ell}}{(\rho(k) + 1)(\ell + 1)} \quad (5)$$

$$\rho(k) = q^{\text{in}}(k - k^{\text{fwd}})/\mu(k) \quad (6)$$

Equation (4) states that the relaxation time consists of two parts: one that is exogenous (term  $\tau_{DQ}$  defined by Equation (5)) and one that is endogenous (i.e., the term within parenthesis of Equation (4)). The expression for  $\tau_{DQ}$  is derived based on insights from the work of Newell (1982, Chap. 5), which studied the transient behavior of an M/M/1/ℓ queue. The underlying intuition behind these formulations will be discussed during the conference.

## 2.2 Upstream link boundary conditions

Let us now describe the main ideas underlying the approximation of the probability governing the links upstream boundary conditions,  $P(UQ(k) < \ell)$ . Note from Figure 1 that the outflow from  $UQ(k)$  is governed by the outflow from  $LO(k)$  queue, and more specifically from the last cell of  $LO(k)$  queue (denoted  $LLO(k)$ ). When  $LLO(k)$  is empty (i.e.,  $LLO(k) = 0$ ),  $UQ(k)$  can be viewed as a Poisson arrival process with rate  $q^{UQ}(k) = \sum_{r=0}^{k-1} q^{\text{in}}(r) - \sum_{r=0}^{k-k^{\text{bwd}}-1} q^{\text{out}}(r)$ , where  $q^{\text{in}}(r)$  (resp.  $q^{\text{out}}(r)$ ) represents the expected inflow to (resp. outflow from) the link during time interval  $r$ . When  $LLO(k)$  is not empty (i.e.,  $LLO(k) > 0$ ), we approximate  $UQ(k)$  as an M/G/ℓ/ℓ queue. The description of how we approximate the distributional parameters of these queueing systems will be presented in detail at the conference.

## 2.3 Expected inflow and outflow of the link

We now describe how we compute the expected inflow and outflow of the link at time step  $k$ . By definition:

$$q^{\text{in}}(k) = \lambda(k)P(UQ(k) < \ell) = \lambda(k)(1 - P(UQ(k) = \ell)) \quad (7)$$

$$q^{\text{out}}(k) = \mu(k)P(DQ(k) > 0) = \mu(k)(1 - P(DQ(k) = 0)) \quad (8)$$

Since we keep track of  $P(UQ(k) = \ell)$  and  $P(DQ(k) = 0)$  over time, we can proceed to compute  $q^{\text{in}}(k)$  and  $q^{\text{out}}(k)$  through Equations (7) and (8), respectively.

### 3. VALIDATION RESULTS AND OPTIMIZATION CASE STUDY

To validate the model, we have carried out experiments with a variety of demand and supply settings on a single link. We have validated the proposed analytical approach versus simulation-based estimates obtained from an event-based simulator used in validating both Osorio and Flötteröd (2015) and Lu and Osorio (2017). The simulator implements an exact version of the stochastic link transmission model presented in Osorio and Flötteröd (2015). Namely, the forward lag and the backward lag are explicitly implemented for each vehicle. We also benchmark the approach versus the analytical models of Osorio and Flötteröd (2015) (which is the most accurate, yet also the least efficient) and of Lu and Osorio (2017) (which is less accurate than Osorio and Flötteröd (2015), but more efficient). The validation experiments indicate that the proposed model captures the dynamics of the link boundaries correctly.

Table 1. Average runtime (in min) per iteration

| Initial points                          | Point 1 | Point 2 | Point 3 | Point 4 |
|-----------------------------------------|---------|---------|---------|---------|
| Benchmark model of Lu and Osorio (2017) | 144.98  | 146.14  | 144.37  | 149.38  |
| Proposed model                          | 1.27    | 1.26    | 1.28    | 1.28    |

We have also used the proposed model to address an optimization problem for a large-scale network, and have benchmarked it versus optimization results obtained from the model of Lu and Osorio (2017). We consider a signal control problem for the Swiss city of Lausanne, which is modeled as a set of 902 lanes and 231 intersections. The signal control problem optimizes the signal plans of 17 intersections distributed throughout the city. The objective is to minimize the average spillback probability. This average is computed over all lanes of the network and over time. The spillback probability of a given lane during time interval  $k$  corresponds to  $P(UQ(k) = \ell)$ . The optimal signal plans derived by each model (proposed and benchmark) are evaluated with a microscopic traffic simulation model of Lausanne. It is shown that the optimal signal plans proposed by both models have comparable objective function performance. They also yield similar average lane queue-lengths and average trip travel times. However, the gain in computational runtime of the proposed model is significant. Table 1 shows the average runtime in minutes per iteration of the optimization algorithm. Each column corresponds to statistics obtained when initializing the algorithms with different initial points. The proposed model improves efficiency by 2 orders of magnitude, on average. Detailed numerical results from this case study will be presented at the conference.

### REFERENCES

- Chong, L. and Osorio, C. (2017). A simulation-based optimization algorithm for dynamic large-scale urban transportation problems, Accepted to Transportation Science . Available at: <http://web.mit.edu/osorioc/www/papers/choOso15dynSO.pdf> .
- Lu, J. and Osorio, C. (2017). A probabilistic traffic-theoretic network loading model suitable for large-scale network analysis, Accepted to Transportation Science . Available at: <http://web.mit.edu/osorioc/www/papers/luOso17.pdf> .
- Morse, P. (1958). Queues, inventories and maintenance; the analysis of operational systems with variable demand and supply, Wiley, New York, USA.
- Newell, C. (1982). Applications of queueing theory, Chapman and Hall, New York, USA.
- Newell, G. (1993). A simplified theory of kinematic waves in highway traffic, part I: general theory, Transportation Research Part B 27(4): 281–287.
- Osorio, C. and Flötteröd, G. (2015). Capturing dependency among link boundaries in a stochastic

dynamic network loading model, *Transportation Science* 49(2): 420–431.

Yperman, I., Tampere, C. and Immers, B. (2007). A kinematic wave dynamic network loading model including intersection delays, *Transportation Research Board Annual Meeting*, Washington DC, USA.

# INTERFACING A DYNAMIC DISCRETE CHOICE BASED TRAVEL DEMAND MODEL TO AN AGENT BASED TRAFFIC SIMULATION MODEL

Mohammad SALEEM<sup>a</sup>, Oskar Blom VÄSTBERG<sup>b</sup> and Anders KARLSTRÖM<sup>c</sup>

<sup>a</sup> *Department of Urban Planning and Environment, KTH Royal Institute of Technology, Sweden*

*Email: [salee@kth.se](mailto:salee@kth.se)*

<sup>b</sup> *Department of Urban Planning and Environment, KTH Royal Institute of Technology, Sweden*

*Email: [oskar.vastberg@abe.kth.se](mailto:oskar.vastberg@abe.kth.se)*

<sup>c</sup> *Department of Urban Planning and Environment, KTH Royal Institute of Technology, Sweden*

*Email: [anders.karlstrom@abe.kth.se](mailto:anders.karlstrom@abe.kth.se)*

This paper presents the ongoing development of SCAPER, which is a random utility based discrete choice travel demand model, capable of consistently incorporating time decisions throughout the day. It discusses how computational speed of SCAPER is improved using an importance based sampling of locations, and how it influences the simulation results. The paper also discusses interfacing SCAPER to MATSim simulation framework, and estimating the model with travel times, travel distances and travel costs produced by the Stockholm demand (simulated) over Stockholm network using MATSim. The paper also presents preliminary results, identifies limitations and highlights future directions.

**Keywords:** SCAPER, MATSim, Dynamic Discrete Choice, Travel Demand Model, Traffic Simulation

## 1. INTRODUCTION

This paper presents the interface of a dynamic discrete choice model of travel demand (SCAPER) with an agent based traffic simulation model (MATSim). In SCAPER, individuals are assumed to make decisions sequentially in time about whether to stay or travel, starting from home in the morning and ending at home in the evening. The choice is made based on a random utility, calculated as the sum of one-stage utility of an action and the expected future utility in the reached state. In previous work, an estimation procedure has been proposed using sampling of alternative travel patterns (Västberg et. al., 2016). The model was estimated based on recorded travel diaries and the subsequent simulation results indicated that the model was able to reproduce departure times, trip lengths, number of trips distribution and mode shares with in the sample very well.

In order to use SCAPER for forecasting in combination with a traffic simulation model, there were two main problems that had to be solved: Firstly, the model is computationally very demanding to evaluate. Simulating a daily travel pattern for one individual takes close to 10 seconds when 1240 locations (zones) are considered for each trip. Secondly, the level-of-service attributes used for estimation are not the same as the level-of-service attributes obtained from the traffic simulator. It would be desirable if the modelling system would reproduce observed behavior from travel diaries in the fixed point with estimated parameters.

To speed up the simulation process, this paper proposes using “sampling of locations”. A simple MNL model is estimated from which locations are sampled for each individual. The resulting travel patterns are compared with observed behavior to determine the number of locations that has to be sampled in order to reach an acceptable level of bias. To avoid the discrepancy in travel times reported in travel diaries, the proposed model uses travel times, travel distances and travel costs generated through a dynamic traffic flow simulation model for the city of Stockholm, using MATSim.

## 2. MODEL

Demand for travel comes from demand for activity participation. With this realization, the travel demand models have evolved recently from trip and tour based models into activity based model. The problem with activity based models is that if one considers a full day activity schedule; there are a huge number of possible ways to plan a day. Although one could exclude certain unnecessary alternatives, but it is not straightforward, as the choice of activities is influenced by the preference for time at a given time. This means, for example, that the decision to select a time to leave for work in the morning is effectively

dictated by the preference of staying home in the morning versus the preference of staying home in the evening, provided the work schedule is flexible. Most of the activity based models create full day activity schedules, without taking the dependency of preference of time over time of day. This is where they fall short. For an overview of activity based models, one could read Pinjari and Bhat (2011) or Rasouli and Timmermans (2014).

All aspects of travel pattern are interconnected, and so a correct representation of time is crucial in activity based models as some aspects of time are fixed and cannot be violated. The choice of travel decision can be explained as trade-offs to spend a limited amount of available time, and it makes sense to assume that travelers consider full day travel pattern, while deciding what activities to engage in. If a state  $s_t$  represents a location and time of day, and an action  $a_t$  defines activity type, duration and the mode of transport that moves the traveler to a new state  $s_{t+1}$ ; then a full day path actually is a sequence of actions starting in the morning and ending in the evening. The travel times vary, and hence add some kind of randomness to the model. States are also only partially visible, as unexpected needs or opportunities may also arise, adding further stochasticity to the model. A rational agent following a policy  $\pi$  in a stochastic environment, starting in a state  $s_t$ , will take action  $a_t$  in state  $s_t$ , that maximizes the expected future utility:

$$V(s_0) = \max_{\pi} E_s \left\{ \sum_{t=0}^T \beta^t u(s_t, a_t) \mid s_0 = s \right\} \quad (1)$$

Where  $u(s_t, a_t)$  is a one-stage utility function and  $\beta$  is the discount factor. We assume  $\beta$  as 1, and assume  $u(s_t, a_t)$  to be additively separable into  $u(x_t, a_t) + \varepsilon_t$ . We assume that a state  $s_t$  can be separated in  $(x_t, \varepsilon_t)$ , such that  $\varepsilon_t$  is Gumbel distributed and  $\varepsilon_t \sim G(-\gamma, 1)$ , meaning the mean of  $\varepsilon$  is 1 rather than  $\gamma$ . The error term  $\varepsilon_t$  captures all uncertainty of the system. It is i.i.d. over alternatives and time. With these assumptions, the expected value function becomes:

$$EV(x_t) = \int V(s_{t+1}) p(ds_{t+1} | s_t, a_t) = \log \left( \sum_{a \in A(x_t)} e^{u(x_t, a) + EV(x_{t+1})} \right) \quad (2)$$

And the probability that an alternative  $a$  is chosen when in state  $x$ , can be given by:

$$P(a_t | x_t) = e^{u(a_t, x_t) + EV(x_{t+1}) - EV(x_t)} \quad (3)$$

For an individual  $n$ , the one stage utility can be obtained by adding the utility of performing the activity and (dis)utility of travelling to the activity location. The travel disutility is in turn dependent on the mode, time of day, origin and destination. The utility of performing an activity is calculated from a constant utility for starting the activity and a duration based utility.

During simulation, each individual is assumed to consider all possible locations for each new action, meaning that the computational time increases quadratically with number of locations. Therefore, it makes sense to sample locations for computational benefit, using an auxiliary model based on importance sampling. Similar approach has been employed in the past by numerous authors e.g. Bradley et. al. (2010) and Liao et. al (2013). We sample locations according to weights obtained through a MNL model, given below.

$$P(l) = \frac{e^{\ln(L_l \cdot \theta_L) + \ln(E_l \cdot \theta_E) + D_{H,l} \cdot \theta_{DH} + D_{W,l} \cdot \theta_{DW} + \ln(D_{H,l} \cdot \theta_{DH}) + \ln(D_{W,l} \cdot \theta_{DW})}}{\sum_{l=1}^N e^{\ln(L_l \cdot \theta_L) + \ln(E_l \cdot \theta_E) + D_{H,l} \cdot \theta_{DH} + D_{W,l} \cdot \theta_{DW} + \ln(D_{H,l} \cdot \theta_{DH}) + \ln(D_{W,l} \cdot \theta_{DW})}} \quad (4)$$

Where  $L_l$  represents number of people living at location  $l$ ,  $E_l$  represents number of people employed at location  $l$ ,  $D_{H,l}$  represents distance to home from location  $l$  and  $D_{W,l}$  represents distance to work from location  $l$ ; while  $\theta_L$ ,  $\theta_W$ ,  $\theta_{DH}$  and  $\theta_{DW}$  represent the corresponding estimated parameters.

### 3. INTERFACING SCAPER WITH MATSIM

As the details about MATSim can be found elsewhere (Horni *et al.*, 2016), only a brief overview is provided. MATSim is an activity-based multi-agent traffic simulation framework. It uses a microscopic description of travel demand and performs fast mesoscopic simulation of traffic flows and the congestion resulting from those traffic flows. For this work, the demand used by MATSim is synthetic in nature, but based on real census data, hence fairly representative of the actual Stockholm demand. Each traveler of the synthetic demand has one or more travel plans for the day, representing its intentions for the day. In order to do a performance-based comparison of plans, MATSim allocates a real-valued utility value to each plan. The utility is calculated through positive contributions by performing utility maximizing activities e.g. work, leisure etc., and negative contribution by travelling. We have replaced the default utility (scoring) functions of MATSim according to the utility functions of SCAPER, in order to add consistency to the integration approach. In SCAPER, travel times and distances are aggregate at zone level, while in MATSim travel times and distances are disaggregate (at a granularity of individual locations). We aim to remove this inconsistency in our future work.

Figure 1 shows the overall architecture of the integration of MATSim and SCAPER. Based on the initial generated synthetic demand, MATSim generates initial level-of-service matrices; based on which, Scaper generates travel patterns consisting of number of trips as well as mode, destination, departure time and purpose for all trips. MATSim then simulates the demand with SCAPER generated travel patterns, and iterates route choice until an approximate stochastic equilibrium is achieved, such that adjusting routes does not yield further improvements on average. At equilibrium, new level-of-service matrices are calculated, which are then fed to SCAPER to generate new travel patterns. MATSim simulates car and PT traffic only. Every other mode is teleported with a pre-specified speed per mode. For the simulation of PT traffic, we use the information provided by Stockholms Läns Landsting (SLL) about transit lines, transit routes, transit vehicles and transit time table.

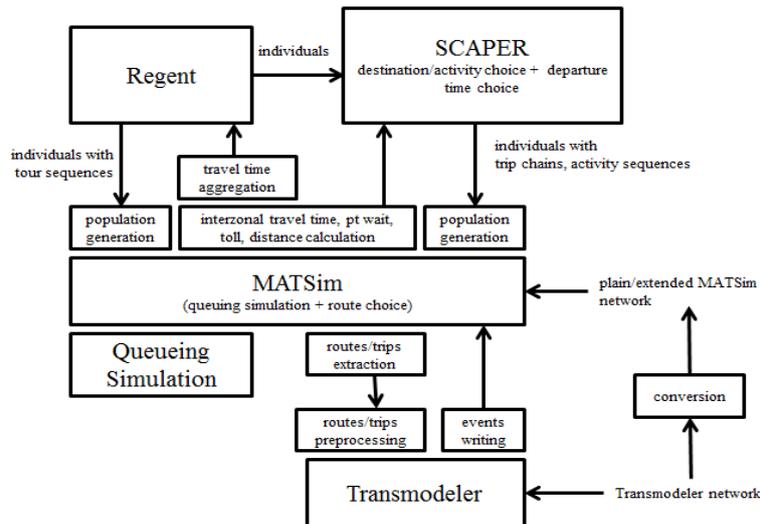


Figure 1. Overall Architecture

The interactive iteration between SCAPER and MATSim continues until convergence is achieved. We assume convergence in SCAPER-MATSim interaction when the difference between different attributes of alternatives in observed data and simulated data do not change systematically anymore and only stochastic variations exist.

#### 4. RESULTS

We estimated the model based on Stockholm travel survey from 2004, in which the users have reported full day travel diaries. The estimation results indicate that most parameters are significant and have the expected sign. Cost is negative indicating that it is preferable to spend time participating in activities as compared to travelling. Home time is valued high early in the morning and late in the night. Jointly with estimation, we also simulated the Stockholm car owning demand (61,000 sampled individuals). For simulation, we have sampled 80 locations out of 1240 locations. It creates a bias to the system, but the bias is within an acceptable threshold. For all important attributes, the bias is under 2 percent. We also assume that 80 is a reasonable number of locations available to any individual on any given day for performing activities. The computational benefit achieved by sampling 80 locations out of 1240 locations is that each individual is handled (simulated) 130 times faster. The simulation results indicate that the overall system (SCAPER-MATSim interaction) converges within five iterations. This is indicated by the rather huge difference between the observed values and simulated values in the early iterations, which gets stable after 3 to 5 iterations and only stochastic variations exist after that, as shown in Figure 2.

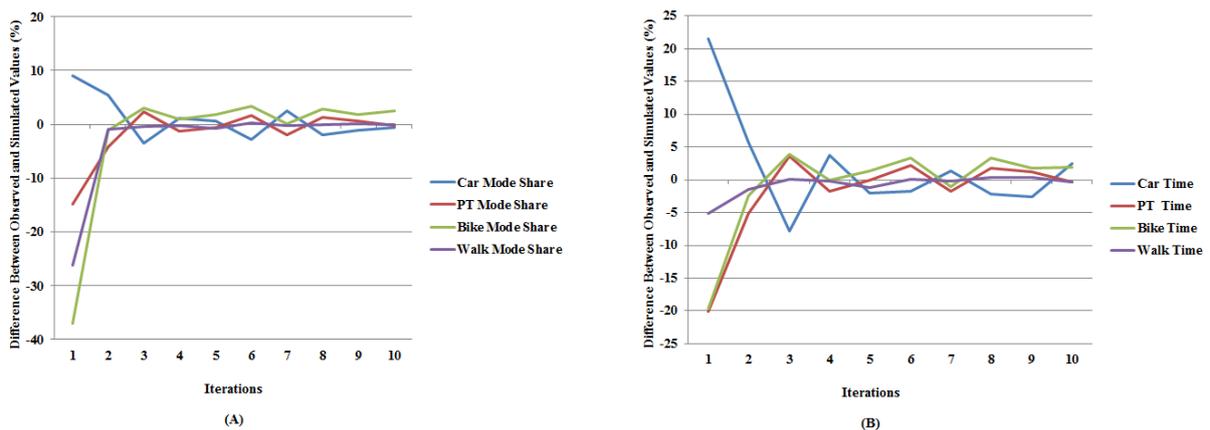


Figure 2. Difference Between Observed Values and Simulated Values Over Iterations

Table 1 shows the values for mode shares and mode times, both for the observed data and simulated data. As one can see, in the converged setting, the difference is very small, indicating the simulated results are able to reproduce the observed behavior.

Table 1. Result Statistics

| Quantity of Interest | Observed Values | Simulated Values | Difference (%) |
|----------------------|-----------------|------------------|----------------|
| Car Mode Share       | 1.03681         | 1.01677          | 1.93257        |
| PT Mode Share        | 1.05715         | 1.07582          | -1.76564       |
| Bike Mode Share      | 0.242816        | 0.240421         | 0.986139       |
| Walk Mode Share      | 0.35712         | 0.35848          | -0.380805      |
| Car Time             | 14.2581         | 14.0202          | 1.66883        |
| PT Time              | 29.3087         | 29.8187          | -1.74013       |
| Bike Time            | 5.23743         | 5.22748          | 0.190012       |
| Walk Time            | 9.85045         | 9.89375          | -0.439575      |

We also estimated the model based on Stockholm travel survey from 2006, in which too the users have reported full day travel diaries. The survey also contains congestion toll, as introduced in central parts

of Stockholm. The application of toll has some interesting effects. It is observed that Car mode share goes down and PT mode share goes up; this is intuitively justifiable since people switch to public transport to avoid paying toll. We also observe a shift in departure times. Figure 3 shows the departure time curves for with and without toll settings. When toll is applied, the departure time curve shifts a little to avoid peak hours in morning and evening. This happens as people tend to leave before the peak hour when toll is the highest. This again is justifiable intuitively as the toll is highest during the peak hours. The total number of departures by car decreases by 6%, when toll is applied. The number of passes through the toll cordons decreases by 45%. Overall people are less happy (average utility difference of 1.98%) if toll is applied, due to the extra financial burden over people moving through central Stockholm. This difference is small, as the financial burden is compensated by an improvement in travel conditions.

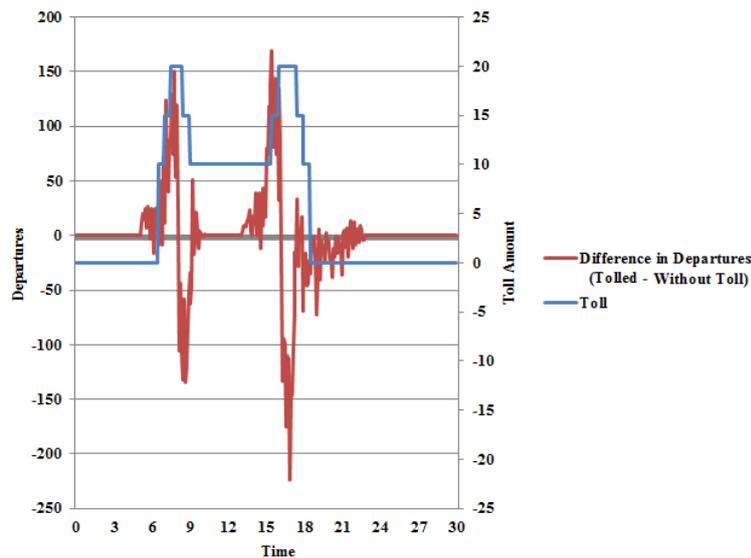


Figure 3. Difference in Number of Departures between With-toll and Without-toll Setting

The results could be improved by feeding PT travel times and PT wait times produced in MATSim to SCAPER. Right now we only feed travel times, travel distances and travel costs for Car traffic. For PT and Bike traffic we use data recorded in travel surveys. We are working on feeding PT travel times, PT wait times, and PT costs to SCAPER, to improve the results further.

## REFERENCES

- Blom-Västberg, O., Karlström, A., Jonsson, D. and Sundberg, M. (2016) *Including time in a travel demand model using dynamic discrete choice*. University Library of Munich, Germany.
- Bradley, M., Bowman, J.L. and Griesenbeck, B. (2010) SACSIM: An applied activity-based model system with fine-level spatial and temporal resolution. *Journal of Choice Modelling*, 3(1), pp. 5-31.
- Horni, A., Nagel, K. and Axhausen, K.W. (2016) *The multi-agent transport simulation MATSim*. London: Ubiquity Press.
- Liao, F., Arentze, T. and Timmermans, H. (2013) Incorporating space-time constraints and activity-travel time profiles in a multi-state supernetwork approach to individual activity-travel scheduling. *Transportation Research Part B: Methodological*, 55(0), pp. 41-58.
- Pinjari, A.R. and Bhat, C.R. (2011) Activity-based travel demand analysis. *A Handbook of Transport Economics*, eds A. De Palma, R. Lindsey, E. Quinet, and R. Vickerman, pp. 213-248.
- Rasouli, S. and Timmermans, H. (2014) Activity-based models of travel demand: Promises, progress and prospects. *International Journal of Urban Sciences*, 18(1), pp. 31-60.

# A PROBABILISTIC SENSITIVITY ANALYSIS GUIDED CROSS ENTROPY METHOD FOR EFFICIENT CALIBRATION OF TRAFFIC MODELS

X. X. XIE<sup>a</sup>, S. W. HU<sup>a</sup>, K. Y. FU<sup>b</sup>, R. X. ZHONG<sup>a,\*</sup>

<sup>a</sup> *Guangdong Provincial Key Laboratory of Intelligent Transportation Systems, School of Engineering, Sun Yat-sen University, Guangzhou, China.*

*Email: [xiexx6@mail2.sysu.edu.cn](mailto:xiexx6@mail2.sysu.edu.cn)*

*Email: [husw3@mail2.sysu.edu.cn](mailto:husw3@mail2.sysu.edu.cn)*

*\*Email: [zhrenxin@mail.sysu.edu.cn](mailto:zhrenxin@mail.sysu.edu.cn)*

<sup>b</sup> *Shenzhen Urban Transport Planning Center, Shenzhen, China.*

*Email: [fukaiky@sutpc.net](mailto:fukaiky@sutpc.net)*

Calibration is a vital procedure before traffic models can be used as tools to accurately simulate traffic flow behaviors. In general, a common objective for model calibration is to minimize the difference between model simulation and observed data. However, such a method may result in many local optima if the underlying traffic models is of complex structure. Moreover, as the model complexity increases, the required computational effort to accurately calibrate traffic model also significantly increases. To tackle this, a probabilistic sensitivity analysis (PSA)-guided-cross-entropy method (CEM) calibration framework is proposed to improve the efficiency of calibration of traffic models. The proposed framework is applied to calibrate the cell transmission model using the data of U.K. M25 Motorway. Results illustrate that this framework can achieve an optimum with approximately a quarter of simulation evaluations of the original CEM. This confirms its great potentials.

**Keywords:** Probabilistic Sensitivity Analysis, Cross Entropy Method, Calibration of Traffic Model

## 1. INTRODUCTION

Calibration is a vital procedure before traffic simulation models can be used as tools to accurately represent traffic flow behaviors. In general, the common objective for model calibration is to minimize the difference between model simulation and observed data. However, such a method may result in a large number of local optima due to the complex structures of the optimization problems and the number of different combinations of the set of parameters. Moreover, as the model complexity increases, the required computational effort to accurately calibrate traffic model also significantly increases.

Since there are many local optima during calibration due to the complex structures of models and the number of different parameters' combinations, the random search methods are regarded as popular and convenient methods to compensate it (Ngoduy and Maher, 2012; Ciuffo and Punzo, 2014; Zhong et al., 2014, 2016). Spiliopoulou et al. (2015) showed that Nelder-Mead algorithm, the genetic algorithm and cross-entropy methods are able to converge to robust model parameters set with different performances. Ngoduy and Maher (2012) applied the cross-entropy method (CEM) to calibrate a second order macroscopic model, with the purpose of solving the continuous multi-external optimization problem. Zhong et al. (2016) proposed a calibration framework based on the CEM and the Probabilistic Sensitivity Analysis (PSA) and used it to calibrate the intelligent driving model which evidenced the CEM could find an optima sufficiently close to the global solution. As a global sensitivity tool, PSA based on the cross entropy distance have recently gained significant attention to analyze the probabilistic density function of the objective function (Liu et al., 2006; Lüdtke et al., 2008; Majda and Gershgorin, 2010, 2011). In the PSA, the cross entropy distance is used to calculate the difference between two probability density functions to consider the cases within a heavy-tail and the nonlinear or non-Gaussian distribution (Zhong et al., 2016). However, to obtain relatively satisfying calibration results, the CEM and PSA techniques require a large number of simulation evaluations and hence a heavy computational burden.

To this end, a PSA-guided-CEM calibration framework is proposed to improve the efficiency of calibration of Cell Transmission Model (CTM) based traffic models. The PSA and the CEM are integrated into a unified framework since both them apply the cross-entropy distance. In the former, the cross entropy distance is applied to measure the difference between two probability density functions (PDFs) to identify the important and unimportant parameters. Then, it is applied in CEM to converge

the samples around the importance samples in each iteration. Along this framework, reducing the dimension of calibration problem by the results of PSA is the main strategy to improve the efficiency of calibration with the CEM.

## 2. MODEL FORMULATION

The goal of the calibration with CTM is to determine the parameters of fundamental diagram to minimize the difference between model simulation and observed data. There are five parameters including the free flow speed  $v_f$ (mph), the maximum allowable flow  $Q_M$ (vph), the critical density  $k_c$ (vpm), the backward congestion wave speed  $v_j$ (mph), and the jam density  $k_j$ (vpm) to calibrate in CTM. The decisive parameters of CTM only contain three of the five since the other two parameters can be calculated based on the relationship of the triangle limitations in fundamental diagram. In this paper,  $v_f$ ,  $v_j$ , and  $Q_M$  are considered as the decision parameters since these three parameters determine the slope and vertex of the triangle fundamental diagram. Moreover, a mean average percentage error is used as an objective function to measure the difference between the simulation density and the observed density, shown as below:

$$O(x) = \frac{1}{C_N} \sum_{i=1}^{C_N} \sum_{t=1}^{T_N} \left| \frac{\hat{\rho}_i(t) - \rho_i(t)}{\rho_i(t)} \right| \quad (1)$$

where  $O(x)$  denotes the mean average percentage error of all cells with parameters  $x = [v_f, v_j, Q_M]$ ,  $\hat{\rho}_i(t)$  is the simulation density of cell  $i$  in time interval  $t$ ,  $i = 1, 2, \dots, C_N$ , and  $t = 1, 2, \dots, T_N$ ,  $\rho_i(t)$  denotes the observed density.

## 3. THE PROBABILISTIC SENSITIVITY ANALYSIS

As previously discussed, the estimation of parameters of the CTM is subject to various issues such as observability, perturbations, uncertainty or errors. Hence their sensitivity analysis is a crucial question in determining which parameter directions are the most or least sensitive to perturbations, uncertainty and data errors. This is essential for the selection of the most important parameters and the identification of their probabilistic characteristics (Zhong et al., 2016). PSA approach evaluates the impact of a random variable on the performance index function by measuring the K–L distance between two probability density functions of the performance index function, obtained before and after the variation reduction of the chosen random variable. For discrete probability distributions  $p_k$  and  $q_k$ , the K–L distance of  $q$  from  $p$  is defined to be:

$$D_{KL}(p||q) = \sum_{k=1}^K p_k \ln \frac{p_k}{q_k} \quad (2)$$

The K–L distance is connected to the cross entropy  $H(p, q)$  (for the distributions  $p$  and  $q$ ) through the following relationship:

$$H(p, q) = H(p) + D_{KL}(p||q) \quad (3)$$

where  $H(p)$  is the entropy of  $p$ .

## 4. THE CROSS-ENTROPY METHOD

The cross-entropy method (CEM) is a general Monte Carlo approach to combinatorial and continuous multi-extremal optimization and importance sampling. The CEM approach can be broken down into two key steps:

1. Generate a number of trial parameter sets randomly according to the chosen distributions.
2. Based on the values of the objective function associated with each trial parameter set, update the probability distribution used to generate the random trial sets according to the principle of “importance sampling”.

The estimation via the importance sampling is described as follows. Consider the general problem of estimating the quantity:

$$l = E_u[H(X)] = \int H(x)f(x; u)dx \quad (4)$$

where  $H$  is some performance function and  $X$  is the random variable.  $f(x; u)$  denotes a member of some parametric family of PDFs, parameterized by  $u$ . Using importance sampling this quantity can be estimated as:

$$\hat{l} = \frac{1}{N} \sum_{i=1}^N H(X_i) \frac{f(X_i; u)}{g(X_i)} \quad (5)$$

where  $X_1, \dots, X_N$  is a random sample from  $g$ . For positive  $H$ , the theoretically optimal importance sampling density (PDF) is given by

$$g^* = H(x)f(x; u)/l \quad (6)$$

This, however, depends on the unknown  $l$ . The CEM aims to approximate the optimal PDF by adaptively selecting members of the parametric family that are closest (in the Kullback–Leibler sense) to the optimal PDF  $g^*$ .

## 5. THE PSA-GUIDED-CEM CALIBRATION FRAMEWORK

The connection between the CEM and the PSA is the cross entropy distance, which is generally used to measure the difference between two PDFs. The CEM applies the cross entropy distance to update the distribution of the next iteration by minimizing the distance between the important sample and the current sample while the PSA applies the cross entropy to measure the difference between the original distribution and derived distribution, before and after uncertainty reduction of the random variable. Figure 1 shows the procedure of the PSA-guided-CEM calibration framework. The first step is to perform the PSA to identify the important and unimportant parameters of the overall model. The second step is to define the model with only the important influence parameters as a simplified model and then use CEM to obtain the Pareto optimum solution sufficiently close to the global optimum value. The last step is to calibrate the full model with the Pareto solution as inputs. In this way, reducing the size of problem by the PSA for calibrating with the simplified models may improve the calibration efficiency. On the other hand, calibrating the full model around the Pareto solution may ensure a more accurate solution, since the dependency among parameters would be beneficial to finding the global optimal solution from the Pareto optimal solutions.

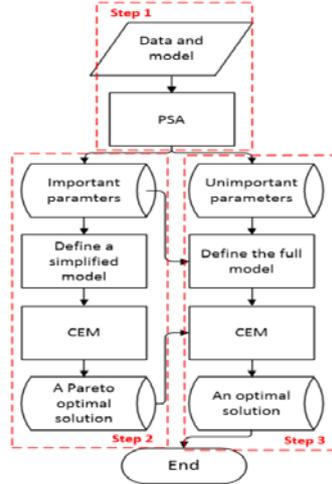


Figure 1: The procedure of PSA-guided-CEM calibration framework

## 6. EMPIRICAL STUDY

The proposed PSA-guided-CEM calibration framework can be used to calibrate the CTM. The study segment is located on U.K. M25 Motorway between Junctions 14 and 16 with clockwise direction, and the total length is 20 km. The motorway is split into six subsections and each section is analyzed separately. According to the first step in the PSA-guided-CEM calibration framework, the important degree sorting parameters are  $Q_M$ ,  $v_j$  and  $v_f$ . Hence, the important parameters combinations of  $Q_M$  and  $v_j$  will first be calibrated while keeping  $v_f$  to its mean value. When the important parameter combinations converges to a certain extent,  $v_f$  will be calibrated by adding its uncertainty to model to obtain the overall solution.

The calibration performances of Original Cross Entropy Method (OCEM), Modified Cross Entropy

Method (MCEM) and PSA-CEM are compared in terms of the Best Objective Value (BOV) and the Simulation Evaluations (SE). The BOV measures the difference between simulation and observed data while the SE records the number of samples to present the efficiency of each method. With the chosen traffic data, Table 1 records the mean performance among 5 calibration repetitions in each subsection using OCEM, MCEM, and PSA-CEM respectively. The results of all subsections indicate the MCEM takes less SE than the OCEM. However, the PSA-CEM takes the least SE, almost the half of the OCEM takes, to achieve the similar BOV in each subsection. The reason may be that reducing the dimension of model before the calibration takes less SE (the number of samples) to cover the whole uncertainty of the objective function.

Table 1: The mean performance of three different methods in terms of Best Objective Value (BOV) and Simulation Evaluations (SE).

|             |     | OCEM   | MCEM   | PSA-CEM |
|-------------|-----|--------|--------|---------|
| Subsection1 | SE  | 40000  | 20516  | 13688   |
|             | BOV | 11.61% | 11.60% | 11.58%  |
| Subsection2 | SE  | 44400  | 20955  | 15924   |
|             | BOV | 9.70%  | 9.70%  | 9.83%   |
| Subsection3 | SE  | 33400  | 18276  | 12662   |
|             | BOV | 11.52% | 11.58% | 11.62%  |
| Subsection4 | SE  | 50800  | 27091  | 18245   |
|             | BOV | 10.04% | 10.01% | 10.33%  |
| Subsection5 | SE  | 50800  | 36604  | 24282   |
|             | BOV | 9.67%  | 9.45%  | 9.84%   |
| Subsection6 | SE  | 47600  | 26676  | 18089   |
|             | BOV | 8.09%  | 8.07%  | 8.24%   |

The empirical study confirms the great potentials of the PSA-guided-CEM calibration framework. The PSA-CEM is used to calibrate the macroscopic traffic models with the identification of important parameters combinations and the fitting curve of repeating samples. In comparison with other methods, the PSA-CEM achieves a similar calibration optimum but with the least simulation evaluations. As a result, the PSA-CEM is a systemic and efficient calibration framework for calibration of traffic models.

## REFERENCES

- Ngoduy, D., & Maher, M. J. (2012). Calibration of second order traffic models using continuous cross entropy method. *Transportation Research Part C: Emerging Technologies*, 24, 102-121.
- Ciuffo, B., & Punzo, V. (2014). "No Free Lunch" Theorems Applied to the Calibration of Traffic Simulation Models. *IEEE Transactions on Intelligent Transportation Systems*, 15(2), 553-562.
- Zhong, R., Chen, C., Chow, A. H., Pan, T., Yuan, F., & He, Z. (2016). Automatic calibration of fundamental diagram for first-order macroscopic freeway traffic models. *Journal of Advanced Transportation*, 50(3), 363-385.
- Zhong, R. X., Fu, K. Y., Sumalee, A., Ngoduy, D., & Lam, W. H. K. (2016). A cross-entropy method and probabilistic sensitivity analysis framework for calibrating microscopic traffic models. *Transportation Research Part C: Emerging Technologies*, 63, 147-169.
- Spiliopoulou, A., Papamichail, I., Papageorgiou, M., Tyrinopoulos, Y., & Chrysoulakis, J. (2017). Macroscopic traffic flow model calibration using different optimization algorithms. *Operational Research*, 17(1), 145-164.
- Ciuffo, B., & Azevedo, C. L. (2014). A sensitivity-analysis-based approach for the calibration of traffic simulation models. *IEEE Transactions on Intelligent Transportation Systems*, 15(3), 1298-1309.
- Liu, H., Chen, W., & Sudjianto, A. (2006). Relative entropy based method for probabilistic sensitivity analysis in engineering design. *Journal of Mechanical Design*, 128(2), 326-336.
- Lüdtke, N., Panzeri, S., Brown, M., Broomhead, D. S., Knowles, J., Montemurro, M. A., & Kell, D. B. (2008). Information-theoretic sensitivity analysis: a general method for credit assignment in complex networks. *Journal of The Royal Society Interface*, 5(19), 223-235.

Majda, A. J., & Gershgorin, B. (2010). Quantifying uncertainty in climate change science through empirical information theory. *Proceedings of the National Academy of Sciences*, 107(34), 14958-14963.

**Session B8**  
**Public Transport II**

# COMMUTER'S DEPARTURE TIME CHOICE WITH FARE-REWARD SCHEME IN A MULTI-TO-SINGLE MASS TRANSIT SYSTEM

Yili TANG<sup>a,b</sup>, Hai YANG<sup>a</sup> and Liu CHENZHE<sup>a</sup>  
<sup>a</sup> *Department of Civil and Environmental Engineering,  
The Hong Kong University of Science and Technology, China*  
<sup>b</sup> *Email: [ytangap@connect.ust.hk](mailto:ytangap@connect.ust.hk)*

This paper introduces a fare-reward scheme to manage a commuter's departure time choice in a multi-to-single transit system, which aims at incentivizing commuter's shift in departure time to the shoulder periods of the peak period to relieve queueing congestion in transit stations. A framework of the commuter departure time equilibrium model in a multi-to-single mass transit system is provided. A bi-level model with fare-reward scheme (FRS) is then introduced that rewards a commuter with one free trip during shoulder periods after a certain number of paid trips during the peak hours. The upper-level model is to minimize the total system trip costs by determining the optimal reward ratio (the number of paid trips required for one free trip) and the fare intervals. The lower-level model is the commuter departure time equilibrium model for both commuters with rewards and without rewards. Our study indicates that with the optimal reward ratio 0.42 in a numerical example, the FRS results in a reduction of 41.46% in queueing time cost and decreases 8.99% total system time costs.

**Keywords:** Departure Time Choice, Equilibrium, Fare-Reward Scheme, Mass Transit, Queueing Congestion

## 1. INTRODUCTION

Peak-hour transit crowding and congestion have been increasingly concerned in metropolitan areas. Starting from 1950s (Vickrey, 1955; Vuchic, 1969), there has been a substantial stream of development of research in transit system optimization. Notably, Mohring (1972) developed a microeconomic model of transit system to optimize the transit services with fixed demand. Existing studies on transit service optimization and demand management are generally investigated with considerations on transit capacity choice, scheduling and fare pricing. Transit authorities usually adjust service frequency to accommodate variable passenger demand. Applying the bottleneck model in mass transit, Kraus and Yoshida (2002) considered optimal fare and service frequency to minimize long term system cost. There are few studies focus on the model of commuter departure equilibrium in transit system. de Cea and Fernández (1993) investigated the equilibrium assignment model in a limited transit line network. To consider the in-vehicle crowding cost, Huang et al. (2004) modelled the equilibrium of urban mass transit system during peak period. In line with the work, Tian and Huang (2007) explored equilibrium properties on a mass transit system considering rigid capacity constraint.

Various fare pricing schemes are also considered to manage the peak period congestion involving peak-fare charging, off-peak discount and their combination. Metropolitan areas such as Washington DC and London have introduced time-based fares with peak and off-peak fare difference. Moreover, several incentive-based fare-free policies demonstrated the reduction of peak ridership. For instance, the Melbourne's 'Early Bird' Scheme allows commuters travel free to the central business district if they finish the trip before 7 a.m. Surveys show 23% of off-peak ridership is shifted from rush-hour period by an average of 42 minutes (Currie, 2010). However, a fundamental problem of the above pricing strategies is to offer monetary compensation or other incentives at the expenses of the government or the transit operator. The discontinued fare-free strategy for off-peak hours in Mercer County, New Jersey reports a loss of around one-fourth of revenues from the experiments (Perone, 2002). With limited source of funding, governments or transit authorities would rather charge a high fare to passengers than avoid revenue losses.

Recently, Yang and Tang (2018) proposed a novel fare-reward scheme (FRS) for managing peak-hour congestion in a single origin-destination urban rail transit bottleneck. Under the proposed FRS, a commuter is rewarded one free trip during pre-specified shoulder periods after a certain number of paid trips during the peak period. Namely, a certain proportion of commuters will be incentivized to shift

their departure time to shoulder periods during daily peak hours and hence passenger flow congestion in the transit system will be smoothed out over time. The planner can determine the reward ratio and free fare intervals to reduce commuters' trip costs and control the number of free commuters.

This paper makes one important step forward to the work of Yang and Tang (2018) by implementing the fare-reward scheme in a many-to-one network and solve for the commuting equilibrium of departure time choices. Section 2 introduces the framework of many-origin and one-destination transit system with capacity constraints and problems of departure time equilibrium with a uniform fare. Section 3 investigates a bi-level model to solve for the optimal fare-reward scheme and fare intervals for reducing the congestions in the system. Section 4 provides the numerical examples and concludes.

## 2. THE MODEL IN THE ABSENCE OF THE FARE-REWARD SCHEME

As show in Figure 1, we consider an urban rail line with multiple origins and a single destination where each train is running with capacity  $s$  and headway  $h$ . During the peak period, a total number of  $N$  commuters travel with  $M$  uniformly service runs through the rail line. A train departs from the furthest station  $S_1$  and stop by  $S_2 \dots S_K$  stations to the destination  $D$ . The commuter demand in each station is denoted by  $N_k$ ,  $k = 1, 2, 3 \dots K$ . The in-vehicle time between station  $k$  and station  $k+1$  is denoted by  $d_k$ .

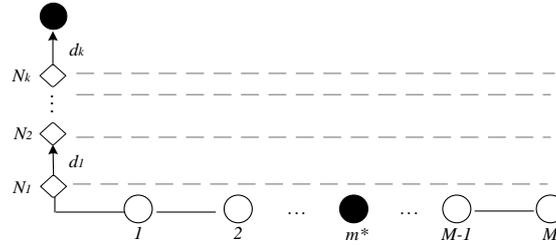


Figure 1. The multi-to-single transit system

Individual commuters are assumed to be homogeneous and have an identical working start time which is the departure time of service run  $m^*$ . In this case, a commuter who takes a service run labeled as  $m$  at station  $k$  encounters a uniform fare cost  $p_0$ , a queueing time  $q_m^k$ , a crowding cost  $g_m^k$ , an in-vehicle time  $t^k$ , a schedule delay  $e_m$  if they arrive at work earlier or  $l_m$  if they arrive late. With homogeneous commuters, the generalized trip costs inclusive of time cost, fare cost and crowding cost is given by

$$TTC_m^k = q_m^k + c_m^k + \alpha t^k + \beta e_m + \gamma l_m + p_0 \quad (1)$$

The unit cost  $\alpha$ ,  $\beta$  and  $\gamma$  are the shadow values of arriving early and arriving late respectively. As in Tian et al. (2007), the crowding cost  $c_m^k$  is related to the density of the commuters taking the same service run and the in-vehicle time. Hence,  $c_m^k$  can be written as

$$c_m^k = \sum_{t=k}^K g \left( \sum_{z=1}^t n_m^z \right) d_t \quad (2)$$

where  $n_m^z$  is the number of commuters from station  $S_z$  taking service run  $m$ .  $g(\cdot)$  is the crowding cost function which is a monotonically increasing function of the density of the commuters in the same train.

In equilibrium, commuters from the same station have the same generalized cost given by (1). Commuters from different stations may have different equilibrium costs. The equilibrium commuter departure time distribution can be derived by solving the following minimization problem

$$TTC(\mathbf{n}) = \sum_{t=1}^K \left( \sum_{m \in M} G \left( \sum_{z=1}^t n_m^z \right) \right) d_t + \sum_{m \in M} \left( \sum_{k=1}^K n_m^k \right) (\delta(m) + T^k) + \sum_{m \in M} \sum_{k=1}^K n_m^k \cdot p_0 \quad (3)$$

$$\sum_{m \in M} n_m^k = N_k, k = 1, 2, \dots, K \quad (4)$$

$$\sum_{t=1}^K n_m^t \leq s, m = 1, 2, \dots, M \quad (5)$$

$$n_m^k \geq 0, = 1, 2, \dots, M \quad (6)$$

The first term of the objective function (3) is the integral of crowding cost function which has no economic interpretation. The second term  $\delta(m)$  is the aggregate schedule delay cost and  $T^k$  is the in-vehicle time of commuters in the transit system respectively. The third term is the transit revenue. Constraint (4) is the demand constraint for each station and constraint (5) represents the rigid capacity. The commuter departure time equilibrium is given by the first order condition of the problem (3).

### 3. THE MODEL IN THE PRESENCE OF THE FARE-REWARD SCHEME

Under the fare-reward scheme (FRS), one assigned free trip is rewarded to commuters during shoulder period after certain paid journeys within peak period. A reward ratio  $\lambda$  is defined to control the free trips which is the number of free trips to the total number of the trips. It is also equivalent to the ratio of the number of rewarded commuters to the total number of commuters on each day during the peak period. The peak period of interest is divided into two free fare intervals (FFI) and one uniform fare interval (UFI). UFI is the central period within the peak period spanning the time interval  $[t_i, t_j]$  that contains service runs from  $i$  to  $j$  and includes the work starting time  $t^*$ , as shown in Figure 2. FFI includes the two shoulder intervals before and after the UFI.

The commuters' departure time choices depend on the original fare, the reward ratio and the UFI determination. Commuters with rewards are defined as free ride commuters while commuters without rewards are defined as non-free ride commuters. We assume the free ride commuters and non-free ride commuters travel in FFI and UFI respectively, as shown in Figure 2.

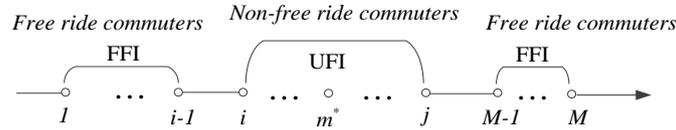


Figure 2. Commuters travel choice after the fare-reward scheme.

The model with the fare-reward scheme is a bi-level optimization. The upper-level model is to determine the optimal reward ratio and UFI to minimize the system total trip costs, specified as follows:

$$\min_{\lambda, UFI} \sum_{k \in K} u_{FFI}^k (1 - \lambda) N_k + u_{UFI}^k \lambda N_k \quad (7)$$

$$(1 - \lambda) N p = N p_0 \quad (8)$$

$$u_{UFI}^k + p \geq u_{FFI}^k \quad (9)$$

where the equilibrium trip costs for free ride commuters  $u_{FFI}^k$  and for non-free ride commuters  $u_{UFI}^k$  can be obtained by solving the lower-level model which is the commuter departure distribution model as illustrated in Section 2. Constraint (8) represents the revenue neutrality for transit authority. Constraint (9) represents the incentive for commuters with reward which indicates that the trip costs of traveling in FFI is no greater than those in UFI.

The upper-level model is the transportation planning problem associated to the determination of the reward ratio and the UFI, with the policy maker's objective of optimizing the system performance in total trip costs. The lower-level model is the transit assignment model for the commuter departure time

distribution.

#### 4. RESULTS AND CONCLUSION

In this section, we turn to numerical examples to investigate the system performance with the characteristics shown in Table 1. We consider an urban rail transit line with 5 origins and 1 destination. With reference to MTR Corporation Limited (2015), a set of parameters transit service has been applied to evaluate the system performance of the fare-reward scheme and compare it with the results of original transit bottleneck model. The unit costs of schedule delay early and late are chosen with due considerations of the relative values and Hong Kong value of time (Census and Statistics Department, 2017; Small et al., 2005; Ubbels et al., 2005). The uniform original fare is 11 HKD for all stations.

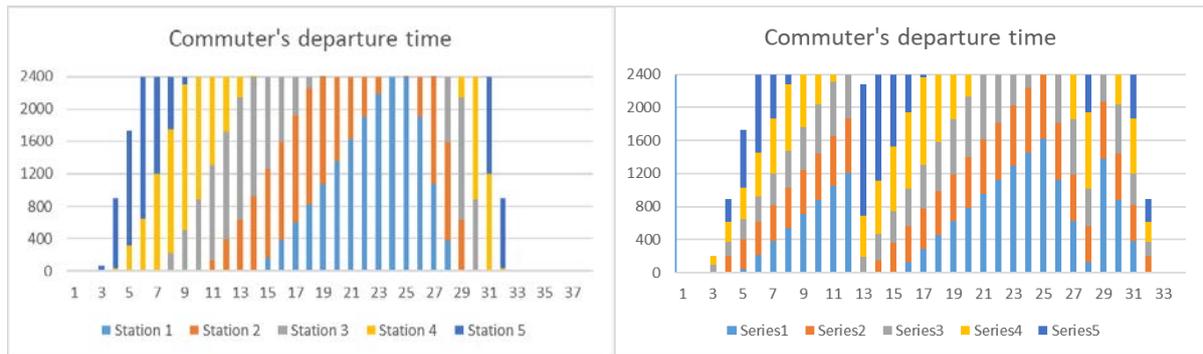
Table 1. Transit system and fare-reward scheme characteristics

| Parameter<br>s | K | M  | h<br>(hour) | $p_0$<br>(HKD) | $(\beta, \gamma)$<br>(HKD/h) | $d_1 = d_2$<br>(hour) | $N_i$<br>$\times 10^3$          | $g(n)$<br>(HKD/h) | $s$<br>$\times 10^3$ |
|----------------|---|----|-------------|----------------|------------------------------|-----------------------|---------------------------------|-------------------|----------------------|
| Value          | 5 | 46 | 2/60        | 11             | (65,195)                     | 0.4167                | 18.3/14.6/<br>12.9/12.1/8.<br>1 | 150/s             | 2.4                  |

The commuter departure time distribution before and after the fare-reward scheme is showed in Figure 3. The system performances are summarized in Table 2. Results indicate the optimal reward ratio is 0.42 with the optimal UFI from 13<sup>th</sup> – 28<sup>th</sup> service runs. The total system time costs decrease 8.99% and total queuing time costs decrease 41.46%. The total trip costs decrease 4.14% which is smaller than the system time costs because of the unchanged fare costs.

Table 2. System performance with the fare-reward scheme

| Terms      | Total trip costs | Total time costs | Total crowding costs | Total queueing costs | Total schedule delay costs | $\lambda^*$ | UFI                                 |
|------------|------------------|------------------|----------------------|----------------------|----------------------------|-------------|-------------------------------------|
| Efficiency | 4.14%            | 8.99%            | 10.96%               | 41.46%               | -0.16%                     | 0.42        | 13 <sup>th</sup> – 28 <sup>th</sup> |



(a)

(b)

Figure 3. Commuter departure time distribution (a) without FRS and (b) in the presence of FRS

The preliminary results in the numerical example reveal interesting findings. Furthermore, the equilibrium properties will be examined by looking into the relationship of departure time choice of commuters from different stations, and the system performance with respect to the original fare with be investigated. The existing and further analysis will be useful to better understand and explore the new fare-reward scheme in transit systems.

## REFERENCES

- Census and Statistics Department. (2017). 2016 report on annual earnings and hours survey. Hong Kong: Retrieved from <https://www.statistics.gov.hk/pub/B10500142016AN16B0100.pdf>.
- Currie, G. (2010). Quick and effective solution to rail overcrowding: Free early bird ticket experience in Melbourne, Australia. *Transportation Research Record*, (2146) 35-42. doi:10.3141/2146-05
- de Cea, J., and Fernández, E. (1993). Transit assignment for congested public transport systems: An equilibrium model. *Transportation Science*, 27(2), 133-147.
- Huang, H., Tian, Q., Yang, H., & Gao, Z. (2004). Modeling the equilibrium bus riding behavior in morning rush hour. *Proceedings of the Ninth Annual Conference of the Hong Kong Society of Transportation Studies*. Hong Kong, 434-442.
- Kraus, M., and Yoshida, Y. (2002). The commuter's time-of-use decision and optimal pricing and service in urban mass transit. *Journal of Urban Economics*, 51(1), 170-195.
- Mohring, H. (1972). Optimization and scale economies in urban bus transportation. *The American Economic Review*, 62(4), 591-604.
- MTR Corporation Limited. (2015). Annual report. (). Hong Kong: Retrieved from <http://www.mtr.com.hk/en/corporate/investor/2015frpt.html>
- Perone, J. S. (2002). Advantages and disadvantages of fare-free transit policy . ( No. NCTR-473-133).National Center for Transit Research, Center for Urban Transportation Research, University of South Florida.
- Small, K. A., Winston, C., and Yan, J. (2005). Uncovering the distribution of motorists' preferences for travel time and reliability. *Econometrica*, 73(4), 1367-1382.
- Tian, Q., & Huang, H. (2007). Commuting equilibria on a mass transit system with capacity constraints. *Transportation and Traffic Theory*.
- Tian, Q., Huang, H., and Yang, H. (2007). Equilibrium properties of the morning peak-period commuting in a many-to-one mass transit system. *Transportation Research Part B: Methodological*, 41(6), 616-631.
- Ubbels, B., Tseng, Y., & Verhoef, E. T. (2005). Value of time, schedule delay and reliability-estimates based on choice behaviour of Dutch commuters facing congestion. *ERSA Conference Papers*, (ersa05p202)
- Vickrey, W. (1955). Some implications of marginal cost pricing for public utilities. *The American Economic Review*, 45(2), 605-620.
- Vuchic, V. R. (1969). Rapid transit interstation spacings for maximum number of passengers. *Transportation Science*, 3(3), 214-232.
- Yang, H., and Tang, Y. (2018). Managing rail transit peak-hour congestion with a fare-reward scheme. *Transportation Research Part B: Methodological*, 110, 122-136.

# BUS SCHEDULE OPTIMIZATION FOR MULTIMODAL EMERGENCY EVACUATIONS

Xia YANG <sup>a</sup>, Xiaozheng (Sean) HE <sup>b</sup> and Xuegang (Jeff) BAN <sup>c</sup>

<sup>a</sup> *Department of Civil Engineering, SUNY Polytechnic Institute, US*

<sup>b</sup> *Department of Civil and Environmental Engineering, Rensselaer Polytechnic Institute, US*  
Email: [hex6@rpi.edu](mailto:hex6@rpi.edu)

<sup>c</sup> *Department of Civil and Environmental Engineering, University of Washington, Seattle, US*  
*College of Transport and Communications, Shanghai Maritime University, Shanghai, China*  
Email: [banx@uw.edu](mailto:banx@uw.edu)

Modeling emergency evacuation could help provide guidance on evacuation planning and mitigate the consequences of disasters. In the era of connected and autonomous vehicles, evacuation planners and operators will be able to dynamically and systematically guide the evacuations for the benefit of the entire system. This paper focuses on modeling bus schedule optimization in multimodal emergency evacuations while assuming that buses are given priorities in evacuations and have fixed link travel times. The model also considers evacuation priorities by giving different weights to the waiting times at different origins based on their emergency agitations. Finally the proposed evacuation model is formulated as a mixed integer programming by discretizing it over time. Some numerical experiments are then followed to investigate the proposed model and the solution algorithm.

**Keywords:** Bus Schedule Optimization, Dynamic Traffic Assignment, System Optimum, Multi-modal Emergency Evacuation, Mixed Integer Linear Programming

## EXTENDED ABSTRACT

Modeling emergency evacuation could help provide guidance on evacuation planning and mitigate the consequences of disasters by identifying the minimum time required for evacuations, the network operational capacities, the bottleneck locations, and the utilization of the multimodal transportation system (Abdegawad et al., 2010; Yuan and Puchalsky, 2014; Yang et al., 2017). There have been abundant emergency evacuation models developed in literature; an in-depth review can be found in Murray-Tuite and Wolshon (2013), from which it can be found that those models vary in the assumptions made on evacuees' choices of whether to evacuate or not (Murray-Tuite and Wolshon, 2013), the departure times (US Army Corps of Engineers, 1999; Lindell 2008; Pel et al., 2012), the evacuation modes (Renne et al. 2009; White et al. 2008; Murray-Tuite and Wolshon, 2013; Yuan and Puchalsky, 2014), the evacuation destinations (Chen, 2004; Cheng and Wilmot, 2009; Yuan et al, 2006), and the evacuation routes (Pel et al., 2012).

The emerging connected and autonomous vehicles will enable evacuation planners and operators to dynamically and systematically guide the evacuations (e.g., by influencing the evacuees' travel choice behavior) for the benefit of the entire system. This motivates researchers to study evacuations based on dynamic system optimum (DSO) principle. The double-queue-based macroscopic model has recently gained popularity in recent evacuation models (Ma et al., 2014b; Yang et al. 2017) because of its capability of capturing real traffic dynamics such as queue spillback as well as its advantage of computation efficiency over the meso- or micro-scopic models (Osorio et al., 2011; Osorio, 2015; Ma et al., 2014a). In the DSO-based evacuation modeling, Ma et al. (2014b) only considered cars/autos as the evacuation mode while Yang et al. (2017) extended the study by considering multiple transportation modes and their interactions, capacities of the transportations system and accessible shelters. However, Yang et al. (2017) assumed that the buses follow the regular peak-hour routes and schedules for the sake of customers' familiarities and bus operation convenience.

To further improve emergency evacuation efficiency, this paper focuses on optimizing bus schedule in multimodal emergency evacuations as it would be practical in the era of connected and autonomous vehicles. It is also assumed that buses will be given priorities in evacuations so their road link travel times are fixed. As a result, the node-based flow model can be applied to model the bus-based flow

dynamics and bus capacities. The proposed model also considers the evacuation priorities by giving different weights to the waiting times at different origins based on their emergency agitations. Discretized over time the proposed evacuation model is then formulated as a mixed integer programming. Finally some numerical experiments are conducted to investigate the proposed model and the solution algorithm.

## REFERENCES

- Abdegawad, H., Abdulhai, B., and Wahba, M. (2010) Multiobjective optimization for multimodal evacuation. *Transportation Research Record: Journal of the Transportation Research Board*, 2196, pp. 21-33.
- Chen, X. and Zhan, F.B. (2004) Agent-based modeling and simulation of urban evacuation: relative effectiveness of simultaneous and staged evacuation strategies. *Journal of the Operational Research Society*, 59(1), pp. 25-33.
- Cheng, G. and Wilmot, C.G. (2009) Time-dependent travel cost impact on hurricane evacuation destination choice models. *Proceedings of the 88th Annual Meeting Transportation Research Board*, Washington, DC: TRID.
- Lindell, M.K. (2008) EMBLEM2: An empirically based large scale evacuation time estimate model. *Transportation Research Part A: Policy and Practice*, 42, pp. 140–154.
- Ma, R., Ban, X. and Pang, J.S. (2014a) Continuous-time dynamic system optimal for single-destination traffic networks with queue spillbacks. *Transportation Research Part B: Methodological*, 68, pp. 98-122.
- Ma, R., Earle, B., Wetmore, S. and Ban, X. (2014b) Evacuation modeling using dynamic system optimal principle: A case study. *Proceedings of the 5th International Symposium on Dynamic Traffic Assignment*, Salerno, Italy: DTA Symposium.
- Murray-Tuite, P. and Wolshon, B. (2013) Evacuation transportation modeling: An overview of research, development, and practice. *Transportation Research Part C: Emerging Technologies*, 27, pp. 24-44.
- Osorio, C. and Flotterod, G. (2015) Capturing dependency among link boundaries in a stochastic dynamic network loading model. *Transportation Science*, 49(2), pp. 420-431.
- Osorio, C., Flotterod, G. and Bierlaire, M. (2011) Dynamic network loading: a stochastic differentiable model that derives link state distributions. *Transportation Research Part B: Methodological*, 45(9), pp. 1410–1423.
- Pel, A.J., Bliemer, M.C.J. and Hoogendoorn, S.P. (2012) A review on travel behavior modelling in dynamic traffic simulation models for evacuations. *Transportation*, 39(1), pp. 97–123.
- Renne, J.L., Sanchez, T.W., Jenkins, P. and Peterson, R. (2009) Challenge of evacuating the carless in five major US cities: identifying the key issues. *Transportation Research Record: Journal of the Transportation Research Board*, 2119, pp. 36–44.
- US Army Corps of Engineers (1999). Northwest Florida hurricane evacuation study technical data report. Washington, DC. Retrieved from: [https://coast.noaa.gov/hes/docs/hes/NW FLORIDA HES TECHNICAL DATA REPORT.pdf](https://coast.noaa.gov/hes/docs/hes/NW_FLORIDA_HES_TECHNICAL_DATA_REPORT.pdf)
- White, R.A. et al. (2008) The role of transit in emergency evacuation. Washington, D.C.: *Transportation Research Board of the National Academies*.
- Yang, X. (2016). Modeling multimodal transportation network emergency evacuation considering evacuee' cooperative behavior. *Transportation Research Procedia*, 23, pp. 380–399.
- Yuan, F. and Puchalsky, C.M. (2014) Multimodal evacuation simulation and scenario analysis in dense urban area: a Philadelphia case study. *Transportation Research Record: Journal of the Transportation Research Board*, 2532, pp. 91-98.
- Yuan, F., Han L., Chin, S.M. and Hwang, H. (2006) Proposed framework for simultaneous optimization of evacuation traffic destination and route assignment. *Transportation Research Record: Journal of the Transportation Research Board*, 1964, pp. 40-48.

**Session C1**  
**Within-Day Equilibrium II**

# A LINK-BASED PROBIT MODEL FOR DYNAMIC TRAFFIC ASSIGNMENT

Genaro PEQUE, Jr<sup>a</sup>, Junji URATA<sup>b</sup> and Takamasa IRYO<sup>c</sup>

<sup>a,b,c</sup> *Department of Civil Engineering,  
Kobe University, Japan*

<sup>a</sup> *Email: [gpequejr@panda.kobe-u.ac.jp](mailto:gpequejr@panda.kobe-u.ac.jp)*

<sup>b</sup> *Email: [urata@person.kobe-u.ac.jp](mailto:urata@person.kobe-u.ac.jp)*

<sup>c</sup> *Email: [iryu@kobe-u.ac.jp](mailto:iryu@kobe-u.ac.jp)*

We propose a link-based multinomial probit (MNP) model for large-scale dynamic traffic assignment simulation where link travel times are perturbed by random errors from unknown distributions. Additionally, link correlations, calculated proportional to the length of the congestion spillback, are also considered. Using the central limit theorem, the repeated sampling of the additive link travel times for each path at each iteration produces an approximation for the route-based MNP model.

**Keywords:** Probit Model, Multivariate Normal Distribution, Central Limit Theorem

## 1. INTRODUCTION

In the multinomial probit (MNP) model, a trip maker  $i$  chooses a path  $k$  from his/her set of paths  $K$  to travel to his/her destination. The trip maker  $i$  assigns a corresponding (perceived) utility,  $U^i$ , for each of these paths; typically the negative value of the path travel time,  $U_k^i = -T_k$ . The utility is defined as,

$$U_k^i = V_k^i + \varepsilon_k^i, \forall i \in I, \forall k \in K, \quad (1)$$

where  $V^i$  is a systematic utility and  $\varepsilon^i$  is a normally distributed random term with zero mean and an arbitrary covariance matrix, multivariate  $N(0, \Sigma^i), \Sigma^i = [\text{cov}(\varepsilon_k^i, \varepsilon_j^i)], \forall i \in I, j \neq k, (j, k) \in K \times K$ . Since the MNP choice probabilities cannot be expressed in closed form, there are many ways to approximate it (Sheffi and Powell, 1981). Compared to multinomial logit (MNL) models, the MNP model can theoretically address its major deficiencies, such as unrealistic overestimation of flows on highly overlapped paths, underestimation of flows on more independent paths and the effect on absolute values of utilities (i.e. a 10-minute difference in a network with two paths has the same effect for a short or long trip), at the cost of analytical tractability. Its application was first formulated by Daganzo and Sheffi (1977), Sheffi and Powell (1981), Sheffi and Powell (1982) and further developed by McFadden's (1989) method of simulated moments.

In Sheffi and Powell (1982), the MNP covariance structure is formulated by pertinently relating the matrix elements to the network topology which means there isn't a need to explicitly evaluate and maintain covariance matrices. Link travel times,  $T_l$ , were modeled as both random and flow-dependent variables. The actual link travel times,  $t_l$ , were perturbed by random errors perceived by the trip makers. More formally, the link travel times are defined as,

$$T_l \sim N(t_l, \theta t_l^0), \forall l \in L, \quad (2)$$

where the mean,  $\mu$ , is given by the actual link travel time,  $t_l$ , and the variance,  $\sigma^2$ , is composed of the link  $l$ 's free flow travel time,  $t_l^0$ , and a constant,  $\theta$ . The vector of path travel times,  $T_k$ , is multivariate normally distributed with moments,

$$\mathbb{E}[T_k] = t_k = \sum_l t_l \delta_{l,k}, \forall k \in K, \quad (3a)$$

and

$$\text{cov}[T_k, T_j] = \theta \sum_l t_l \delta_{l,k} \delta_{l,j}, \forall j, k \in K, \quad (3b)$$

where  $\delta_{a,b}$  is an element of the link-path incidence matrix which is equal to 1 if the link  $l$  belongs to path  $k$  and 0, otherwise.

Thus, the path choice probability can be expressed as,

$$P_k = \Pr \left[ T_k = \min_{j \in K} T_j \right], \forall j, k \in K. \quad (4)$$

Using the MNP model, it is possible to get negative draws from the normal distribution (e.g. the normal distribution extends from the negative infinity to the positive infinity). Hence, studies by Nielsen (2000), Ramming (2002), Bierlaire and Frejinger (2005) and Prato and Bekhor (2006) avoided this by either truncating the negative draws or approximating the normal distribution using a different distribution.

## 2. PROPOSED MODEL

### 2.1 A link-based Probit Model

In this paper, we allow for different types of distributions to avoid negative draws from a normal distribution. More formally, link travel times are defined as,

$$T_l = t_l + \text{Dist}(0, \hat{\theta}_l d_l), \quad (5)$$

where  $\text{Dist}(\cdot)$  is an unknown distribution with zero mean,  $d_l$  is the link length and  $\hat{\theta}_l$  is a constant. The variance-covariance matrix is represented by,

$$\Sigma = \hat{\theta}_l \begin{pmatrix} d_l^1 & \dots & d_l^{1|L|} \\ \vdots & \ddots & \vdots \\ d_l^{|L|1} & \dots & d_l^{|L|} \end{pmatrix}, \quad (6)$$

where  $|L|$  is the number of links and  $d_l^m : \hat{l} \rightarrow \mathbb{R}$  is a function related to the maximum length of the congestion spillback,  $\hat{l}$ , from link  $m$  to link  $l$  and is updated after each network loading.

At this time, the calculation of the variance-covariance is straightforward,

- (a) Produce a vector of  $K$  random draws which represents the variance of each link
- (b) Link correlations are calculated using a “backward-step” from the destination noting the maximum length of the spillback to each of the downstream links until the origin is reached.

### 2.2 Approximation using the Central Limit Theorem

To approximate the route-based MNP model using the link-based MNP model, we rely on the central limit theorem (CLT) which establishes that, in most situations, when independent random variables are added, their properly normalized sum tends toward a normal distribution even if the original variables themselves are not normally distributed. The CLT is central to this research since the summation of non-stable distributions produce a mixture distribution. However, a repeated sampling from this mixture distribution tends toward a normal distribution (figure 1).

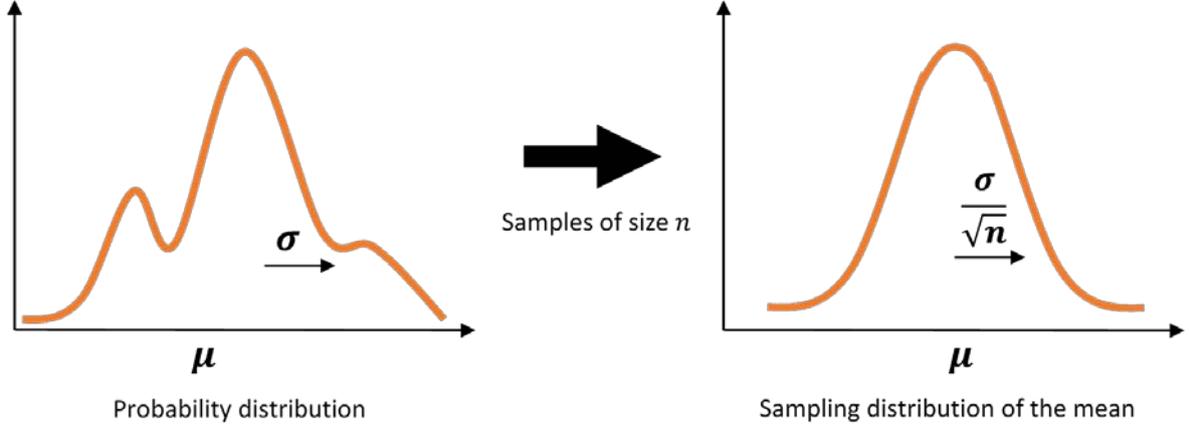


Figure 1. Whatever the form of the probability distribution (left), the sampling distribution tends to a normal distribution (right), and its dispersion is given by the central limit theorem

The classical CLT requires that random variables are independently and identically distributed (i.i.d.) but we want to assume that link random variables are not i.i.d. There are two CLT's that satisfy our requirement (given that they comply with certain conditions) that distributions converge to Normal distributions as the number of samples approach infinity, namely, the Lyapunov CLT and the Lindeberg CLT.

Let,

$$s_n = \sum_{i=1}^n \sigma_i^2, \quad (7)$$

where  $\sigma_i^2 = \text{Var}[T_i]$  and  $\mu_i = \mathbb{E}[T_i]$ .

The Lyapunov CLT requires that for some  $\alpha > 0$ , if Lyapunov's condition,

$$\lim_{n \rightarrow \infty} \frac{1}{s_n^{2+\alpha}} \sum_{i=1}^n \mathbb{E}[|T_i - \mu_i|^{2+\alpha}] = 0, \quad (8)$$

is satisfied, then the sum of  $\frac{T_i - \mu_i}{s_n}$  converges in distribution to a standard normal random variable as  $n \rightarrow \infty$ ,

$$\frac{1}{s_n} \sum_{i=1}^n (T_i - \mathbb{E}[T_i]) \xrightarrow{d} N(0,1). \quad (9)$$

In this paper, the weaker Lindeberg condition is used. Using the same notations above, the Lindeberg condition which requires that for every  $\varepsilon > 0$ ,

$$\lim_{n \rightarrow \infty} \frac{1}{s_n^2} \sum_{i=1}^n \mathbb{E}[\mathbb{I}_{\{|T_i - \mu_i| > \varepsilon s_n\}} (T_i - \mu_i)^2] = 0, \quad (10)$$

where  $\mathbb{I}_{\{\bullet\}}$  is the indicator function. Then the distribution of the standardized sums converges towards the standard normal distribution.

## REFERENCES

- Bierlaire, M and Frejinger, E. (2005) Route choice models with subpath components. *Proceedings of the 5th Swiss Transport Research Conference*, Ascona, Switzerland
- Daganzo, C. and Sheffi, Y. (1977) On stochastic models of traffic assignment. *Transportation Science*, Vol.11, pp.253-274
- Haghani, M., Shahhoseini, Z., and Sarvi, M. (2015) Probit-Based Traffic Assignment: A Comparative Study between Link-Based Simulation Algorithm and Path-Based Assignment and Generalization to Random-Coefficient Approach. *International Journal of Transportation Engineering*, Volume 2-3, pp.179-198
- McFadden, D. (1989) A method of simulated moments for estimation of discrete response models without numerical integration. *Econometrica* 57, 995-1026
- Nielsen, O. (2000) A stochastic transit assignment model considering differences in passengers utility functions. *Transportation Research Part B*, 34(5), 377-402
- Prato, C. and Bekhor, S. (2006) Applying branch and bound technique to route choice set generation, *Transportation Research Record*, 1985, 19-28
- Ramming, S. (2002) Network knowledge and route choice. Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, USA
- Sheffi, Y. and Powell, W. (1981) A comparison of stochastic and deterministic traffic assignment over congested networks. *Transportation Research*, Vol.15B, pp. 52-64
- Sheffi, Y. and Powell, W. (1982) An algorithm for the equilibrium assignment problem for random link times. *Networks*, 12(2), pp. 191-207

# A DYNAMIC SYSTEM OPTIMAL MODEL FOR HYBRID NETWORK OF RESERVOIRS AND CONNECTING LINKS

Qian GE<sup>a</sup>, Ke HAN<sup>b</sup>, Haoran FU<sup>c</sup> and Daisuke FUKUDA<sup>d</sup>

<sup>a</sup> *Department of Civil and Environmental Engineering,  
Tokyo Institute of Technology, Japan  
Email: [ge.qian@plan.cv.titech.ac.jp](mailto:ge.qian@plan.cv.titech.ac.jp)*

<sup>b</sup> *Department of Civil and Environmental Engineering,  
Imperial College London, U.K.  
Email: [k.han@imperial.ac.uk](mailto:k.han@imperial.ac.uk)*

<sup>c</sup> *Graduate School of Information Sciences,  
Tohoku University, Sendai, Japan.  
Email: [holanfu@gmail.com](mailto:holanfu@gmail.com)*

<sup>d</sup> *Department of Civil and Environmental Engineering,  
Tokyo Institute of Technology, Japan  
Email: [fukuda@plan.cv.titech.ac.jp](mailto:fukuda@plan.cv.titech.ac.jp)*

This paper focuses on dynamic system-optimal (DSO) traffic assignment problems on networks. The DSO is such an assignment strategy that the network-wide travel cost is minimized for a given table of origin-destination trip demands. Under certain conditions, the DSO can be formulated as linear programs in terms of links and nodes where vehicles' behavior could be described by state space models. Other formulations include mixed integer programs and nonlinear programs. However, these approaches suffer from the curse of dimensionality as the network size expands. The Macroscopic Fundamental Diagram (MFD) serves as an efficient modeling tool for large-scale network with simplified network flow dynamics while capturing the aggregate network density-throughput relationships. Typically, a MFD is used to model a region with relatively homogeneous traffic density; and, for large and inhomogeneous networks, a spatial segregation method is usually required. However, when performing clustering or partition method to general networks, the presence of certain links presents a problem as they could not be clustered into any reservoir. Intra-city streets are usually proximal in space and could be reasonably clustered, but inter-city arterials show different characteristics. When modeling the entire (multi-region) network, we seek a combination of both link model and reservoir model, the latter being governed by the MFD. We also several relaxation and reformulation schemes for hard constraints of the proposed DSO model by introducing special order set (SOS) variables. A numerical experiment is in progress to validate the proposed framework.

**Keywords:** Dynamic System Optimal, Hybrid Network, SOS Variables

## 1. INTRODUCTION

This paper focuses on dynamic system-optimal (DSO) traffic assignment problems on networks. The DSO is such an assignment strategy that the network-wide travel cost is minimized for a given table of origin-destination trip demands. There are a number of studies on DSO type assignment based on the seminal work of Ziliaskopoulos [2000], Peeta and Mahmassani [1995]. Under certain conditions, the DSO can be formulated as linear programs in terms of links and nodes where vehicles' behavior could be described by state space models. Other formulations include mixed integer programs and nonlinear programs. However, these approaches suffer from the curse of dimensionality as the network size expands. The Macroscopic Fundamental Diagram (MFD) serves as an efficient modeling tool for large-scale network with simplified network flow dynamics while capturing the aggregate network density-throughput relationships. Typically, a MFD is used to model a region with relatively homogeneous traffic density; and, for large and inhomogeneous networks, a spatial segregation method is usually required. There exist several studies on DSO for pure reservoir network [Yildirimoglu et al., 2015]. However, when performing clustering or partition method to general networks [Saeedmanesh and Geroliminis, 2016], the presence of certain links presents a problem as they could not be clustered into any reservoir. As shown in the example of Figure 1, intra-city streets are usually proximal in space and could be reasonably clustered, but inter-city arterials show different characteristics. When modeling the entire (multi-region) network, we seek a combination of both link model and reservoir model, the latter being governed by the MFD.

In our modeling framework, a network is viewed as a combination of reservoirs and connecting links, to be modeled separately. This study proposes a DSO model for networks consisting of reservoirs and connecting links.

## 2. METHODOLOGY

An example of hybrid network of reservoirs and connecting links is shown in Figure 2. A reservoir may have multiple entrance-exit (EE) pairs. Given the MFD of each reservoir, we can estimate the demand and supply for the reservoir. Meanwhile, link transmission model is applied to describe traffic dynamics for each connecting link. This problem is addressed in Section 2.1 following Geroliminis et al. [2013]. The cross-reservoir flow is determined by the traffic dynamics on the junctions in Section 2.2. Based on these dynamical models, we build a single destination system optimal dynamic traffic assignment model for hybrid reservoir/connecting link network in Section 2.3. Due to the nonlinearity of the MFD, Section 3 will discuss several relaxation and reformulation schemes for hard constraints of the proposed DSO model by introducing special order set (SOS) variables. Prospective results and issues are discussed in the Section 4.

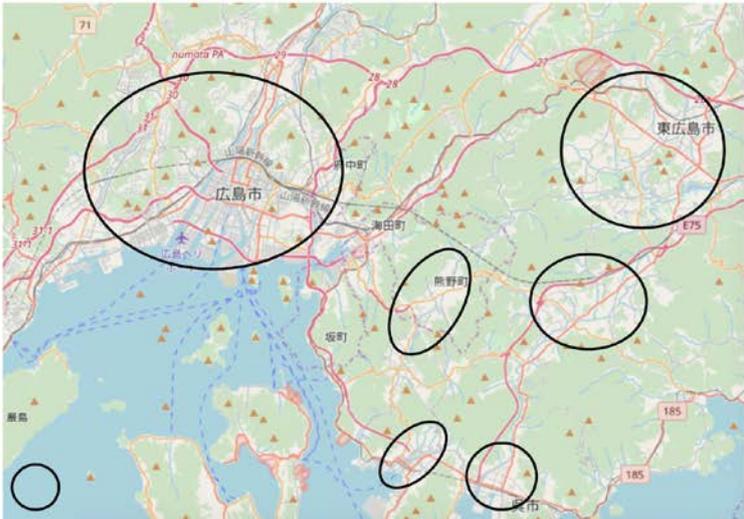


Figure 1. Reservoirs and connecting links (map of Hiroshima). Hiroshima consists of several urban areas connected by arterials. Each urban area could be viewed as a reservoir.

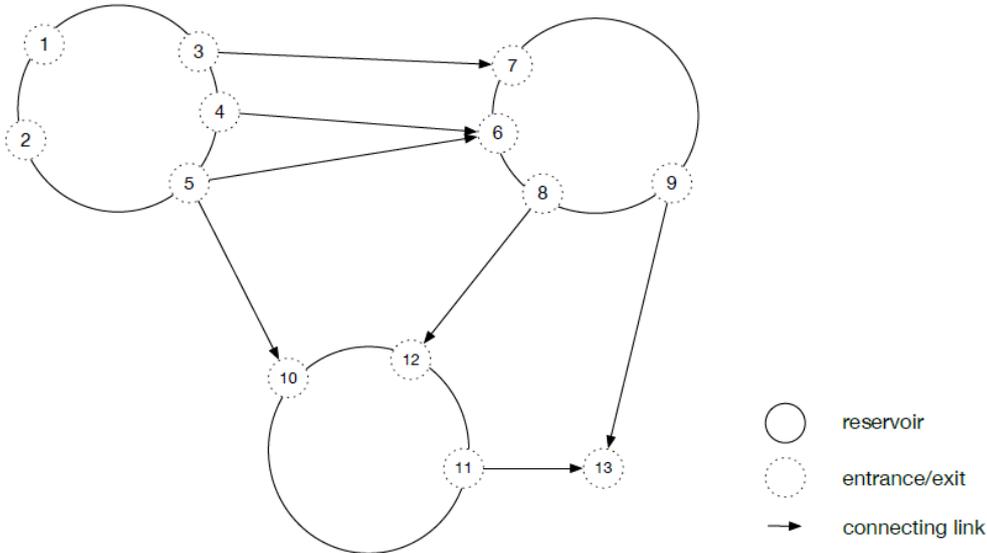


Figure 2. Hybrid network of reservoir and connecting links. 1, 2 are origin nodes, 13 is the destination node, 3-7, 4-6, 5-6, 5-10, . . . and are connecting links. Diverge and merge problem need to be considered at node 5 and 6 respectively.

## 2.1. Traffic dynamics in reservoirs and connecting links

The macroscopic fundamental diagram (MFD) describes the relationship between vehicle accumulation and flow in reservoirs (Daganzo [2007], Geroliminis and Daganzo [2008]). Consider a reservoir with  $i = 1, \dots, m$  en-trances and  $j = 1, \dots, n$  exits. This section builds a discrete model for the traffic dynamics in reservoirs following Geroliminis et al. [2013] and Yildirimoglu et al. [2015]. We take the heterogeneity of EE pairs into consideration when modeling the traffic in reservoir. More explicitly, we track the vehicle movement by the entrance and exit. Let  $t = 1, \dots, T$  be the time horizons and  $\Delta t$  be the time resolution. The vehicle accumulation of reservoir is  $n_r^t$ . The set of reservoir and connecting links are denoted as  $R$  and  $L$ . The upstream and downstream connecting links for any reservoir are  $L^u$  and  $L^d$ .  $o_1, o_2 \dots o_n$  and  $d$  are origins and destination. If there were no capacity constraint at the boundaries of reservoir, the traffic dynamics in reservoir  $r$  could be described as<sup>1</sup>:

$$n_r^{t+1} = n_r^t + \sum_{i \in I} q_{ir}^t - \sum_{j \in J} f_{jr}^t + \sum_{i \in I} f_{ir}^t \quad \forall i \in I \quad (16)$$

$$D_r^t = \min \{ G(n_r^t) \Delta t, \tilde{F}_r \Delta t \} \quad (17)$$

$$S_r^t = \tilde{N}_r - n_r^t \quad (18)$$

$$D_{jr}^t = \beta_j D_r^t \quad (19)$$

$$S_{ir}^t = \alpha_i S_r^t \quad (20)$$

$$\sum_j \beta_j = 1 \quad (21)$$

$$\sum_i \alpha_i = 1 \quad (22)$$

Moreover, we adopt the link transmission model (LTM) [Han et al., 2016, Yperman, 2007, Jin, 2015] to describe the traffic dynamics within each the connecting link:

$$n_l^{t+1} = n_l^t - f_{l.}^t + f_{.l}^t \quad \forall l \in L \quad (23)$$

$$\Delta t \sum_{k=1}^{t-\Delta t^w} f_{l.}^k - \Delta t \sum_{k=1}^t f_{.l}^k + \tilde{\rho}_l \xi_l > 0 \Rightarrow S_l^t = f_{l.}^{t-\Delta t^w} \quad (24)$$

$$\Delta t \sum_{k=1}^{t-\Delta t^w} f_{l.}^k - \Delta t \sum_{k=1}^t f_{.l}^k + \tilde{\rho}_l \xi_l = 0 \Rightarrow S_l^t = \tilde{F}_l \quad (25)$$

$$\Delta t \sum_{k=1}^{t-\Delta t^y} f_{.l}^k - \Delta t \sum_{k=1}^t f_{l.}^k > 0 \Rightarrow D_l^t = f_{.l}^{t-\Delta t^y} \quad (26)$$

$$\Delta t \sum_{k=1}^{t-\Delta t^y} f_{.l}^k - \Delta t \sum_{k=1}^t f_{l.}^k > 0 \Rightarrow D_l^t = \tilde{F}_l \quad (27)$$

In the previous equations,

- (1) Equation (1) represents the conservation of flow in reservoir where  $q_{ir}^t$  are OD demand from entrance  $i$ .  $f_{jr}^t$  is the flow out from the reservoir from exit  $j$ ,  $f_{ir}^t$  are flows enter the reservoir from  $i$ .
- (2) Equations (2) and (3) determines the demand and supply of reservoir based on MFD, where  $G(n_r^t)$  is vehicle accumulation-flux relationship in the reservoir.  $\tilde{F}_r$  is the boundary capacity of the reservoir.  $\tilde{N}_r$  is the bound of vehicle accumulation in reservoir.

<sup>1</sup> We only consider the inter-reservoir traffics in this study

- (3) Equations (4) and (5) assigns the supply and demand of reservoir to each entrance and exit where  $\alpha_i \geq 0$  and  $\beta_j \geq 0$  are predetermined ratio.
- (4) Equation (8) is the conservation of flow in connecting links, where  $f_l^t$  is the flow to downstream reservoirs and  $f_l^t$  is the flow from upstream reservoirs.
- (5) Equations (9) to (12) describe the link transmission model for a triangle fundamental diagram [Han et al., 2016] where  $f_{.l}^t$  and  $f_{.l}^t$  are inflow and outflow,  $D_l^t$  and  $S_l^t$  are demand and supply of link  $l$ ,  $v$  is the free flow speed,  $w$  is the backwards shock wave speed,  $\xi_l$  is the length of link.
- $$\Delta_l^v = \left\lceil \frac{\xi_l}{\Delta t v} \right\rceil \quad \text{and} \quad \Delta_l^w = \left\lceil \frac{\xi_l}{\Delta t w} \right\rceil. \quad [\bullet] \text{ is rounding operation.}$$

## 2.2. Diverge and merge problems for reservoir-link interaction

Though the previous equations could track the traffic dynamics in reservoirs and connecting link, the cross-boundary flow has not yet explored. Diverge and merge problem exist in the hybrid network, as shown in Figure 3. We could simply apply the results of Daganzo [1995] and Han et al. [2016] to calculate the flux since we have already known the demand and supply of reservoirs and connecting links.

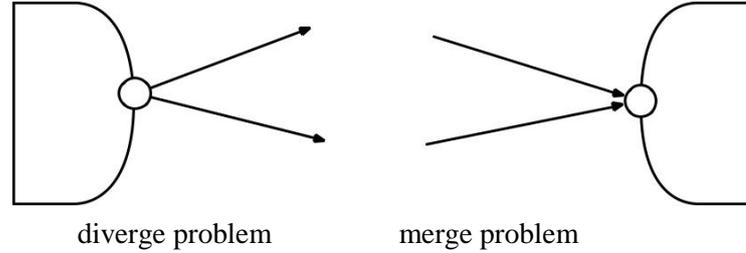


Figure 3. Diverge and merge problem. Diverge problem occurs when multiple connecting links are available for an exit, and merge problem occurs when flows on multiple links share the same entrance

For the exit node with multiple downstream links  $v \in V_d$ , we have the following solution:

$$f_{jr}^t = \min\{D_{jr}^t, \max\{\dots, \frac{S_l^t}{\gamma_{jl}} \dots\}\} \quad (28)$$

$$f_{jl}^t = \gamma_{jl} f_{jr}^t \quad (29)$$

$$\sum_l \gamma_{jl} = 1 \quad \forall l \in L^* \quad (30)$$

where  $\gamma_{jl} \geq 0$  are predefined distribution ratios.

For the entrance node with multiple upstream links  $v \in V_u$ , we have:

$$f_{ir}^t = \min\{\sum_{l \in L'} D_l^t, S_{ir}^t\} \quad (31)$$

$$f_{il}^t = \gamma_{il} f_{ir}^t \quad (32)$$

$$\sum_l \gamma_{il} = 1 \quad \forall l \in L' \quad (33)$$

where  $\gamma_{il} \geq 0$  are predefined turning ratios.

## 2.3. DSO problem for hybrid network

Based on previous discussions on the dynamical model in hybrid network, we present an optimization formulation of the single destination dynamic system optimal (SD-DSO) model based on Ziliaskopoulos [2000] and Nie [2011].

$$[\text{SD-DSO-HN}] \quad \min \sum_{t=1}^T \sum_{r \in R} n_r^t \Delta t + \sum_{t=1}^T \sum_{l \in L} n_l^t \Delta t \quad (34)$$

subject to:

$$\begin{aligned} & \text{Equations (1) to (7)} \quad \forall r \in R \\ & \text{Equations (8) to (12)} \quad \forall l \in L \\ & \text{Equations (13) to (15)} \quad \forall v \in V_d \\ & \text{Equations (16) to (18)} \quad \forall v \in V_u \\ & f_{lt}^t = D_l^t \quad \forall l, d \text{ on } l \end{aligned}$$

In this formulation, we minimize the dwelling time of each trip in reservoirs and connecting links and hence minimizes the total travel costs of all OD trips. However, this problem is a linear programming problem with hard constraints. The MFD relationship of vehicle accumulation and flux is usually nonlinear. The constraints imposed by the LTM are IF-THEN constraints. The next section will discuss relaxation schemes and reformulation techniques for this problem.

### 3. SOLUTION

The hard constraints in the [SD-DSO-HN] is could be relaxed and reformulated by SOS variables [Beale and Forrest, 1976]. A Special Ordered Set of type One (SOS1) variable is set of variables where only one of them could be non-zero in a feasible solution. SOS1 variables could help reformulate min / max constraints in [SD-DSO-HN]. A Special Ordered Set of type Two (SOS2) variable is set of variables where only two adjacent members of them could be non-zero in a feasible solution. SOS2 variables could help reformulate piecewise linear function constraints.

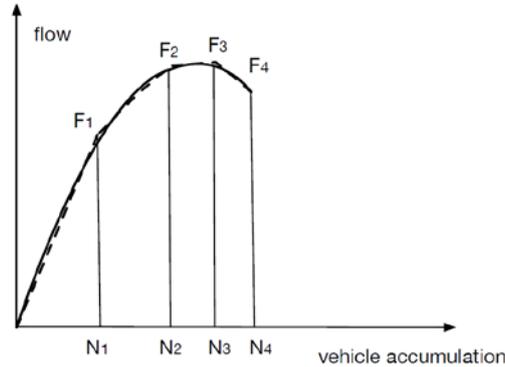


Figure 4. Piecewise linear approximation of the MFD relation

#### 3.1. Approximating nonlinear MFD

In general,  $G(n)$  is a nonlinear function and imposes hard constraints for the DSO problem. Previous studies [Merchant and Nemhauser, 1978, Nie, 2011] indicate that this function could be approximated by a series of piecewise linear functions and coined it as ordered solution property (OSP). For the example in Figure 4, we have an approximation for  $G(n)$ <sup>2</sup>

$$\tilde{G}(n) = \begin{cases} \frac{F_1}{N_1} n & 0 \leq n \leq N_1 \\ F_1 \frac{n-N_1}{N_2-N_1} + F_2 \frac{N_2-n}{N_2-N_1} & N_1 \leq n \leq N_2 \\ F_2 \frac{n-N_2}{N_3-N_2} + F_3 \frac{N_3-n}{N_3-N_2} & N_2 \leq n \leq N_3 \\ F_3 \frac{n-N_3}{N_4-N_3} + F_4 \frac{N_4-n}{N_4-N_3} & N_3 \leq n \leq N_4 \end{cases} \quad (35)$$

<sup>2</sup> We omit the reservoir  $r$  and time horizon  $t$  in the notations for this section for simplicity.

where  $(N_m, F_m)$   $m = 1, 2, 3 \dots$  are breaking points of the piecewise function. This study proposes a novel method to tackle the OSP problem using SOS2 variables. Let  $p_m$ ,  $m = 1, 2, 3 \dots$  be a continuous variable, then the following constraints:

$$\sum_m p_m = 1, \quad m = 1, 2, \dots \quad (36)$$

$$n = \sum_m p_m N_m, \quad m = 1, 2, \dots \quad (37)$$

$$\tilde{G}(n^t) = \sum_m p_m F_m, \quad m = 1, 2, \dots \quad (38)$$

$$SOS2: [\dots p_m \dots], \quad m = 1, 2, \dots \quad (39)$$

$$p_m \geq 0, \quad m = 1, 2, \dots \quad (40)$$

is equivalent to Equation (20) but could be solved efficiently with branch-and-bound method.

### 3.2. Reformulating LTM constraints

We use *Big-M* method to formulate the IF-THEN constraints of LTM, following Han et al. [2016]. Two pairs of binary variables would be introduced to indicate the traffic condition at the entrance and exit of a link.

Finally, we could get an MILP formulation with SOS variables for this problem.

## 4. PROSPECTIVE RESULTS AND DISCUSSIONS

This study develops a novel model of dynamic system optimal traffic assignment for hybrid network of reservoirs and connecting links:

- The traffic movement in each reservoir is described by a MFD-based dynamical model, while the traffic dynamics on connecting links is described by link transmission model.
- The DSO model involves hard constraints of nonlinear equation and IF-THEN constraints, SOS2 variables and Big-M method have been applied to convert the hard problem to mixed integer linear programming problem. It could be tackled by most commercial or open source solvers for MILP.
- This formulation doesn't rely on the cells, and hence the number of variables have been reduced.

We have finished the modeling at present. A case study is in progress to validate the proposed DSO framework.

## REFERENCES

- Beale, E. and Forrest, J.J. (1976) Global optimization using special ordered sets. *Mathematical Programming*, 10(1), pp. 52–69.
- Daganzo, C.F. (1995) The cell transmission model, Part II: Network traffic. *Transportation Research Part B: Methodological*, 29(2), pp. 79–93.
- Daganzo, C.F. (2007) Urban gridlock: Macroscopic modeling and mitigation approaches. *Transportation Research Part B: Methodological*, 41(1), pp. 49–62.
- Geroliminis, N. and Daganzo, C.F. (2008) Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings. *Transportation Research Part B: Methodological*, 42(9), pp. 759–770.
- Geroliminis, N., Haddad, J. and Ramezani, M. (2013) Optimal perimeter control for two urban regions with macroscopic fundamental diagrams: A model predictive approach. *IEEE Transactions on Intelligent Transportation Systems*, 14(1), pp. 348–359.

- Han, K., Liu, H., Gayah, V.V., Friesz, T.L. and Yao, T. (2016) A robust optimization approach for dynamic traffic signal control with emission considerations. *Transportation Research Part C: Emerging Technologies*, 70, pp. 3–26.
- Jin, W.-L. (2015) Continuous formulations and analytical properties of the link transmission model. *Transportation Research Part B: Methodological*, 74, pp. 88–103.
- Merchant, D.K. and Nemhauser, G.L. (1978) A model and an algorithm for the dynamic traffic assignment problems. *Transportation Science*, 12(3), pp. 183–199.
- Nie, Y.M. (2011) A cell-based Merchant–Nemhauser model for the system optimum dynamic traffic assignment problem. *Transportation Research Part B: Methodological*, 45(2), pp. 329–342.
- Peeta, S. and Mahmassani, H.S. (1995) System optimal and user equilibrium time-dependent traffic assignment in congested networks. *Annals of Operations Research*, 60(1), pp. 81–113.
- Saeedmanesh, M. and Geroliminis, N. (2016) Clustering of heterogeneous networks with directional flows based on “Snake” similarities. *Transportation Research Part B: Methodological*, 91, pp. 250–269.
- Yildirimoglu, M., Ramezani, M. and Geroliminis, N. (2015) Equilibrium analysis and route guidance in large-scale networks with MFD dynamics. *Transportation Research*, 9, pp. 185–204.
- Yperman, I. (2007) The link transmission model for dynamic network loading. PhD Thesis, Katholieke Universiteit Leuven, Netherlands.
- Ziliaskopoulos, A.K. (2000) A linear programming model for the single destination system optimum dynamic traffic assignment problem. *Transportation Science*, 34(1), pp. 37–49.

# ROUTE CHOICE EQUILIBRATION ALGORITHM FOR STOCHASTIC DYNAMIC TRAFFIC ASSIGNMENT WITH FULL ROUTE SET

J. VERSTRAETE<sup>a\*</sup>, W. HIMPE<sup>b</sup> and C. M. J. TAMPÈRE<sup>c</sup>  
*L-Mob (Leuven Mobility Research Centre), KU Leuven, Belgium*

*Email : [Jeroen.Verstraete@kuleuven.be](mailto:Jeroen.Verstraete@kuleuven.be), [Willem.Himpe@kuleuven.be](mailto:Willem.Himpe@kuleuven.be), [Chris.Tampere@kuleuven.be](mailto:Chris.Tampere@kuleuven.be)*

*\*SB PhD fellow at FWO*

This paper develops a stochastic logit route equilibration algorithm for dynamic traffic assignment based on an implicit full and stable route set. This algorithm generalizes a recursive formulation and solution of the turning rates at nodes that was previously proposed for static assignment (Bell, M.G.H., 1995. Alternatives to Dial's logit assignment algorithm. *Transp. Res. Part B* 29, 287–295). Since a full route set comes at the cost of possibly including illogical routes (e.g. cycles), we propose and test conceptually some pragmatic ways of minimizing path choice probabilities for such illogical paths. We show on a mid-sized network how our algorithm indeed accelerates convergence to stochastic equilibrium.

**Keywords:** Route Choice, Stochastic User Equilibrium, Link Transmission Model, Dynamic Traffic Assignment, Recursive Logit

## 1. ROUTE SETS

In every traffic assignment model, a route set needs to be determined. A route set is a subset of all possible routes between an origin destination pair, containing only routes that are plausible. Theoretically in stochastic route choice models, the route set should consider every possible route (albeit that many eventually may have extremely small probability of being used).

Let us consider a stochastic user equilibrium (including the extreme case of approaching zero variance for a deterministic solution). Enumerating all routes in a network is in general infeasible in any real-sized network. However, not including a relevant route in the subset has a negative impact on the quality of the assignment. Most existing algorithms work with an explicit route set, where for every origin-destination pair the plausible subset paths is explicitly enumerated. Implicit approaches on the contrary, avoid explicit enumeration by only defining turn fractions at nodes. These can include all possible all turns (as will be done in this research) or only a subset of possible turns (like Dial's algorithm; (Dial and Voorhees, 1971)). In this latter case, the topological order of nodes determines which turns are plausible, herewith implicitly defining a route set: only turns that connect a lower to a higher ranked node in the topological order are considered. As a consequence, with a change of congestion levels, the topological order and hence also the route set may change.

In addition to the behavioral plausibility and theoretical completeness of the route set, definition of the route set also has important consequences on convergence of equilibration algorithms. More precisely – and as illustrated in the full paper – stability of the route set over iterations is highly determinant for convergence speed. For this reason, in literature two concepts exist for dealing with a route set during iterations of an equilibration algorithm. First, a fixed route set can be predetermined at the beginning of the algorithm, which gives advantages for computation time and convergence speed, but increases the chance of missing a relevant route. Under a fixed route set, a converged equilibrium is the solution of a convex combination of route fractions, a problem easily solved by generalized techniques. To decrease the chance of missing a relevant route – especially in highly congested networks – an alternative approach is to have a flexible route set that is updated over iterations (similar to column generation techniques for combinatorial problems). The problem of finding a consistent route set over iterations is on the other hand more difficult and could potentially deteriorate convergence to an equilibrium solution.

In literature, many efforts start from the concern for a behaviorally and theoretically plausible path set (e.g. having refined path cost correlation structures; (Smits et al., 2014)), often without explicit consideration of its impact on equilibration. This paper explores the alternative starting point, motivated by our observation how an (implicitly defined) full (and hence by definition fixed/stable) route set significantly accelerates equilibration algorithms in static networks. This paper develops a stochastic

route equilibration algorithm for dynamic traffic assignment based on an implicit full and stable route set. We show on a mid-sized network how this indeed smoothens convergence to stochastic equilibrium. However this comes at the cost of possibly including illogical routes (e.g. cycles) and requires a recursive formulation and solution of the turning rates at nodes (Bell, 1995; Fosgerau et al., 2013). We propose and test conceptually some pragmatic ways of minimizing path choice probabilities for such illogical paths. Moreover, we show how the recursive solution in DTA can be solved, and offers potential for developing warm-started algorithms with arbitrarily large path-choice time discretization intervals.

## 2. FULL ROUTE SET

The full route set is by definition fixed over iterations and guarantees that in equilibrium, all relevant routes are part of the solution. As in real-sized networks, a combinatorial large or even infinite number of routes exists, a full route set has to be inevitably implicit. (Bell, 1995) was the first to acknowledge and exploit this fact for formulating an full-route-set based alternative to Dial's famous implicit algorithm for stochastic static assignment. Bell's work went relatively unnoticed, until (Fosgerau et al., 2013) recently took inspiration from it to formulate the so-called recursive logit model. They used the implicit full route set formulation to guarantee consistency between choice set generation and parameter estimation for a logit route choice model.

The idea of Bell is simple and elegant. Let  $W$  be a matrix of weights constructed as follows:

$$w_{n,m} = \exp(-\alpha * Cost_{n,m}) \quad (1)$$

Then  $W$  expresses the weights between each pair of two links  $m$  and  $n$  directly (with a weight of 0 if the path doesn't exist). Bell shows then that  $W^2$  expresses the combined weights of all the routes between each pair of links consisting of exactly 2 links. The combined weights of all routes between each pair of links consisting of any number of links is  $W + W^2 + W^3 + \dots = (I - W)^{-1} - I$ . This can only be calculated if  $(I - W)$  can be inverted, thus not every network (with its specific parameters) can have such combined matrix of weights. The probability that a link  $k$  (connecting node  $r$  to node  $s$ ) is used for an OD-pair  $i$  to  $j$  can be calculated by.

$$P_{ijk} = w_{ir} * \exp(-\alpha * Cost_k * \frac{w_{sj}}{w_{ij}}) \quad (2)$$

The path choice is thus reduced to a sequential link choice model. Fosgerau more explicitly states the conditional probability for a traveler  $n$  going to a destination  $d$  will use link  $k$  given the traveler is on link  $a$ . The Expected Maximum Perceived Utility on link  $k$  is:

$$V_k^d = E[ \max_{a \in A_k} (v_{ka} + \mu \varepsilon + V_a^d) ] \quad (3)$$

With  $v_{ka}$  the utility for going from link  $k$  to  $a$ ,  $\varepsilon$  a random utility component,  $\mu$  a scale parameter,  $V_a^d$  the expected downstream utility at link  $a$  and  $A_k$  the sets of links that are accessible from link  $k$ . The Expected Maximum Perceived Utility function in a logit model can be analytically solved as a logsum. Equation 3 then becomes:

$$V_k^d = \mu \ln \sum_{a \in A_k} \exp(\frac{1}{\mu} (v_{ka} + V_a^d)) \quad (4)$$

An important difference between the formulation of Fosgerau and Dial lies in the absence of any (topological) order. This means that  $V_k^d$  can depend on itself, be it directly by a U-turn or indirectly through circular dependencies. This leads to a system of equations that under certain (mild) conditions can be solved.

The algorithm developed in this paper for the dynamic traffic assignment, that has the recursive logit formulation as its foundation, consists of two large parts. The first is to determine the maximum perceived utility per time step. The second one is to determine the turning fractions (with these maximum perceived utilities (MPU)). Recursive Logit works with destination based flows, therefore these two parts are done in sequence for every destination. Other iteration schemes may be possible.

To determine the (transformed) MPU for all time steps, we start with the last time step, then other time steps are calculated in an upwind order. For the last time step, we assume that utilities over paths remain constant over time; hence conditions are stationary and a static implicit RL assignment applies, quantifying MPU's of the last time step following eqs. (1)-(4). For other time steps in upwind order, the dynamic counterpart of eq(4) can be applied, as it defines MPU at a time  $t$  to be depending only on  $V_a^d$  values of later time steps  $t^* = t + \text{travel time between } k \text{ and } a$  ( $t \leq t^*$ ). When at a given point we need an MPU at time  $t^*$ , which is in-between two known time steps,  $t_1 < t^* < t_2$ , a linear interpolation is made. To reduce complexity, the implementation in this paper assumes small time steps such that the MPU at time  $t$  only depends on later time steps ( $t < t_1$ ) and eq(4) becomes an explicit set of equations. Note however that conceptually it is easy to extend this algorithm to larger time steps, for which interpolations may be needed between MPU at later times and the ones currently being calculated, which would require implicit (iterative) solutions of the system of eqs(4). This however has the advantage of allowing arbitrarily large time increments, and of initiating the algorithm with approximate solutions (i.e. warm starting, for instance when repeated, near-identical computations of the same network are done).

Once all maximum perceived utilities are calculated, the last thing remaining is to calculate the turning fractions (or probabilities) out of node  $k$  towards all nodes  $a$  of the forward star  $A(k)$ , following standard multinomial logit probabilities:

$$P^d(a|k) = \frac{\exp\left(\frac{1}{\mu}(v_{ka} + V_a^d)\right)}{\sum_{a' \in A_k} \left(\exp\left(\frac{1}{\mu}(v_{ka'} + V_{a'}^d)\right)\right)} \quad (5)$$

The output of this algorithm are (destination based) turning fractions over time at each node. Any dynamic network loading (DNL) algorithm can apply these turning fractions to compute link flows and new travel times. In our experiments, we used as a DNL the iterative Link Transmission Model (Himpe et al., 2016). With new travel times, the MPU's can be recomputed in the next equilibration iteration until convergence.

Because the route set is stable over iterations, it turns out that the equilibration algorithm can update the turning rates at every iteration using proportional steps, whereas traditionally a decreasing step size (method of successive averages, MSA) is required. As a result, the stochastic equilibrium algorithm converges smoothly (no zig-zag) and much faster than MSA. Figure 1 shows the convergence of stochastic DTA route equilibration in the network of Rotterdam. The network consists of 560 links and 331 node, where 44 nodes serves as origins and destinations. To show the potential of the algorithm, different step sizes are plotted: decreasing MSA step, a proportional step size of 0.5, and a full step size.

One drawback of the full route set however, is that by using all possible routes, some illogical routes are considered. Examples are routes with cycles or unlikely shortcuts like off-on-ramp combinations. A pragmatic way of preventing this is to add penalties to the utilities  $v_{ka}$  in eqs (4) and (5), representing the cost of traveling from node  $k$  to node  $a$ , hence the turning cost on  $k$  towards  $a$  + cost of the link between these nodes. Such penalties can reduce the probability of illogical routes to negligible fractions. We illustrate this approach by using a penalty for a U-turn (which usually should not occur in plausible routes, as consecutive U-turns would form unrealistic cycles). Unlikely shortcuts could be prevented by adding penalties for transitions between road hierarchies. In ongoing future work, we will empirically develop a set of (turning) penalties that best predicts observed route choices in floating car data.

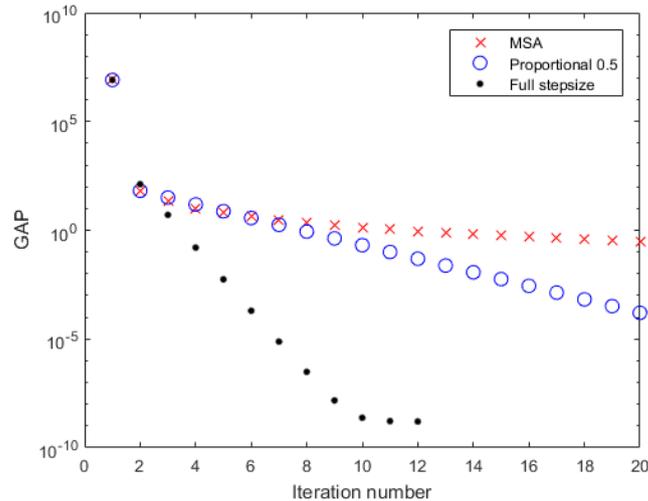


Figure 11: Convergence of recursive logit DTA with different step sizes

### 3. CONCLUSION

Explicit route sets in stochastic dynamic route choice equilibrium, have some down sides, they do not contain every possible route and with a flexible route set the convergence problem is harder to solve. A fixed route set makes the algorithm more stable and efficient. An implicit full route set guarantees that no possible route is missed, because it considers all possible routes (can be an infinity number). In this research this is done by using recursive logit, which calculates the probability of each turn given the traveler's destination. The utility of a turn depends on the MPU of all possible paths downstream of the chosen link towards the destination plus the utility of making the turn and travelling that link. Apart from obvious cost components like the turn delay and link travel time, additional penalties can be defined to minimize the probability of choosing implausible routes. Examples of such penalties are a toll or a Boolean indicating an increase (or decrease) in link hierarchy. The latter could for instance be used to avoid unrealistic routes that leave a higher road category over a very short distance (e.g. off-on-ramp combinations; or unrealistic rat-running through residential streets).

This paper has shown that the static procedure described by (Fosgerau et al., 2013) can be generalized to a dynamic traffic assignment by interpolating the maximum perceived utilities in an upwind order. After the maximum perceived utilities are determined, the probabilities of each turn can be calculated. These are then the input of an DNL. By using a fixed full route set, the algorithm appears to be more stable. This results in fewer iterations needed for convergence, as now a fixed proportional update step size can be used instead of one that decreases over iterations like MSA.

### REFERENCES

- Bell, M.G.H., 1995. Alternatives to Dial's logit assignment algorithm. *Transp. Res. Part B* 29, 287–295.
- Dial, R.B., Voorhees, A.M., 1971. PROBABILISTIC MULTIPATH TRAFFIC ASSIGNMENT WHICH OBVIATES PATH ENUMERATION The problem. *Transpn Res* 5, 83–111.
- Fosgerau, M., Frejinger, E., Karlstrom, A., 2013. A link based network route choice model with unrestricted choice set. *Transp. Res. Part B Methodol.* 56, 70–80.
- Himpe, W., Corthout, R., Tampère, M.J.C., 2016. An efficient iterative link transmission model. *Transp. Res. Part B Methodol.*, Within-day Dynamics in Transportation Networks 92, Part B, 170–190.
- Smits, E.-S., Bliemer, M., Pel, A., Arem, B., 2014. On Route Choice Models with Closed-Form Probability Expressions, in: *Transportation Research Board Annual Meeting*. Presented at the TRB Annual Meeting, Washington DC.

**Session C2**  
**Dynamic Traffic Control II**

**Session C3**  
**Day-to-Day Dynamics II**

# ONE IS NOT LIKE THE OTHER: DETERMINISTIC VS STOCHASTIC DAY-TO-DAY DTA MODELS: PART 1

K. PARRY<sup>a</sup>, A. MAHMOODJANLOU<sup>a</sup> and M. L. HAZELTON<sup>a</sup>

<sup>a</sup> *Massey University, New Zealand*

*Email: [k.parry@massey.ac.nz](mailto:k.parry@massey.ac.nz)*

Day-to-day dynamic traffic assignment models are a powerful tool for predicting how a system evolves after structural changes to a traffic system. One particular application is to gauge the effectiveness of an intervention aimed at shifting the system towards a desired state; for example, introducing free public transport to encourage people to switch from using private vehicles for travel.

The deterministic version of these models, where the modelled properties remain the same if the starting conditions are unchanged, is well-studied. Their stochastic counterpart, where these properties can vary widely even under equal initial conditions, is much less well understood.

We demonstrate via illustrative examples some basic differences between the deterministic and stochastic day-to-day models, for example how reaching a state of equilibrium expresses itself. This work shows that forecasts of patterns of travel flow over an extended period of time can differ significantly between a stochastic model and its (seemingly) analogous deterministic counterpart. This has important implications for network control.

**Keywords:** Traffic Assignment, Deterministic, Stochastic, Equilibrium, Long-Term

## 1. BACKGROUND

Evaluation of proposed urban planning strategies involve various tools, one of the most important being day-to-day dynamic traffic assignment models. These models translate travel demand through the network, observed over a sequence of time points, into flows and travel times on the individual routes and links in the system. One particular application is to gauge the effectiveness of an intervention aimed at shifting the system towards a desired state; for example, introducing free public transport to encourage people to stop using private vehicles for travel.

Day-to-day dynamic traffic assignment models come in two varieties. On the one hand, we have deterministic models which focus on the identifiable source of variation, and where the modelled properties remain the same if the starting conditions are unchanged. On the other hand, we have the stochastic version of these models where these properties can vary widely even under equal initial conditions because of unpredictable sources of variability. While stochastic day-to-day assignment models have more flexibility to reflect the variability we observe in travel pattern data more realistically, deterministic models generally have an advantage in terms of computational efficiency and the tractability of their mathematical properties.

Consequently, deterministic models are well studied (Cantarella et al., 2015; Watling, 1999; He and Liu, 2010; Bie and Lo, 2010; Friesz et al., 1994; M.J.Smith, 1984). Stochastic models have also received considerable attention (K. Parry et al., 2016; Watling and Cantarella, 2013; Hazelton and Watling, 2004; T. Akamatsu, 1996; Davis and Nihan, 1993; Cantarella and Cascetta, 1995). However, there remains much that is not well understood about their properties. For example, we know that the mean processes of certain classes of Markov traffic assignment models describe natural deterministic dynamic processes in systems with unique (deterministic) fixed points (Cantarella and Cascetta, 1995), but we still understand relatively little about relationships between such deterministic and stochastic assignment models, in particular in the presence of multiple (deterministic) equilibria.

## 2. EQUILIBRIUM ANALYSIS

In this talk we present, based on the two-route network with multiple equilibria discussed in Smith et al. (2014), how reaching equilibrium expresses itself differently between the two types of dynamic traffic

assignment models. We also specify both the deterministic and stochastic model following Smith et al. (2014), and assume that the properties of the system on day  $t + 1$  are fully regulated by the costs generated from flows on day  $t$ .

## 2.1. Deterministic model

In a deterministic model, we suggest that drivers swap from one route to another at a given rate  $c_r(\mathbf{x}^t) - c_s(\mathbf{x}^t)$  on any given day  $t$ , where  $\mathbf{x}^t$  are the route flows on day  $t$ ,  $r$  and  $s$  index different routes. If switching routes is cheaper then the difference will be positive. We then define  $\Delta_{r,s}$  as the swap vector from route  $r$  to route  $s$  with  $-1$  in the  $r$ th place,  $+1$  in the  $s$ th place and  $0$  elsewhere.

The overall rate of change in the route flows is thus:

$$d(\mathbf{x}^t) = \sum k \cdot [c_r(\mathbf{x}^t) - c_s(\mathbf{x}^t)]_+ \cdot x_r^t \cdot \Delta_{r,s} \quad (41)$$

where  $k \in (0, 1)$  and  $x_+ = \max(0, x)$ . The parameter  $k$  determines the proportion of users that are prepared to switch on each day. A value closer to  $1$  implies a higher level of swapping between routes from day to day.

Given route volumes  $\mathbf{x}^t$  on the previous day, the route flows on day  $t + 1$  are

$$\mathbf{x}^{t+1} = \mathbf{x}^t + d(\mathbf{x}^t). \quad (42)$$

A sufficiently small  $k$  value negates the issue of the elements of  $\mathbf{x}$  potentially taking on negative values.

Equilibria are fixed points in the state space, that is, a traffic system will reach equilibrium if and only if

$$[c_r(\mathbf{x}^t) - c_s(\mathbf{x}^t)]_+ \cdot x_r^t = 0 \quad (43)$$

for all pairs of routes  $r$  and  $s$  serving the same origin-destination pair. After equilibrium is reached, the flow pattern no longer changes from day to day. This occurs when either the flow on route  $r$  is zero, or there is no cheaper alternative route  $s$  available, correspondingly to Wardrop's user equilibrium. Our plots mirror the findings given by Cantarella et al. (2015), who demonstrate that the existence of multiple modes of transport will lead a deterministic model to converge towards multiple fixed system states, called equilibrium points, as a result of interactions between the modes.

## 2.2. Stochastic model

When taking a stochastic modelling approach, the sequence of route flows  $\{\mathbf{x}^t : t = 1, 2, \dots\}$  is modelled as Markov chain (E.Cascetta, 1989). We use a logit model to define the route choice probabilities:

$$q_{t+1} = 1 / [1 + \sum_{s \sim r} \exp(\beta [c_r(\mathbf{x}^t) - c_s(\mathbf{x}^t)])] \quad (44)$$

where  $s \sim r$  indicates that routes  $r$  and  $s$  serve the same origin-destination pair and  $\beta$  is a sensitivity measure of travellers to changes in the system. The parameter  $\beta$  can take on any positive value, with larger values representing more reactive users.

As we are considering a network with only two routes, we can model route flows for a sequence of time intervals indexed by  $t = 1, 2, \dots$  using a Binomial distribution:

$$\mathbf{x}^{t+1} \sim \text{Bin}(N, \mathbf{q}^{t+1}). \quad (45)$$

We will see that in the stochastic case, an equilibrium corresponds to a fixed distribution in the space of probability measures on the route flow vector. We also illustrate that the flow pattern does not remain constant in a stochastic equilibrium: the flow vectors  $\mathbf{x}^t$  and  $\mathbf{x}^{t+1}$  will almost certainly differ. It is the pattern of variation, as characterized by the underlying probability distribution, that is invariant.

This talk shows that forecasts of patterns of travel flow over an extended period of time can differ significantly between a stochastic model and its (seemingly) analogous deterministic counterpart. This has important implications for network control.

## REFERENCES

- Akamatsu, T. (1996) Cyclic flows, Markov process and stochastic traffic assignment. *Transportation Research Part B: Methodological*, 30, pp. 369-386.
- Bie, J. and Lo, H.K. (2010) Stability and attraction domains of traffic equilibria in a day-to-day dynamical system formulation. *Transportation Research Part B: Methodological*, 44, pp. 90-107.
- Cantarella, G.E. and Cascetta, E. (1995) Dynamic processes and equilibrium in transportation networks: Towards a unifying theory. *Transportation Science*, 29, pp. 305-329.
- Cantarella, G.E., Belona, P., and Watling, D.P. (2015) Day-to-day dynamics and equilibrium stability in a two-mode transport system with responsive bus operator strategies. *Networks and Spatial Economics*, 15, pp. 485-506.
- Cascetta, E. (1989) A stochastic process approach to the analysis of temporal dynamics in transportation networks. *Transportation Research, Part B: Methodological*, 23, pp. 1-17.
- Davis, G.A. and Nihan, N.L. (1993) Large population approximations of a general stochastic traffic assignment model. *Operations Research*, 41, pp. 169-178.
- Friesz, T.L., Bernstein, D., Mehta, N.J., Tobin, R.L. and Ganjalizadeh, S. (1994) Day-to-day dynamic disequilibria and idealized traveller information systems. *Operations Research*, 42, pp. 102-113.
- Hazelton, M.L. and Watling, D.P. (2004) Computation of equilibrium distributions of Markov traffic assignment models. *Transportation Science*, 38, pp.331-342.
- He, X. and Liu, H.X. (2010) A link-based day-to-day traffic assignment model. *Transportation Research, Part B: Methodological*, 44, pp. 597-608.
- Parry, K., Watling, D.P. and Hazelton, M.L. (2016) A new class of doubly stochastic day-to-day dynamic traffic assignment models. *EURO Journal on Transportation and Logistics*, 5, pp. 5-23.
- Smith, M.J. (1984) The stability of a dynamic model of traffic assignment- an application of a method of Lyapunov. *Transportation Science*, 18, pp. 245-252.
- Smith, M.J., Hazelton, M.L., Lo, H.K., Cantarella, G.E. and Watling, D.P. (2014) The long-term behaviour of day-to-day traffic assignment models. *Transportmetrica A: Transport Science*, 10, pp. 647-660
- Watling, D.P. (1999) Stability of the stochastic assignment problem: a dynamical systems approach. *Transportation Research Part B: Methodological*, 33, pp. 281-312.
- Watling, D.P. and Cantarella, G.E. (2013) Modelling sources of variation in transportation systems: theoretical foundations of day-to-day dynamic models. *Transportmetrica B: Transport Dynamics*, 1, pp. 3-32.

# ONE IS NOT LIKE THE OTHER: DETERMINISTIC VS STOCHASTIC DAY-TO-DAY DTA MODELS: PART 2

A. MAHMOODJANLOU<sup>a</sup>, M. L. HAZELTON<sup>a</sup> and K. PARRY<sup>a</sup>

<sup>a</sup> *Massey University, New Zealand*

*Email: [a.mahmoodjanlou@massey.ac.nz](mailto:a.mahmoodjanlou@massey.ac.nz)*

In deterministic models we know that moving the system into the basin of attraction for a desired state guarantees convergence to that equilibrium point after a fixed number of steps/timepoints. However, the situation is more complex when considering stochastic day-to-day dynamic models.

In this talk we examine the length of time required for a stochastic model to reach some pre-specified state (travel pattern). The expected value of this time (which is itself a random variable) is called the mean first passage time (MFPT). Using logit route choice with a Markov process model for traffic assignment, we examine how MFPT depends on system properties like sensitivity to cost differences. We present a theorem which shows that it is possible to find stochastic day-to-day models for which attainment of a desirable target pattern of traffic flow occurs with probability one irrespective of initial state of the system, but for which the expected time to achieve this state (as expressed by the MFPT) is infinite.

The talk finishes with a brief proof of the theorem and some interesting numerical results.

**Keywords:** Traffic Assignment, Deterministic, Stochastic, Equilibria, Long-Term

## 1. INTRODUCTION

Traffic dynamic assignment models are critical tools for system management and urban planning. The predicted traffic flows from these models allow us to make forecasts on travel behaviour and hence decisions about long-term network improvements or adverse events like the failure of a bridge, or to shorter-term network control measures such as the imposition of tolls or refinement of traffic signalling schemes (He and Liu, 2010).

A variety of types of assignment model are now available. The first to be developed were static equilibrium models like Wardrop or Deterministic User Equilibrium (DUE) (Wardrop, 1952) and the Stochastic User Equilibrium (SUE) (Daganzo, 1983). Such models do not capture variation in patterns of traffic flow, which may occur as a result of identifiable changes in conditions and/or may because haphazard and unpredictable variation in traffic flows (Watling and Hazelton, 2003; Watling and Cantarella, 2013). In response, researchers have paid increasing attention over the past 40 years on models that describe the dynamic behaviour of traffic systems, with a significant body of work focusing on day-to-day variation.

The natural variation in flows from day-to-day can be captured in two different ways. In deterministic traffic assignment models we focus on the identifiable sources of variability, while in stochastic models the predicted flows account for unknown uncertainty as well. A critical issue in the context of traffic planning and network control concerns the convergence of such models to fixed equilibrium points (Cantarella et al., 2015; Hazelton, 2002; Horowitz, 1984). In the first part of this presentation we demonstrate that in systems with multiple equilibria the predicted performance of the network may vary substantially between the deterministic and stochastic versions of them.

We now consider further the natural goal of seeking to influence travellers so as to direct the overall flow toward an optimal fixed point. This issue was studied by Bie and Lo (2010), who examined how the state space of such a system may be partitioned into multiple basins of attraction. For deterministic models, once the system is in such a basin, it will converge towards the equilibrium point within. From this follows that in principle, deterministic models lead us to believe that an intervention that shifts the system into the basin of attraction for our desired equilibrium will guarantee success. While stochastic

models also allow us to assume a successful shift to a desired state, the time needed to do so may become impractically large due to the possibility of the system switching to an entirely different part of the state space (Watling, 1996).

In this talk we take a closer look at stochastic models and define a random variable  $T$ , that measures the number of steps it takes to get from a starting state  $i$  in the system to a desired state  $j$  for the first time:

$$T_{ij} = \min \{t, X_t = j | X_0 = i\}$$

We are interested in the expected number of time-steps for reaching state  $j$  for the first time given that the chain was initially in state  $i$ , which we refer to as mean first passage time (MFPT). The concept of MFPT is important to understanding the convergence of a stochastic process to a deterministic process when we have a large number of travelers as well as increasing values of  $\beta$ . To this end we present the MFPT for a number of varying demand levels and increasing values of  $\beta$ .

We then develop a theorem that characterizes the conditions under which the convergence of long-run dynamic behaviour in stochastic models occurs.

Theorem: Let  $M > 0$  be given. Then exists a day-to-day stochastic model for which both of the following hold:

1. The flow vector will reach some pre-specified route flow pattern  $x_0$  at some time point with probability 1;
2. The expected time (i.e. mean number of days) for the system to reach  $x_0$  is greater than  $M$ .

In terms of network, suppose that the route flow pattern  $x_0$  is some desirable state that we wish to reach. Then the theorem indicates that we can ensure that  $x_0$  is eventually attained by the system, but that it may take an arbitrarily large number of days for this to happen. The proof works by constructing a system based on the two-route network with multiple equilibria discussed in Smith et al. (2014) where the two routes correspond to two modes, bus and car. Our talk will finish with more details on the proof of this theorem.

Our overall goal is raise awareness of circumstances in which the choice between apparently similar models can have a profound effect on forecasts of future system behaviour, and a corresponding impact on decision making when faced with alternative schemes for network modification and control.

## REFERENCES

- Bie, J. and Lo, H.K. (2010) Stability and attraction domains of traffic equilibria in a day-to-day dynamical system formulation. *Transportation Research Part B: Methodological*, 44, pp. 90-107.
- Cantarella, G.E., Belona, P. and Watling, D.P. (2015) Day-to-day dynamics and equilibrium stability in a twomode transport system with responsive bus operator strategies. *Networks and Spatial Economics*, 15, pp. 485-506.
- Daganzo, C.F. (1983) Stochastic network equilibrium with multiple vehicle types and asymmetric, indefinite link cost jacobians. *Transportation Science*, 17, pp. 282-300.
- Hazelton, M.L. (2002) Day-to-day variation in Markovian traffic assignment models. *Transportation Research Part B: Methodological*, 36, pp. 637-648.
- He, X. and Liu, H.X. (2010) A link-based day-to-day traffic assignment model. *Transportation Research Part B: Methodological*, 44, pp. 597-608.

- Horowitz, J.L. (1984) The stability of stochastic equilibrium in a two-link transportation network. *Transportation Research Part B: Methodological*, 18, pp. 13-28.
- Smith, M.J., Hazelton, M.L., Lo, H.K., Cantarella, G.E. and Watling, D.P. (2014) The long-term behaviour of day-to-day traffic assignment models. *Transportmetrica A: Transport Science*, 10(7), pp. 647-660.
- Wardrop, J.G. (1952) Some theoretical aspects of road traffic research. *Proceedings of the Institute of Civil Engineering Part II*, 1, pp. 325-378.
- Watling, D.P. (1996) Asymmetric problems and stochastic process models of traffic assignment. *Transportation Research Part B: Methodological*, 30, 339-357.
- Watling, D.P. and Cantarella, G.E. (2013) Modelling sources of variation in transportation systems: theoretical foundations of day-to-day dynamic models. *Transportmetrica B: Transport Dynamics*, 1, pp. 3-32.
- Watling, D.P. and Hazelton, M.L. (2003) The dynamics and equilibria of day-to-day assignment models. *Networks and Spatial Economics*, 3, pp. 349-370.

# DAY-TO-DAY NETWORK DYNAMICS OVER DIFFERENT SCALES: FROM MICRO TO MACRO

W. LI<sup>a</sup> and N. JIA<sup>b</sup>

<sup>a</sup> *College of Management and Economics,  
Tianjin University, China  
Email: [liweimeng@tju.edu.cn](mailto:liweimeng@tju.edu.cn)*

<sup>b</sup> *College of Management and Economics,  
Tianjin University, China  
Email: [jia\\_ning@tju.edu.cn](mailto:jia_ning@tju.edu.cn)*

Day-to-day traffic (DTD) dynamical systems have attracted much attention recently as a powerful tool describing the evolution of network flows over time. Since travelers' choice behavior is the driven force of the variation of network flow, there must be some links between individual-level choice behavior and network-level flow dynamics. This paper proposed a unifying framework which connects agent-based model, stochastic process model and deterministic model together, where the deterministic model is the most aggregative model, the agent-based model is the least aggregative model, and the stochastic process model is in between. The same network dynamics can be modeled over the three scales. We have proved that several well-known day-to-day dynamic models, including Smith's model(Smith,1984), Logit model(Sheffi,1984), Inertia Logit model(Cantarella *et al.*,2016), Jin(2007), Nagurney(1997) and Kumar *et al.*(2015), all of them can be re-defined over three scales. Some theoretical characteristic, e.g. equilibrium and stability, are discussed over three scales. The relationships between different models are confirmed by numerical examples. Some virtual experiments-based data are used to verify the model, and results suggest that real day-to-day dynamics follows the proposed framework.

**Keywords:** Day-To-Day Network Dynamics, Agent-Based Modeling, Stochastic Process

## 1. BACKGROUND

Day-to-day traffic dynamic models can generally be classified into two kinds: macro and micro models. The former class concentrates on the network-level, i.e. network flow dynamics(Smith,1984; Yang *et al.*,2009; Kumar *et al.*,2015), while the latter class is constructed from the perspective of an individual traveler's route choice behavior (Djavadian *et al.*,2017; Tian *et al.*,2010), which is well-known as agent-based model.The macro models can be further divided into deterministic models (He *et al.*,2012; Jing *et al.*, 2010) and stochastic ones (Smith *et al.*,2014; Watling *et al.*,2015; Cantarella *et al.*,2016) according to whether stochastic factors are considered. Since travelers' choice behavior is the driven force of the variation of network flow, there must be some links between individual-level choice behavior and network-level flow dynamics. However, studies on the link between network flow dynamics and individual choice behavior are still elusive. A few existing results can be seen in Jin(2007), Wei *et al.*(2016).To making up the present theory, this paper makes some effects on it.

Contributions of this work include: A unifying framework is proposed connecting agent-based model, stochastic process model and deterministic model together. These models over different scales have their own emphases on describing the phenomenon of traffic flow change, where the deterministic model is the most aggregative model, the agent-based model is the least aggregative model, and the stochastic process model is in between.The same network dynamics can be modeled over the three scales. We have proved that several well-known day-to-day dynamic models, including Smith's model(Smith,1984), Logit model(Sheffi,1984), Inertia Logit model(Cantarella *et al.*,2016), Jin(2007), Nagurney *et al.*(1997) and Kumar *et al.* (2015), all of them can be re-defined over three scales. Some theoretical characteristic, e.g. equilibrium and stability, are discussed over three scales. The relationships between different models are confirmed by numerical examples. Some virtual experiments-based data are used to verify the model, results suggest that real day-to-day dynamics follows the proposed framework.

## 2. THE PAPER CONTENT

### 2.1 The modeling approach of the unifying framework

#### 2.1.1 The agent-based model

Given an  $n$ -path network,  $f_k^t$  is the flow volume on path  $k$  at time  $t$ , i.e. the number of agents who choose path  $k$  at time  $t$ . The flow volume at time  $t$  is notated as  $F^t = (f_1^t, f_2^t, \dots, f_n^t)^T$ . Each agent has a decision vector  $\vec{P}_j^t = (p_{j1}^t, p_{j2}^t, \dots, p_{jn}^t)^T$ , where  $p_{jk}^t$  is the probability of choosing path  $k$  at time  $t$  for agents in the path  $j$ . Agents make decisions according to decision vector. After each choice, agents update their decision vector by the travel costs they have experienced. The probability updating rule is  $\vec{P}^t = \Phi(F^t)$ , meaning that  $\vec{P}^t$  relates to previous flow assignment.

#### 2.1.2 The stochastic process model

According to the model above, an agent on path  $j$  chooses path  $i$  with probability  $p_{ji}$ , and does not choose  $i$  with probability  $1 - p_{ji}$ . And agents on the same path have the same decision vectors. Then the number of agents choosing path  $i$  on path  $j$  follows a bi-nominal distribution  $\text{Binomial}(f_j, p_{ji})$ . The stochastic evolution process of the system is:

$$f_i^{t+1} = \sum_{j=1}^n \Delta h_{ji}^t \quad (1)$$

$$\Delta h_{ji}^t \sim \text{Binomial}(f_j^t, p_{ji}^t) \quad (2)$$

$$\vec{P}^t = \Phi(\vec{F}^t) \quad (3)$$

$\vec{P}^t$  is the decision vector for agents on path  $j$ . That is, given flow distribution  $\vec{F}^t$ , route flow  $\vec{F}^{t+1}$  follows the Bi-nominal Poisson Distribution, in which distribution parameters are determined by and only by  $\vec{F}^t$ . Thus the distribution of  $\vec{F}^{t+1}$  depends on and only on  $\vec{F}^t$ , which indicates that the evolution of the system is a Markov process. For specific probability updating rule, the existence of limited distribution could be determined according the theorem given by Stokoy *et al.* (1989). This model can be viewed as a mesoscopic form, between the agent-based model and the deterministic model.

#### 2.1.3 The deterministic model

The stochastic process model could make connections with deterministic dynamic model by taking expectations (Watling *et al.*, 2015):  $\Delta h_{ji}^t \sim \text{Binomial}(f_j^t, p_{ji}^t) \Rightarrow E(\Delta h_{ji}^t) = f_j^t \cdot p_{ji}^t$ .

By this method, the deterministic dynamic model can be notated as

$$f_i^{t+1} = \sum_{j=1}^n \Delta h_{ji}^t \quad (4)$$

$$\Delta h_{ji}^t \sim f_j^t \cdot p_{ji}^t \quad (5)$$

$$\vec{P}^t = \Phi(\vec{F}^t) \quad (6)$$

If  $\Phi(\vec{F}^t)$  is specific, the macroscopic path flow swapping rules could be deduced. The deterministic model is the most aggregated model. Generally it can be viewed as a macro model of traffic flow.

### 2.1.4 The equilibriums over different scales

## 2.2 The explanation of existing models under the unifying framework

In this section, this paper proves refined forms of some popular models over these three scopes and analyses their properties respectively. The form of Smith's model is the simplest among them Let's take this as an example to explain the analysis process. That of other models are presented in full paper in the future.

### 2.2.1 Smith's model

Let's suppose probability updating rule as

$$p_{ji}^t = \begin{cases} 1 - \sum_{j=1, j \neq i}^n \eta(C_j^t - C_i^t)_+ & \text{if } j = i \\ 0 < \eta(C_j^t - C_i^t)_+ < 1 & \text{if } j \neq i \end{cases} \quad (7)$$

where  $C_j^t$  is the travel cost of currently used path. For any real number  $X_+$ ,  $X_+ = \max\{0, X\}$ . We suppose  $C_j^t = c_j^t(f_j^t)$ .  $c_j^t$  is a monotone increasing function. Agents change their route choice probabilities according to this rule.

At individual level, the dynamic system is equilibrium if and only if

$$F_j^t(C_j^t - C_i^t)_+ = 0 \quad \forall j, i \in \{1, 2, \dots, n\} \quad (8)$$

From the perspective of stochastic process, the traffic flow has a limit distribution. The proof is in Appendix. From a deterministic perspective, the path swapping rule could be obtained by mathematical derivation.

$$\begin{aligned} \Delta f_i^t &= E \left( \sum_{j=1}^n \Delta h_{ji}^t - \sum_{j=1}^n \Delta h_{ij}^t \right) \\ &= \sum_{j=1}^n (E(\Delta h_{ji}^t) - E(\Delta h_{ij}^t)) \\ &= \sum_{j=1}^n (f_j^t p_{ji}^t - f_i^t p_{ij}^t) \end{aligned} \quad (9)$$

$$\begin{aligned} &= \eta \cdot \sum_{j=1}^n (f_j^t [C_j - C_i]_+ - f_i^t [C_i - C_j]_+) \\ f_j^{t+1} - f_j^t &= \eta \cdot \sum_{j=1}^n (f_j^t [C_j - C_i]_+ - f_i^t [C_i - C_j]_+) \end{aligned} \quad (10)$$

The continuous form of path swapping rule is:

$$\dot{f}_j(t) = \eta \cdot \sum_{j=1}^n (f_j^t [C_j - C_i]_+ - f_i^t [C_i - C_j]_+) \quad (11)$$

Let  $\dot{F}(t) = (\dot{f}_1(t), \dot{f}_2(t), \dots, \dot{f}_n(t))$ ,  $F(t) = \Psi(F)$ . According to Smith(1984), the dynamic system is equilibrium if and only if  $\Psi'(F) = 0$ . The dynamic system is stable. The proof has given in Smith(1984).

### 2.2.2 Numerical experiment of Smith's model

The numerical experiment results of this model is given. The flow proportion evolution of the three level models from macro to micro is presented as Figure.1. Figure.2 presents the stochastic process model's frequency distribution of limit flow proportion with different traveller number (30,300,3000).

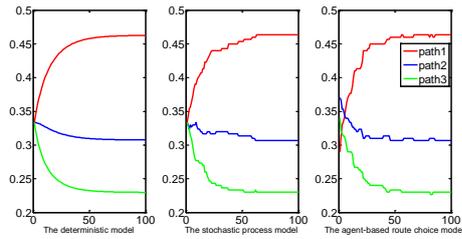


Figure 1. The flow proportion evolution of the three models from macro to micro

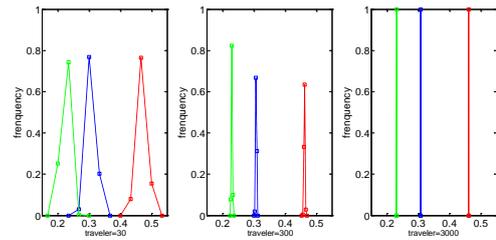


Figure 2. The frequency distribution of limit flow proportion with different traveler number

### 2.3 Empirical evidence

Essentially, real world network flow dynamics are agent-based, where agents are real travelers. According to the proposed framework, it should be described over the three scales. In this part, we collect some experimental data from a laboratory experiment of route choice and compare them with theoretical prediction. Because the subjects in the experiment are not many, the experimental dynamics can be well-reproduced by a stochastic process model. Comparison between experimental and theoretical route flow distributions is presented as Fig.3. It suggests that real day-to-day dynamics follows the proposed framework.

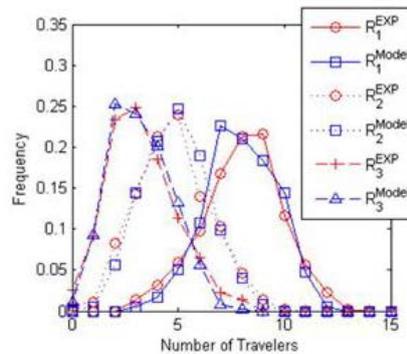


Figure 3. Comparison between experimental and theoretical route flow distributions

### REFERENCES

- Brown, G. W., & Von Neumann, J. (1950). Solutions of games by differential equations. *Contributions to the Theory Games*, 27(4), 73-79.
- Cantarella, G.E., Watling, D.P. (2016) A general stochastic process for day-to-day dynamic traffic assignment: Formulation, asymptotic behaviour, and stability analysis. *Transportation Research Part B Methodological* 92, 3-21.
- Djavadian, S., Chow, J.Y.J. (2017). An agent-based day-to-day adjustment process for modeling 'Mobility as a Service' with a two-sided flexible transport market. *Transportation Research Part B Methodological* 104(C), 36-57.
- He, X., Liu, H.X. (2012). Modeling the day-to-day traffic evolution process after an unexpected network disruption. *Transportation Research Part B Methodological* 46(1), 50-71.
- Jin, W.L. (2007). A dynamical system model of the traffic assignment problem. *Transportation Research Part B Methodological* 41(1), 32-48.
- Jing, B., Hong, K.L. (2010). Stability and attraction domains of traffic equilibria in a day-to-day dynamical system formulation. *Transportation Research Part B Methodological* 44(1), 90-107.
- Kumar, A., Peeta, S. (2015). A day-to-day dynamical model for the evolution of path flows under disequilibrium of traffic networks with fixed demand. *Transportation Research Part B Methodological* 80, 235-256.
- Sheffi, Y. (1984). *Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Models*. Prentice-Hall.

- Smith, M., Hazelton, M.L., Hong, K.L., Cantarella, G.E., Watling, D.P. (2014). The long term behaviour of day-to-day traffic assignment models. *Transportmetrica* 10(7), 647-660.
- Smith, M.J. (1984). The Stability of a Dynamic Model of Traffic Assignment—An Application of a Method of Lyapunov. *Transportation Science* 18(3), 245-252.
- Stokey, N.L., Lucas, R.E. (1989). Recursive Methods in Economic Dynamics. *Canadian Journal of Economics/revue Canadienne D`economique* 23(Jan), 71.
- Tian, L., Huang, H., Liu, T. (2010). Day-To-Day Route Choice Decision Simulation Based on Dynamic Feedback Information. *Journal of Transportation Systems Engineering & Information Technology* 10(4), 79-85.
- Watling, D.P., Cantarella, G.E. (2015). Model Representation & Decision-Making in an Ever-Changing World: The Role of Stochastic Process Models of Transportation Systems. *Networks & Spatial Economics* 15(3), 843-882.
- Wei, F., Jia, N., Ma, S. (2016). Day-to-day traffic dynamics considering social interaction: From individual route choice behavior to a network flow model. *Transportation Research Part B Methodological* 94, 335-354.
- Yang, F., Zhang, D. (2009). Day-to-day stationary link flow pattern. *Transportation Research Part B: Methodological* 43(1), 119-126.

**Session C4**  
**Traffic and Demand Management II & Closing**

# ROBUST MULTI-PERIOD TRADABLE CREDIT SCHEME TO PROMOTE ZERO-EMISSIONS VEHICLES UNDER TRAVEL DEMAND UNCERTAINTY

M. MIRALINAGHI<sup>a</sup> and S. PEETA<sup>b</sup>

<sup>a</sup> *Lyles School of Civil Engineering, Purdue University, USA*

*Email: [smiralin@purdue.edu](mailto:smiralin@purdue.edu)*

<sup>b</sup> *Lyles School of Civil Engineering, Purdue University, USA*

*Email: [Peeta@purdue.edu](mailto:Peeta@purdue.edu)*

This study designs a robust multi-period tradable credit scheme (TCS) to incentivize travelers to shift from internal combustion engine vehicles (ICEVs) to zero-emissions vehicles (ZEVs) over a long-term planning horizon to reduce vehicular emissions. The need for robust design arises because of uncertainty in forecasting travel demand over a planning horizon in the order of several years. The robust multi-period TCS design is formulated as a bi-level model. In the upper level, the CA determines the TCS parameters, i.e. credit allocation and charging schemes, by vehicle mode to minimize the worst-case vehicular emissions rate, i.e. the maximum vehicular emissions rate under the possible travel demand scenarios. Thereby, the incentive to shift to ZEVs is fostered by allocating more credits and charging fewer credits to ZEV travelers compared to ICEV travelers. The upper-level model is a nonlinear mixed integer program with linear and integer constraints. In the lower level, travelers minimize their generalized travel costs under the TCS parameters obtained in the upper level. These parameters are used to determine the mode split between ICEVs and ZEVs using a binomial logit function, and influence route selection based on the difference in credits charged on links for these two modes. The lower-level model is a mathematical program with equilibrium constraints. The bi-level model is solved using a cutting plane method. Numerical experiments illustrate that the proposed TCS design reduces volatility in the realized vehicular emissions rates under different travel demand scenarios compared to a TCS design that does not consider demand uncertainty.

**Keywords:** Multi-Period Tradable Credit Scheme, Zero-Emissions Vehicles, Travel Demand Uncertainty; Robust Design

## 1. INTRODUCTION

Traffic congestion and consequent greenhouse gas (GHG) emissions are major quality-of-life issues in metropolitan areas. The United States has sought to reduce its GHG emissions by 26-28 percent below the 2005 level in 2025 (The White House 2015). The U.S. transportation sector accounted for 26.4% of its GHG emissions in 2016, with cars and trucks accounting for 83.1% of that amount (EPA 2016). Hence, transportation emissions reduction strategies are critical for reducing GHG emissions.

Transportation emissions reduction strategies have been classified into two groups. The first group focuses on mobility management strategies such as congestion pricing and Tradable credit scheme (TCS). Yang and Wang (2011) propose the concept of tradable credit scheme (TCS) which consists of credit allocation and charging schemes. A central authority (CA) determines the number of credits to allocate to travelers (credit allocation scheme) and the number of credits to charge for usage of links (credit charging scheme). To address long-term system-level goals, Miralinaghi and Peeta (2016) propose a multi-period TCS framework that enables the CA to factor long-term fluctuations in travel demand/supply. In this framework, the planning horizon is divided into multiple periods of equal length. The CA determines the credit allocation and charging schemes, referred to as TCS parameters, in each period to achieve system-level goals. Travelers may transfer unused credits to future periods based on several factors such as projected credit prices in future periods, current TCS parameters, and current travel demand and traffic network supply. The second group focuses on promoting alternative fuel vehicles (AFVs) that have fewer emissions compared to internal combustion engine vehicles (ICEVs). Zero-emissions vehicles (ZEVs) are AFVs that do not emit harmful pollutants. Electric and hydrogen vehicles are well-known examples of ZEVs. Despite the potential benefits of ZEVs, their adoption has several barriers such as high retail prices and lack of refueling infrastructure (Krause et al. 2013). Strategies such as fuel tax increase and subsidy programs have been proposed to increase the adoption of ZEVs in practice. However, there is potential political resistance to fuel tax increases (de Palma and

Lindsey 2011). In the U.S., several subsidy programs have been implemented to promote ZEVs. However, they can entail significant monetary burden for the government. For example, CalEPA has invested about \$288 million to fund rebates for ZEVs up to April 2017 (Center for Sustainable Energy 2017). Despite this enormous investment, the adoption rate of ZEVs is still about 4.6% in California (Global Automakers 2017).

This study proposes a hybrid approach combining the two groups of emissions reduction strategies by using the concept of multi-period TCS to encourage the public to purchase and use ZEVs. In this context, the multi-period TCS is used to manage travel demand and travelers' route choices, and promote ZEV adoption by rewarding travelers who generate fewer vehicular emissions. Thereby, a key advantage of the proposed approach is its circumvention of monetary subsidies for the adoption of low-emissions vehicles. Specifically, the study seeks to design a multi-period TCS so that a CA can minimize vehicular emissions over a long-term planning horizon while incentivizing the use of ZEVs for travel purposes. Travel demand is classified into two modes – ZEV and ICEV travelers. For simplicity, other types of alternative-fuel vehicles, e.g., ethanol and natural gas fuel vehicles, are not considered. The TCS parameters of each period are designed to be mode-specific, and consist of: (i) mode-specific origin-destination (O-D) based credit allocation, i.e. credit allocation rate to travelers is based on their mode and O-D pair, and (ii) mode-specific link-based credit charging scheme, i.e. number of credits charged from travelers for using a link is based on the mode. ZEV adoption is rewarded under the multi-period TCS by allocating more credits to ZEV travelers and charging them fewer credits compared to ICEV travelers. The mode split of travelers is captured using a binomial logit function that incorporates the effects of favorable credit allocation and link credit charges for ZEVs. Further, ZEV travelers can sell unused credits to ICEV travelers in a travel credit market and generate monetary gains. This further incentivizes travelers to adopt ZEVs, and contributes further to the higher travel utilities for ZEV travelers compared to ICEV travelers. This ability to leverage the use of a multi-period TCS to promote ZEV usage provides a sustainable behavioral mechanism to achieve the CA's system-level objective of reducing vehicular emissions.

In practice, forecasting travel demand over a long-term planning horizon (in the order of few years) has inherent uncertainty, and the reliability of travel demand forecasts diminishes as the length of the planning horizon increases. This demand uncertainty can be attributed to changes in land use, economic and demographic characteristics. However, it has not been addressed in TCS-design studies, and represents another novel aspect considered in this study. We account for demand uncertainty and illustrate it needs to be explicitly factored as it influences the effectiveness of the multi-period TCS. The robust multi-period TCS design is formulated as a bi-level model. In the upper level, the CA determines the multi-period TCS parameters to minimize vehicular emissions for the travel demand scenario which causes the maximum emissions among all possible demand scenarios, labeled the worst-case vehicular emissions. Also, the generalized travel costs, consisting of travel time and credit consumption costs, are constrained from increasing beyond a predetermined bound in each period after the TCS implementation. The upper-level model is a nonlinear mathematical program with linear and integer constraints. In the lower level, travelers minimize their generalized travel costs through their choice of mode and route under the multi-period TCS parameters obtained in the upper level. These parameters are used to determine the mode split between ICEVs and ZEVs using a binomial logit function, and influence route selection based on the difference in credits charged on links for these two modes. The lower-level model is a mathematical program with equilibrium constraints.

Under the robust TCS design, travel demand for each O-D pair in each period can be its average value or one of the values in its uncertainty set whose probabilities are unknown. The travel demand uncertainty set consists of the possible demand scenarios and includes the average demand scenario. The demand values are often forecasted by the metropolitan planning organizations using econometric methods based on socioeconomic and land use characteristics. Bertsimas and Sim (2003) propose the notion of uncertainty budget which in our study context caps the total number of O-D pairs whose travel demand deviates from their average values in each period. It enables the CA to tradeoff computational burden and accuracy. For example, if the CA allows the demand of several O-D pairs to deviate from their average values, it can add significant computational burden due to the number of possible scenarios.

Hence, the notion of uncertainty budget is used in robust design to enhance computational tractability. Another key advantage of the uncertainty budget in our study context is that it enables capturing the increasing degree of travel demand uncertainty through the planning horizon as long-term forecasts have higher uncertainty compared to short-term forecasts. In other words, the uncertainty budget increases over the planning horizon. The robust multi-period TCS design bi-level model is solved using a cutting-plane method (Lawphongpanich and Hearn 2004).

The following are the contributions of this study. First, to the best of our knowledge, this is the first study that proposes to leverage a multi-period TCS to circumvent the need for monetary subsidies to promote the use of low-emissions vehicles. Second, to our best knowledge, this is the first study to provide a hybrid approach that combines mobility management strategies and strategies that promote alternative fuels to reduce vehicular emissions over the long-term. This hybrid approach synergistically aids the attaining of system-level environmental goals while promoting ZEV usage. Third, the proposed approach leads to sustainable behavior of travelers in practice. That is, it incentivizes travelers to use ZEVs by designing TCS parameters that promote travelers' shift to ZEVs because it leads to the higher utility for them. Fourth, the study proposes a robust TCS design to account for demand uncertainty that is inherent to the long-term planning horizon for optimizing the system against the worst-case vehicular emissions.

## 2. METHODOLOGY

This section presents the upper-level model in which the CA aims to design TCS parameters to minimize the worst-case vehicular emissions. Let  $G(N, A)$  be a directed transportation network, where  $N$  and  $A$  denote the sets of nodes and links of a transportation network, respectively. Let  $w \in W$  denote the set of O-D pairs. The planning horizon is divided into  $T$  independent periods of equal length where the length of each period is in order of a year. Let  $\Gamma$  denote the set of time periods. Travel demand and traffic network supply are constant within each period but vary across periods. We consider a mixed traffic scenario consists of ZEV and ICEV travelers where  $M = \{1, 2\}$  denotes the set of modes, in which mode 1 corresponds to ZEV travelers and mode 2 corresponds to ICEV travelers. Let  $d_w^t$  denote the aggregate travel demand rate of O-D pair  $w$  in time period  $t$ . Let  $v_a^{t,m}$  and  $v_a^t$  denote the flow of mode  $m$  and aggregate flow on link  $a$  in time period  $t$ , respectively. Let  $\mu_w^{t,m}$  denote the minimum generalized travel cost of mode  $m$  for O-D pair  $w$  in time period  $t$ .

Let  $e_a^t$  denote the vehicular emissions function of link  $a$  in period  $t$ . As traffic flow consists of two modes, ZEVs and ICEVs, and only the second mode is the source of vehicular emissions, the system vehicular emissions rate of traffic network is equal to  $\sum_{t \in \Gamma} \sum_{a \in A} v_a^{t,2} e_a^t(v_a^t)$  through the planning horizon. Under the multi-period TCS, the CA allocates credits at the average rate of  $\xi^t$  in each period  $t$ . Let  $\xi = \{\xi^t, \forall t\}$  denote the vector of CA-issued credit rates. This study assumes that the CA implements a mode-specific credit allocation scheme under which travelers of each mode  $m$  for O-D pair  $w$  in time period  $t$  receive credits at the rate of  $n_w^{t,m}$ , i.e.  $\sum_{w \in W} \sum_{m \in M} n_w^{t,m} q_w^{t,m} = \xi^t$ . Let  $\mathbf{u} = \{u_a^{t,m}, a \in A, t \in \Gamma\}$  denote the mode-specific link-based credit charging scheme under which travelers of mode  $m$  are charged  $u_a^{t,m}$  credits for using link  $a$  at any time in period  $t$ .

To account for the uncertainty in aggregate travel demand during the planning horizon, it is assumed that the aggregate travel demand rate of each O-D pair in each period during the planning horizon takes one of several values whose occurrence probabilities are unknown. These values of travel demand rates are forecast by transportation planners. The aggregate travel demand in each period  $t$  is assumed to belong to an uncertainty set  $\mathcal{E}$ . Let  $d_w^{t,s}$  denote the aggregate travel demand rate of O-D pair  $w$  in time period  $t$  under scenario  $s$ , where  $d_w^{t,1}$  denotes the average aggregate travel demand of O-D pair  $w$  in time period  $t$ . Let  $\eta_w^{t,s}$  denote the binary variable indicating whether scenario  $s$  realized for O-D pair  $w$  in time period  $t$  is the worst-case demand scenario. The demand uncertainty set,  $\mathcal{E}$ , can be defined as follows:

$$\mathcal{E} = \{ \mathbf{d} \mid \sum_{s \in \mathcal{S}} d_w^{t,s} \eta_w^{t,s} = d_w^t, \sum_{s \in \mathcal{S}_w} \eta_w^{t,s} = 1, \sum_w \sum_{s \in \mathcal{S}_w - \{1\}} \eta_w^{t,s} \leq \Gamma^t, \eta_w^{t,s} \in \{0, 1\} \} \quad (1)$$

where  $\Gamma$  is the uncertainty budget which allows at most  $\Gamma^t$  travel demand rates to deviate from their average values  $d_w^{t,1}$  in each period. The uncertainty budget is used to factor computational burden, and is also higher for forecasts of travel demand rates further into the future, i.e.  $\Gamma^t \geq \Gamma^{t-1}$  for any period  $t$ . The demand uncertainty set also ensures that only one travel demand scenario is realized for each O-D pair in each time period. Then, the upper-level model is formulated as the following mathematical program:

$$\min_{(\xi, u, n)} \max \sum_{t \in \Gamma} \sum_{a \in A} v_a^{t,2} e_a^t(v_a^t) \quad (2)$$

$$\frac{\mu_w^{t,m}}{\mu_w^{t,UE}} \leq \phi^t \quad \forall t, m, w \quad (3)$$

$$d \in \mathcal{E} \quad (4)$$

where  $\phi^t$  denotes the maximum ratio of generalized travel cost after TCS implementation to the generalized travel cost  $\mu_w^{t,UE}$  without TCS (which is labeled as NoTCS) for any O-D pair in each period  $t$ . The upper-level model (2)-(4) is a nonlinear mathematical program with linear and integer constraints. The objective function (2) minimizes the maximum vehicular emissions that manifest under the possible travel demand scenarios. The decision variables in the upper-level are the mode-specific O-D based credit allocation scheme and mode-specific link-based credit charging scheme during the planning horizon. As the CA aims to increase the market penetration of ZEVs, the objective function can be optimized, that is, the worst-case vehicular emissions can be minimized, by allocating more credits to ZEV travelers and charging them fewer credits compared to ICEV travelers. This increases the credit consumption costs of ICEV travelers, motivating travelers to shift to ZEVs. In the limit, this can imply that all travelers will shift to ZEVs within one period to minimize vehicular emissions. However, this is not sustainable in practice as substantial short-term increases in ICEV generalized travel costs are neither politically acceptable nor realistic from the perspective that all travelers will purchase ZEVs within a period. To account for realism in market behavior and to ensure that ICEV travel cost increases are within acceptable norms, the CA seeks to increase ICEV credit consumption costs gradually through the planning horizon so that travelers can better adapt to the multi-period TCS to promote increased ZEV usage over time. Hence, to prevent significant increases in travel costs after TCS implementation, there is need for a constraint (3) which states that the ratio of generalized travel cost after TCS implementation to the one under NoTCS  $\mu_w^{t,UE}$  for each O-D pair in each time period should not exceed the CA-specified ratio  $\phi^t$ . The CA can gradually increase  $\phi^t$  during the planning horizon to motivate travelers to shift to ZEVs while ensuring that they adapt to TCS implementation in a sustainable manner. Constraint (4) includes a set of integer constraints and states that the travel demand scenario should belong to travel demand uncertainty set  $\mathcal{E}$ .

In the lower-level model, travelers minimize their generalized travel costs under the TCS parameters determined in the upper level, through their mode and route choices. The numerical experiments show that the robust design of multi-period TCS leads to higher market penetration of ZEVs, which minimizes worst-case vehicular emissions over the planning horizon. It also reduces the volatility in realized system emissions rate in practice under different travel demand scenarios compared to the system optimal multi-period TCS design that does not factor demand uncertainty. This increases the reliability in reducing vehicular emissions under travel demand uncertainty.

## REFERENCES

- Bertsimas, Dimitris, and Melvyn Sim. 2003. "Robust Discrete Optimization and Network Flows." *Mathematical Programming* 98 (1–3). Springer-Verlag: 49–71. doi:10.1007/s10107-003-0396-4.
- Center for Sustainable Energy. 2017. "California Air Resources Board Clean Vehicle Rebate Project,

- Rebate Statistics.” <https://cleanvehiclerebate.org/eng/rebate-statistics>.
- de Palma, André, and Robin Lindsey. 2011. “Traffic Congestion Pricing Methodologies and Technologies.” *Transportation Research Part C: Emerging Technologies* 19 (6): 1377–1399.
- EPA. 2016. *Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990-2014*. Washington, D.C. <https://www.epa.gov/sites/production/files/2016-04/documents/us-ghg-inventory-2016-main-text.pdf>.
- Global Automakers. 2017. *Quarterly Market Report*. Washington, D.C. [http://drivingzev.com/sites/default/files/blogs/zev\\_market\\_report\\_q2\\_2017.pdf](http://drivingzev.com/sites/default/files/blogs/zev_market_report_q2_2017.pdf).
- Krause, Rachel M, Sanya R Carley, Bradley W Lane, and John D Graham. 2013. “Perception and Reality: Public Knowledge of Plug-in Electric Vehicles in 21 U.S. Cities.”
- Lawphongpanich, Siriphong, and Donald W. Hearn. 2004. “An MPEC Approach to Second-Best Toll Pricing.” *Mathematical Programming* 101 (1). Springer-Verlag: 33–55.
- Miralinaghi, Mohammad, and Srinivas Peeta. 2016. “Multi-Period Equilibrium Modeling Planning Framework for Tradable Credit Schemes.” *Transportation Research Part E: Logistics and Transportation Review* 93: 177–198. doi:10.1016/j.tre.2016.05.013.
- The White House. 2015. “FACT SHEET: U.S. Reports Its 2025 Emissions Target to the UNFCCC | Whitehouse.gov.” *Office of the Press Secretary*. <https://obamawhitehouse.archives.gov/the-press-office/2015/03/31/fact-sheet-us-reports-its-2025-emissions-target-unfccc>.
- Yang, Hai, and Xiaolei Wang. 2011. “Managing Network Mobility with Tradable Credits.” *Transportation Research Part B* 45 (3): 580–594. doi:10.1016/j.trb.2010.10.002.

# AN ALGORITHM FOR THE USER EQUILIBRIUM WITH AGENT-BASED DYNAMIC TRANSIT ASSIGNMENT

M. HICKMAN<sup>a</sup>

<sup>a</sup> *School of Civil Engineering, The University of Queensland, Australia*

*Email: [m.hickman1@uq.edu.au](mailto:m.hickman1@uq.edu.au)*

We consider dynamic models of network assignment and loading in the context of a public transit network. If individual passengers may be described as agents in the network, their choice of paths may be described using the concept of user equilibrium, where no passenger has an incentive to change paths unilaterally. In cases where the vehicle capacity is fixed, those links at capacity will have an adjusted cost that reflects the shadow price of capacity. The application of this shadow price, however, depends on the passenger's level of priority for available capacity, e.g. if they are already on board or have priority over other passengers to board a vehicle at a stop. Hence, we define an agent-based formulation of the passenger assignment in a capacity-constrained transit network with passenger priorities. We propose a solution algorithm that calculates the shadow price of available capacity, dependent on the passenger priority on each link. This algorithm converges to a user (passenger) equilibrium in the full transit network.

**Keywords:** Transit Assignment, Transit Network Loading, Agent-Based Modelling

## 1. INTRODUCTION

There has been considerable interest in simulation-based methods of network loading and assignment among the research in dynamic traffic assignment (DTA). This has been extended to the areas of dynamic public transport assignment, with a variety of agent-based methods implemented over the last 20 years, including MILATRAS (Wahba and Shalaby, 2009, 2011), MATSim (Horni et al., 2017), Dybus (Nuzzolo et al., 2001), FAST-TRIPS (Khani, 2013; Khani et al., 2015), and BusMezzo (Cats et al., 2010; Cats, 2013).

One concern about these simulation-based assignment and network loading methods is the algorithmic performance. Two elements in this respect consider (1) whether the assignment converges to an equilibrium or other testable passenger objective, and (2) whether the assignment is computationally efficient. Of course many of the existing tools noted above are computationally efficient, in the sense that they have been demonstrated on large real-world networks with hundreds of thousands or millions of agents moving within the public transit network in a simulation period.

In this study, we use the theoretical framework of Nguyen et al. (2001), Hamdouch et al. (2004), and Hamdouch and Lawphongpanich (2008) to define the user equilibrium in a dynamic transit network, and we propose an agent-based algorithm to solve for this user equilibrium with a capacity-constrained, schedule-based transit assignment. We also illustrate the computational performance of this algorithm on the public transit network of Brisbane, Australia.

## 2. USER EQUILIBRIUM CHARACTERISTICS

### 2.1 Multi-Commodity Flow Formulation

The capacity-constrained transit passenger assignment problem makes use of the traditional Multi-Commodity Flow (MCF) problem (Ahuja et al., 1993). Each commodity  $k$  represents the necessary flow  $q_k$  between a single origin  $r_k$  and a single destination  $s_k$ . The cardinality of the set  $K$  is  $|K| =$  the number of origin-destination pairs with positive demand (passengers). The constraints in the MCF formulation are:

$$\sum_{k=1}^K \sum_{p \in P_k} \delta_{ij}(p) \cdot f(p) \leq u_{ij} \quad \forall (i, j) \in A \quad (1)$$

$$\sum_{p \in P_k} f(p) = f_k \quad \forall k = \{1, 2, \dots, K\} \quad (2)$$

$$f(p) \geq 0 \quad \forall p \in P_k, k = \{1, 2, \dots, K\} \quad (3)$$

where,  $P_k$  is the set of paths for origin-destination pair  $k$ , and  $f(p)$  is the flow on path  $p$ . The indicator variable  $\delta_{ij}(p)$  indicates whether link  $(i,j)$  is on path  $p$ ,  $\delta_{ij}(p) = 1$  if path  $p$  includes link  $(i,j)$  and  $\delta_{ij}(p) = 0$  otherwise. Constraint (1) shows that all path flows over link  $(i,j)$  cannot exceed the capacity of the link,  $u_{ij}$ . Constraint (2) shows that the sum of path flows for commodity  $k$  must sum to the total origin-destination flow for commodity  $k$ . Constraint (3) ensures non-negativity of path flows.

This same set of constraints will hold for the capacity-constrained transit assignment problem (Carraraesi et al., 1996; Nguyen et al., 2001). Then, a “user equilibrium” transit assignment is one in which no transit passenger can reduce their travel time (or generalized cost) by unilaterally changing paths. This implies:

- (1) All used paths for an origin-destination pair have a travel time (or generalized cost) that is less than or equal to the travel time (or generalized cost) of any unused path.
- (2) For any origin-destination pair, there is no path with available capacity that has a shorter travel time (lower generalized cost) than that of any used path.

With this definition, the complementary slackness conditions for an optimal solution are:

$$c_p^{\sigma, w} = c_k(p) + \sum_{(i,j) \in p} w_{ij} - \sigma_k \quad \forall p \in P_k, k = \{1, 2, \dots, K\} \quad (4)$$

where  $\sigma_k$  is the dual variable for the flow constraint on commodity  $k$ ,  $w_{ij}$  is the dual variable for the capacity of a link  $(i,j)$ , and  $c_k(p)$  is the cost of path  $p$  for commodity  $k$ . Extending this use of  $w_{ij}$ , we can determine the “adjusted” cost of a used path  $p$ , using:

$$\text{Adjusted cost on path } p = c_k(p) + \sum_{(i,j) \in p} w_{ij}$$

The term  $\sigma_k$  can be interpreted as the *minimum* adjusted cost of a used path  $p$ . One may conclude that a public transit assignment must solve for values of  $w_{ij}$  and the minimum adjusted costs by origin-destination pair  $\sigma_k$ . It is also noteworthy that the values of  $w_{ij}$  are not unique, so that we only are seeking one of many possible solutions that satisfy the complementary slackness conditions.

## 2.2 Passenger Priority

However, the passenger equilibrium condition is complicated by priority (Nguyen et al. 2001; Hamdouch et al. 2004). To borrow language from Nguyen et al., we use:

- *additional passenger*: one more passenger that could be added to the existing flow  $q_k$  who could then seek a path
- *switching passenger*: one passenger among the flow  $q_k$  who could switch paths

Nguyen et al. indicate that we may only be able to find an equilibrium if we consider the “additional passenger”, rather than the “switching passenger”. The fact that passenger  $n$  is “switching” paths should **not** free up space on a trip segment **before** they select a new path. Nguyen et al. define:

- *residual capacity*: the number of passenger spaces remaining on a given vehicle trip segment  $ij$ , which we will denote  $r_{ij}$
- *available capacity*: the number of passenger spaces remaining on a given vehicle trip segment  $ij$  for passengers of priority  $l$ , which we will denote  $a_{ijl}$

In this language, residual capacity holds the usual meaning that a vehicle trip still has empty places on board. The notion of *available* capacity is the number of spaces remaining on board, for passengers with a certain level of priority. In this case, if we let  $u_{ij}$  be the vehicle capacity, there could be zero residual capacity ( $r_{ij} = 0$ ), but there may be available capacity for passengers with priority  $l$  if either: (1) the

number of passengers with  $l$  or higher priority is strictly less than  $u_{ij}$ , or (2) the number of passengers on board with priority lower than  $l$  is positive. That is, a passenger of priority lower than  $l$  could be displaced from the vehicle, if someone of priority  $l$  or higher wants to take that place. Then, even though  $r_{ij} = 0$ , we have  $a_{ijl} > 0$ . The implication is that, when a passenger considers switching paths, the new path should have *available* capacity *before* the switch. Equilibrium means that no passenger can improve his/her travel time (or generalized cost) by unilaterally switching to another path *with positive available capacity*.

Finally, on any given link  $ij$ , let us define  $|L|$  priority levels, with a given priority level  $l \in \{1, 2, 3, \dots, L\}$ , level 1 has highest priority, level 2 has the next-lower priority, and so on until we reach level  $L$ , which has the lowest priority. It is certainly possible that  $L$  will depend on the link  $ij$ .

The following constraints define link priority. In (5),  $\delta_{ijl}(p)$  is 1 if path  $p$  uses link  $ij$  and has priority on that link of  $l$ , and 0 otherwise. The available capacity for link  $ij$  for passengers of priority  $l$  is:

$$a_{ijl} = \text{Max} \left( u_{ij} - \sum_{m=1}^{l-1} \sum_{k=1}^K \sum_{p \in P_k} \delta_{ijm}(p) \cdot f(p), 0 \right) \quad (5)$$

Note that the summation includes up to priority level  $l-1$ , meaning that  $a_{ijl}$  represents all available capacity for passengers at level  $l$ . Within level  $l$ , no passenger has priority over any other passenger on link  $ij$ . Yet, we wish to avoid a passenger in level  $l$  displacing another passenger in level  $l$ , creating iterative assignment that may not converge.

This formulation of the available capacity yields the following constraint set:

$$\sum_{k=1}^K \sum_{p \in P_k} \delta_{ijl}(p) \cdot f(p) \leq a_{ijl} \quad \forall l \in \{1, 2, \dots, L\}, \forall (i, j) \in A \quad (6)$$

Or, rewriting by moving the higher-priority flows to the left-hand side gives:

$$\sum_{k=1}^K \sum_{p \in P_k} \delta_{ijl}(p) \cdot f(p) + \sum_{m=1}^{l-1} \sum_{k=1}^K \sum_{p \in P_k} \delta_{ijm}(p) \cdot f(p) \leq u_{ij} \quad \forall l \in \{1, 2, \dots, L\}, \forall (i, j) \in A \quad (7)$$

This constraint set replaces the capacity constraint set for links  $ij$  in the original MCF. Instead, we now have a set of  $L$  constraints per link  $ij$  that address the priority assignment. These constraints (7) also result in a set of dual prices for each link  $ij$ , with one dual price for each level of priority  $l$ . Let us call this dual variable  $v_{ijl}$ , for each link  $ij$  and for priority level  $l$ .

Let us hypothesize that, for a link  $ij$ , there is a priority level  $m$ , with  $1 \leq m \leq L$ , at which the capacity  $u_{ij}$  is reached. That is, we know all passengers from priority 1 to  $m-1$  are able to use the link, while at least one passenger from priority  $m$  is also able to use the link before it reaches capacity. We then know from complementary slackness that the dual variables  $v_{ijl} = 0$  for all  $l$  up to  $m-1$ , i.e. for  $1 \leq l \leq m-1$ . We also know that subsequent dual variables  $v_{ijl} \geq 0$  for all  $l$  such that  $m \leq l \leq L$ .

With the dual variables for capacity on a link with priority:

$$\text{Adjusted cost on path } p = c_k(p) + \sum_{(i,j) \in p} \delta_{ijm}(p) \sum_{l=1}^m v_{ijl} \quad (8)$$

This means that the adjusted cost includes all of the dual prices on link  $ij$  up to priority level  $m$ , the priority level at which at least one user of path  $p$  with priority  $m$  may use link  $ij$ . These dual prices are all equal to zero, and hence there is no further ‘‘adjustment’’ to the cost on path  $p$  up to priority level  $m-1$ . Furthermore, since the dual variables in (8) are all non-negative, all users of link  $ij$  with priority level of  $m$  or higher (i.e., with zero available capacity on the link) should have at least as high a cost of using link  $ij$  than those with a lower priority level. Specifically,

$$\sum_{l=1}^n v_{ijl} \leq \sum_{l=1}^{n+1} v_{ijl} \quad \forall n \in \{1, 2, \dots, L-1\}, \forall (i, j) \in A \quad (9)$$

This means that lower-priority flows will have added costs of at least  $v_{ijm} \geq 0$ . Also, if  $v_{ijl} > 0$  for some priority level  $l$ , all priority levels  $n > l$  have dual prices of at least  $v_{ijl}$ .

### 3. SOLUTION ALGORITHM

Rather than using penalty methods (Nguyen et al., 2001), we propose solving this problem using column generation. The idea is to generate new columns, which represent the minimum travel time paths with available capacity. Then, we shift passengers from paths with no capacity to paths with available capacity. The subsequent step is to update costs for certain links, recognizing that we have dual prices for capacity. These dual prices are the  $v_{ijl}$ .

Figure 1 shows the algorithmic steps to determine shortest paths with these dual prices. This algorithm follows Ahuja et al. (pp. 668-670) but with a few changes. First, we have to account for available capacity in the column generation when generating a new shortest path for a passenger – see step 3.2.2. Second, we have to be conscious of how the dual prices  $v_{ijl}$  are applied, in light of priority. Specifically, once we have estimated  $v_{ijl}$  for a link, that dual price only applies to the current level and lower levels of passenger priority on that link. Passengers with higher priority on  $ij$  do not experience  $v_{ijl}$ . Of course, passengers with lower priority have a dual price that is at least as big as  $v_{ijl}$ .

1. Initialisation
  - 1.1. Calculate the shortest path for each origin-destination pair  $k$  (i.e., commodity  $k$ )  
Let the resulting value of this shortest path be  $c_k(p)$
  - 1.2. Assign the full flow  $q_k$  to this shortest path, for each origin-destination pair  $k$
2. Determine which links  $(i, j)$  in the network are at or over capacity. Let this set of links be  $E$ . If  $E = \emptyset$ , go to Step 5.
3. Choose a link  $ij \in E$ . If  $E$  is empty, go to step 4.
  - 3.1. Find the passenger  $n$  on link  $ij$  with the lowest priority. Let  $k$  be the origin-destination pair for that passenger  $n$ .
  - 3.2. Find the value of  $w_{ij}$  and a new path  $p'$  for passenger  $n$  as follows:
    - 3.2.1. Identify the connector  $m_{ij}$  that allows passenger  $n$  to use link  $ij$
    - 3.2.2. Calculate the shortest path for origin-destination  $k$ , **without** this connector, and using **only** links with positive *available* capacity for passenger  $n$ . Call this path  $p'$ .
    - 3.2.3. Re-assign passenger  $n$  to path  $p'$ .
    - 3.2.4. If it has not yet been calculated, set  $w_{ij}$  to be the difference in the labels for  $j$  and  $i$  in this new shortest path tree:  $w_{ij} = \pi(j) - \pi(i)$
    - 3.2.5. Re-establish the connector  $m_{ij}$  but now with cost  $w_{ij} - c_{ij}$
    - 3.2.6. For any other connectors with *lower* priority than  $m_{ij}$ , also assign a cost of  $w_{ij} - c_{ij}$
  - 3.3. If link  $ij$  is still over capacity, return to step 3.1. Otherwise, remove link  $ij$  from  $E$ , and go to step 3.
4. Go to Step 2.
5. End

Figure 1. Solution Algorithm

Based on this column generation approach, we obtain an user equilibrium for the schedule-based capacitated transit assignment problem. A check could also be performed at the end, by calculating the shortest path for each origin-destination  $k$ , but including the link dual costs  $v_{ijl}$  as appropriate (i.e., in light of priority). This shortest path calculation provides the final value of  $\sigma_k$ . One may then compare this value of  $\sigma_k$  with the travel times for all passengers on origin-destination pair  $k$ .

The algorithm shown in Figure 1 converges to an equilibrium solution, where no passenger has the ability to unilaterally change their route without incurring costs that are at least as large as their current costs. In the full paper, we consider the quality of this equilibrium, and the computational requirements for the real-world transit network from Brisbane, Australia.

## REFERENCES

- Ahuja, R. K., Magnanti, T. L., & Orlin, J. B. (1993). *Network flows: theory, algorithms, and applications*. Prentice-Hall, Englewood Cliffs, New Jersey (USA).
- Carraresi, P., Malucelli, F., & Pallottino, S. (1996). Regional mass transit assignment with resource constraints. *Transportation Research Part B: Methodological*, 30(2), 81-98.
- Cats, O. (2013). Multi-agent transit operations and assignment model. *Procedia Computer Science*, 19, 809-814.
- Cats, O., Burghout, W., Toledo, T., & Koutsopoulos, H. (2010). Mesoscopic modeling of bus public transportation. *Transportation Research Record: Journal of the Transportation Research Board*, (2188), 9-18.
- Hamdouch, Y., Marcotte, P., & Nguyen, S. (2004). Capacitated transit assignment with loading priorities. *Mathematical programming*, 101(1), 205-230.
- Hamdouch, Y., & Lawphongpanich, S. (2008). Schedule-based transit assignment model with travel strategies and capacity constraints. *Transportation Research Part B: Methodological*, 42(7-8), 663-684.
- Horni, A., Nagel, K., & Axhausen, K. W. (Eds.). (2016). *The multi-agent transport simulation MATSim*. London: Ubiquity Press.
- Khani, A. (2013). *Models and solution algorithms for transit and intermodal passenger assignment (Development of FAST-TrIPs Model)*. The University of Arizona.
- Khani, A., Hickman, M., & Noh, H. (2015). Trip-based path algorithms using the transit network hierarchy. *Networks and Spatial Economics*, 15(3), 635-653.
- Nguyen S., Pallottino S. and Malucelli F. (2001) A modelling framework for the passenger assignment on a transport network with time-tables. *Transportation Science*, 35(3), 238-249.
- Nuzzolo, A., Crisalli, U., & Rosati, L. (2001). Dy-bus: A tool for the schedule based dynamic assignment of urban transit networks. In *Proceedings of 9th World Conference on Transportation Research*.
- Wahba, M., and Shalaby, A. (2009). MILATRAS: a new modeling framework for the transit assignment problem. *Schedule-Based Modeling of Transportation Networks*, 1-24.
- Wahba, M., and Shalaby, A. (2011). Large-scale application of MILATRAS: case study of the Toronto transit network. *Transportation*, 38(6), 889-908.

# AN OPTIMIZATION MODEL FOR A STATION-BASED BIKE SHARING SYSTEM REBALANCING PROBLEM

Yufeng ZHANG<sup>a</sup>, Alireza KHANI<sup>b</sup>

<sup>a</sup> *Department of Civil, Environmental, and Geo-Engineering,  
University of Minnesota, USA*

*Email: [zhan4879@umn.edu](mailto:zhan4879@umn.edu)*

<sup>b</sup> *Department of Civil, Environmental, and Geo-Engineering,  
University of Minnesota, USA*

*Email: [akhani@umn.edu](mailto:akhani@umn.edu)*

This paper formulates a dynamic bike repositioning problem for a station-based bike sharing system. Assuming stable and known time-dependent demand patterns, the model optimizes the number of a fleet of trucks needed for conducting rebalancing tasks and, at the same time, the routes for them traveling between stations for picking up and dropping off bikes so as to maximally meet the demand. The formulation takes the handling time of balancing trucks into consideration. The case study on Twin Cities' bike sharing system shows the efficacy of the proposed formulation.

**Keywords:** Dynamic Bike Repositioning, Vehicle Routing Problem

## 1. INTRODUCTION

A substantial number of cities around the world have been implementing bike-sharing systems, and citizens are embracing these systems because of the convenience of using bikes as an alternative and important transportation mode for short distance traveling. In both station-based and dockless bike-sharing systems, stations' unbalanced demands is by no means a negligible problem that may negatively affect users' experience with the system. A bike user can end up with no available bikes when arrives at a station or finds no bikes in the vicinity of one area, especially during peak usage hours. Given a limited total number of bikes in the system, how to move bikes between stations so as to meet the time-dependent demands to the greatest extent during a certain time period (a planning horizon) becomes an interesting problem to both system management agency and transportation researchers.

In the literature, rebalancing/repositioning problems can be categorized into two branches in terms of balancing strategy with one being static system balancing and the other being dynamic system balancing. Dynamic rebalancing strategy aims to balance bikes under the users' intervention, while static rebalancing strategy ignores users' time-dependent activity that affects the availability of bikes at different stations, which usually operates during the night.

In this study, we formulate a dynamic rebalancing optimization problem, where transshipment and multiple visits to one station by the same/different balancing trucks are allowed. It is assumed that the demand for all the stations at all time intervals are known a priori, which is also relatively stable following the assumption users' activities are less likely to change dramatically. The assumption gives the convenience of utilizing a time expanded bike station network.

The rest of the paper follows the structure: the problem is formulated in section 2, a case study is presented in section 3, and lastly, section 4 briefly summaries the paper and indicates the future work.

## 2. PROBLEM SETUP

### 2.1 Basic assumption and goal

The basic assumption made in the study is that demands at all stations at all time intervals are known a priori, although the demands time interval does not need to be consistent with network discretization time interval. This is a somewhat strong assumption, but it is also reasonable to believe over a certain period (say several months in summer), stations usage patterns are relatively stable. Plenty of literature

exists for estimating demands from historic usage data.

Given the known demands, makespan(maximum time that a truck can be scheduled for operation), decide how many trucks are needed and the dynamic vehicle routing plan that can best satisfy the time-dependent demands.

## 2.2 Parameters and Indices

This section introduces the parameters and indices for variables and sets that are used to formulate the bike repositioning optimization problem as shown in the next section.

$S$ : set of stations;

$\mathcal{V}$ : set of balancing trucks;

$\mathcal{A}$ : set of arcs;

$v$ : index for trucks,  $v \in \mathcal{V}$ ;

$i, j$ : index for stations,  $i, j \in S$ ;

$ij$ : index for arcs,  $ij \in \mathcal{A}$ ;

$R$ : index for the origin depot;

$S$ : index for the destination depot;

$t$ : index for time interval,  $t \in \{0, 1, 2, \dots\}$ ;

$CV^v$ : capacity of balancing truck  $v$ ;

$CS_i$ : capacity of station  $i$ ;

$c_{ij}$ : generalized cost for traversing arc  $ij$ ;

$t_{ij}^v$ : time for traversing arc  $ij$ ;

$T^v$ : maximum cumulative travel time for truck  $v$ ;

$ht$ : average handling time for picking up/delivering bikes from/to stations;

$D_i$ : demand at station  $i$ , if there are pick-ups by users, then  $D_i < 0$ , otherwise,  $D_i > 0$ ;

$I_{i0}$ : initial inventory at station  $i$  and depot  $R$ ;

$\alpha$ : unmet demand penalty adjusting parameter;

$n$ : maximum number of bikes that can be handled at one time interval ( $\Delta t$ ), which equals the planning time horizon length ( $T$ ) over handling time;

$Cap$ : capital cost of a balancing truck;

$\beta$ : capital cost depreciation adjusting factor.

## 2.3 Decision Variables

$x_{ij}^v$ : link usage indicator, 1 if truck  $v$  traverses the arc  $(i, j)$ , 0 otherwise;  $x_{ij}^v \in \{0, 1\}$

$y_{ij}^v$ : number of bikes on truck  $v$  using arc  $(i, j)$ ;  $y_{ij}^v \in \mathbb{Z}_+$

$I_i$ : inventory at station  $i$  at the beginning of each time interval;  $I_i \in \mathbb{Z}_+$

$L_i^v$ : number of bikes loaded to station  $i$  by truck  $v$ ;  $L_i^v \in \mathbb{Z}_+$

$U_i^v$ : number of bikes unloaded from station  $i$  by vehicle  $v$ ;  $U_i^v \in \mathbb{Z}_+$

$d_{ij}^v$ : cumulative time duration of truck  $v$  after traversing arc  $ij$ ;  $d_{ij}^v \in \mathbb{R}_+$

$z_v$ : whether to purchase a balancing truck or not;  $z_v \in \{0, 1\}$

## 2.4 Problem Formulation

The objective function is comprised of three terms, which reflects the total travel costs, unmet demands

shortage, the total capital cost of purchasing balancing vehicles, respectively. Noting that in constraint (13), the unmet demands are only penalized when there is a bike shortage. In other words, excess bikes are not penalized since they are capable of meeting some unexpected demand. Constraint set (1) stipulates the flow conservation and also constrain the number of trucks that are needed. Constraint (9) and (10) takes the handing time into consideration.

$$\bullet \text{ Min } \sum_v \sum_{ij} c_{ij}^v x_{ij}^v - \alpha \sum_i ds_i + \beta Cap \sum_{v \in \mathcal{V}} z_v \quad (1)$$

$$\bullet \text{ s.t. } \sum_j x_{ij}^v - \sum_j x_{ji}^v = \begin{cases} -z_v & i=R \\ 0 & i \in \mathcal{S} \setminus \{R, S\} \\ z_v & i=S \end{cases}, \quad \forall v \in \mathcal{V} \quad (2)$$

$$\bullet \quad y_{ij}^v \leq CV^v x_{ij}^v \quad \forall ij \in \mathcal{A}, \forall v \in \mathcal{V} \quad (3)$$

$$\bullet \quad L_i^v \leq \min \left\{ \sum_{j \in \mathcal{S}} y_{ji}^v, n \right\} \quad \forall i \in \mathcal{S}, \forall v \in \mathcal{V} \quad (4)$$

$$\bullet \quad \sum_{v \in \mathcal{S}} L_i^v \leq CS_i - I_i - D_i \quad \forall i \in \mathcal{S} \quad (5)$$

$$\bullet \quad U_i^v \leq \min \left\{ CV^v - \sum_{j \in \mathcal{S}} y_{ij}^v, n \right\} \quad \forall i \in \mathcal{S}, \forall v \in \mathcal{V} \quad (6)$$

$$\bullet \quad \sum_{v \in \mathcal{V}} U_i^v \leq I_i + D_i \quad \forall i \in \mathcal{S} \quad (7)$$

$$\bullet \quad L_i^v - U_i^v = \sum_{j \in \mathcal{S}} y_{ji}^v - \sum_{j \in \mathcal{S}} y_{ij}^v \quad \forall i \in \mathcal{S}, \forall v \in \mathcal{V} \quad (8)$$

$$\bullet \quad d_{ij}^v \leq T x_{ij}^v \quad \forall ij \in \mathcal{A}, \forall v \in \mathcal{V} \quad (9)$$

$$\bullet \quad \sum_{j \in \mathcal{S}} d_{ij}^v - \sum_{j \in \mathcal{S}} d_{ji}^v = \sum_{j \in \mathcal{S}} x_{ij}^v t_{ij}^v + ht(L_i^v + U_i^v) \quad \forall i \in \mathcal{S}, \forall v \in \mathcal{V} \quad (10)$$

$$\bullet \quad I_i^{(0)} = I_{i0} \quad \forall i \in \mathcal{S} \cup R \quad (11)$$

$$\bullet \quad I_i^{(t+1)} = I_i^{(t)} + \sum_{v \in \mathcal{V}} L_i^v - \sum_{v \in \mathcal{V}} U_i^v + D_i^{(t)} \quad \forall i \in \mathcal{S}, \forall t \neq 0 \quad (12)$$

$$\bullet \quad ds_i = \min \{ I_i + D_i, 0 \} \quad \forall i \in \mathcal{S} \quad (13)$$

$$\bullet \quad x_{ij}^v \in \{0, 1\} \quad \forall ij \in \mathcal{A}, \forall v \in \mathcal{V} \quad (14)$$

$$\bullet \quad y_{ij}^v, I_i^v, L_i^v, U_i^v \in \mathbb{Z}_+ \quad \forall ij \in \mathcal{A}, \forall i \in \mathcal{S}, \forall v \in \mathcal{V} \quad (15)$$

$$\bullet \quad d_{ij}^v \in \mathbb{R}_+ \quad \forall ij \in \mathcal{A}, \forall v \in \mathcal{V} \quad (16)$$

### 3. CASE STUDY

A case study on the Twin Cities' bike-sharing system, NICERIDE, was conducted. This section discusses the time-expanded network for bike sharing system first and then shows some experimental results.

#### 3.1 Time-expanded network

Given a planning horizon time length  $T$  and discretization time level  $\Delta t$ , a bike station is expanded into

$(T/\Delta t + 1)$  stations with each representing the station at a discrete time point. Links only exist between two stations when the actual travel time between two stations is less than or equal to the difference of time index for two stations. For the case study,  $\Delta t$  is set to 2 minutes, and planning horizon time length  $T$  is set to 1 hour.

### 3.2 Routing strategies

The travel time between any two stations is evaluated by the shortest path distance, demands for next planning time horizon for all stations are randomly drawn in the interval  $[-8,5]$ . There are 30 stations in the subset of the bike sharing network, which results in 870 nodes and 13,145 links in total for one-hour planning horizon with the problem settings.

With different initial stocks of all the stations, routing strategies vary, the table summarizes the statistics with varied initial stock (for simplicity, we assume equal initial stock for all stations at time 0).

All the tests were executed on a 3.0 GHz Core i5-4590S computer with the 64-bit version of the Windows 7 operating system and 8GB RAM and the optimization problems were solved by Gurobi 7.5.0.

Table 1. Routing strategy with varied initial stocks

| Initial stocks | # of stations visited | Makespan (min) | Total # of rebalanced bikes | Unmet demand | Computational time (s) | Served extra demand |
|----------------|-----------------------|----------------|-----------------------------|--------------|------------------------|---------------------|
| 5              | 15                    | 29.25          | 18                          | 0            | 159.92                 | 7                   |
| 6              | 10                    | 27             | 8                           | 0            | 23.57                  | 2                   |
| 7              | 10                    | 27             | 8                           | 0            | 13.25                  | 1                   |

Table 1 shows that initial stocks influence the routes of balancing trucks, with more initial stocks, balancing trucks tend to have a fewer number of bikes needed to be rebalanced and the makespan is shorter. A direct following step will be testing how unequal initial stocks and penalty parameters change the routing strategies.

### CONCLUSION AND FUTURE WORK

In this study, an optimization problem for bike-sharing system dynamic rebalancing problem is formulated. The objective is to find optimal routes for a fleet of trucks that can best meet the demands at different stations. The formulation takes the handling time of balancing trucks that are needed for picking up and dropping off bikes into consideration, and if calibrated well the number of balancing trucks can also be optimally decided given a relatively stable time-dependent demand pattern. The next step for this work is to design an efficient and effective algorithm for the problem.

### REFERENCES

- Contardo, C., Morency, C. and Rousseau, L.M., 2012. Balancing a dynamic public bike-sharing system (Vol. 4). Montreal: Cirrelt.
- Button, K.J. (1982) *Transport Economics*. Heineman, London.
- Ho, S.C. and Szeto, W.Y., 2014. Solving a static repositioning problem in bike-sharing systems using iterated tabu search. *Transportation Research Part E: Logistics and Transportation Review*, 69, pp.180-198.
- Bulhões, T., Subramanian, A., Erdoğan, G. and Laporte, G., 2018. The static bike relocation problem with multiple vehicles and visits. *European Journal of Operational Research*, 264(2), pp.508-523.
- Raviv, T., Tzur, M. and Forma, I.A., 2013. Static repositioning in a bike-sharing system: models and solution approaches. *EURO Journal on Transportation and Logistics*, 2(3), pp.187-229.
- Chemla, D., Meunier, F. and Calvo, R.W., 2013. Bike sharing systems: Solving the static rebalancing problem. *Discrete Optimization*, 10(2), pp.120-146.

Pal, Aritra, and Yu Zhang. Free-floating bike sharing: solving real-life large-scale static rebalancing problems. *Transportation Research Part C: Emerging Technologies* 80 (2017): 92-116.

