5-2014

Models and Solution Algorithms for Asymmetric Traffic and Transit Assignment Problems

Donghyung Yook
Utah State University

MODELS AND SOLUTION ALGORITHMS FOR ASYMMETRIC TRAFFIC AND TRANSIT ASSIGNMENT PROBLEMS

by

Donghyung Yook

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

DOCTOR OF PHILOSOPHY

in

Civil & Environmental Engineering

Approved:

_____	_____
Dr. Kevin Heaslip Major Professor	Dr. Anthony Chen Committee Member
_____	_____
Dr. Paul Barr Committee Member	Dr. Randal Martin Committee Member
_____	_____
Dr. Daniel Coster Committee Member	Dr. Mark R. McLellan Vice President for Research and Dean of the School of Graduate Studies

UTAH STATE UNIVERSITY
Logan, Utah

2014

ABSTRACT

Models and Solution Algorithms for Asymmetric

Traffic and Transit Assignment Problems

by

Donghyung Yook, Doctor of Philosophy

Utah State University, 2014

Major Professor: Dr. Kevin P. Heaslip
Department: Civil & Environmental Engineering

Modeling the transportation system is important because it provides a "common ground" for discussing policy and examining the future transportation plan required in practices. Generally, modeling is a simplified representation of the real world; however, this research added value to the modeling practice by investigating the asymmetric interactions observed in the real world in order to explore potential improvements of the transportation modeling.

The Asymmetric Transportation Equilibrium Problem (ATEP) is designed to precisely model actual transportation systems by considering asymmetric interactions of flows. The enhanced representation of the transportation system by the ATEP is promising because there are various asymmetric interactions in real transportation such as intersections, highway ramps, and toll roads and in the structure of the transit fares.

This dissertation characterizes the ATEP with an appropriate solution algorithm and its applications. First, the research investigates the factors affecting the convergence of the ATEP. The double projection method is applied to various asymmetric types and complexities in the different sizes of networks in order to identify the influential factors including demand intensities, network configuration, route composition between modes, and sensitivity of the cost function. Secondly, the research develops an enhancement strategy for improvement in computational speed for the double projection method. The structural characteristics of the ATEP are used to develop the convergence enhancement strategy that significantly reduces the computational burdens.

For the application side, instances of asymmetric interactions observed in in-vehicle crowding and the transit fare structure are modeled to provide a suggestion on policy approach for a transit agency. The direct application of the crowding model into the real network indicates that crowd modeling with multi user classes could influence the public transportation system planning and the revenue achievement of transit agencies. Moreover, addition of the disutility factor, crowding, not always causes the increase of disutility from the transit uses. The application of the non-additive fare structure in the Utah Transit Authority (UTA) network addresses the potential of the distance-based fare structure should the UTA make a transition to this fare structure from their current fare model. The analysis finds that the zero base fare has the highest potential for increasing the transit demand. However, collecting less than $0.50 with a certain buffer distance for the first boarding has potential for attracting the users to UTA's transit market upon the fare structure change.

(209 pages)

PUBLIC ABSTRACT

Models and Solution Algorithms for Asymmetric

Traffic and Transit Assignment Problems

by

Donghyung Yook, Doctor of Philosophy

Utah State University, 2014

Major Professor: Dr. Kevin P. Heaslip
Department: Civil & Environmental Engineering

Generally, modeling is a simplified representation of the real world; however, this research adds value to the modeling practice by investigating the asymmetric interactions observed in the real world in order to explore potential improvements of the transportation modeling. The enhanced representation of the transportation system by the asymmetric transportation equilibrium problem is promising because there are various asymmetric interactions in real transportation such as intersections, highway ramps, toll roads and in the structure of the transit fares.

The dissertation considers the asymmetric interactions of flows in the traffic and transit assignment problems with an efficient solution algorithm. The study begins with characterizing the asymmetrically formulated equilibrium problem in terms of the convergence, and the computational efficiencies followed by demonstrating the enhanced modeling of the real-world transportation problems with asymmetric interactions in a public transportation system.

ACKNOWLEDGMENTS

I would like to thank my adviser, Professor Kevin Heaslip, for his support, guidance, and encouragement throughout my Ph.D course at USU. I also want to thank Professor Anthony Chen for his academic guidance for improving the dissertation. My wife, Jihui, who is always encouraging me with her beautiful smile, and our first baby, Haiwon, and my parents and parents in law in South Korea have been waiting for the completion of this dissertation. I would like to dedicate this one to my family.

On Parents' Day in 2014

Donghyung Yook

CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

NOTATIONS

$\Phi = (N, A)$	network represented by a set of nodes N and a set of arcs A
N	a set of nodes
A	a set of arcs
TA	a set of time expanded arcs
R	a set of origins $R \in N$
S	a set of destinations $S \in N$
b	base fare
β	step-size
c_a	cost of link a
c_{rs}	user cost for an O-D pair connecting an origin r and destination s
C	cost vector
CP_a	capacity of link a
$d(\cdot)$	demand function
$\delta_{rs}^{p,a}$	incidence variable that takes one if a route p connecting an origin r and destination s pass through link a or line l, otherwise zero.
δ_{ea}	one if a line segment e is the first part of line section a
e	elasticity of demand
ε	passenger car equivalent
f_{rs}^{p}	route flows p connecting an O-D pair between r and s

frq_l	frequency of line l
g	in-equality constraints
H	headway
h	equality constraints
iv	in-vehicle transit time
iw	initial waiting time
L	pool of lines
m	mileage fare
M	vehicle types or modes
N	number of population
Ω	a set of feasible route flows
P_{rs}	a set of paths connecting an O-D pair between r and s
$\varphi(\cdot)$	transit operating cost function
φ_l	operating cost per passenger-mile for line l
$\varphi_{iv}, \varphi_{iw}, \varphi_{xf}, \varphi_{xfw}$	weights on the time component
ψ_l	additional fare for using premier service l
ρ	asymmetric complexities
λ	buffer distance
r^l	revenue allocated to line l
s	in-vehicle distance
$s_a(\cdot)$	mode interaction function on link a
sb	subsidy

σ	satisfaction level of a constraint
t	average travel time
t^{free}	free-flow link travel time
t_w	average wait time
τ	threshold value that determines if a route satisfies the UE condition
θ_{iv}	coefficient of in-vehicle time
θ_{iw}	coefficient of initial waiting time
θ_{xf}	coefficient of number of transfers
θ_{μ}	coefficient of fare
θ_{xfw}	coefficient of transfer waiting time
θ_{prk}	coefficient of parking cost
θ_{opr}	coefficient of operating cost
θ_{toll}	coefficient of toll fare
q_{rs}	demand between an origin r and a destination s
q_d	ridership under distance-based fare
μ	fare
μ_x	transfer fare
μ_f	flat fare
μ_d	distance-based fare
u	unit fare

V^m	deterministic utility for taking alternative m
v_a	flows of link a
$v_{a'}$	flows on the link opposite to a
V	flow vector
vc_e	vehicle capacity of line segment e
w	weight factor
ω_{rs}^p	route cost of p connecting an origin r rand a destination s
xf_{rs}	number of transfers
xfw_{rs}	transfer waiting time
ξ	measure of travel time equivalent to one dollar of fare paid
ΔT	change of total ridership
$\Delta\varpi$	difference between willingness to pay by users and actual cost for using transit
$z(\cdot)$	objective function
$\zeta(f)$	amount of flows that deviates from the shortest route under UE condition

CHAPTER 1

INTRODUCTION

Analysis of urban transportation networks begin with estimating flows that traverse a given transportation facility in a unit time. The first step of this analysis constructs a functional relationship between the behaviors of travel units (passengers, vehicles, and pedestrians) and the flows and congestion within the transportation network. The estimated flow patterns within the network provide the researchers with insight on the efficiency and effectiveness of the transportation network. Generally, urban transportation network users travel in a way that minimizes transportation disutility. For example, motorists driving between a given origin and a destination are likely to choose the route with the minimum travel time. For this research, the disutility is represented by travel cost. The Network Equilibrium assignment Problem (NEP) mathematically formulates the functional relationship in that users pursue the minimum travel cost to their destinations in a way that all users sharing the same origin and the destination cannot reduce their travel cost. Namely, the NEP determines the flow patterns that ensure the users have the same minimum travel cost to their destinations (Sheffi 1985).

In the NEP, the assumption on the functional relationship, which limits the impact of the travel cost on a given link to the flow on only that link, has been a long tradition (Dafermos 1972). However, the assumption restricts the model development toward better simulation of the realistic transportation systems. By relaxing this restriction that allows asymmetric relationships, a more realistic model for the NEP is expected. This is

the motivation behind designing the Asymmetric Transportation Equilibrium Problem (ATEP).

Observing various situations of traffic and/or passenger flows interacting asymmetrically in both traffic and public transportation systems is the goal of ATEP. The sources of asymmetric interactions are classified into three groups: links, routes, and modes. The link interaction is when the unbalanced marginal effects of one or more links affect the performance of a given link. When the marginal effects are too minimal and intertwined to be explicitly modeled within the link cost function, the interactions are aggregated in the route level. The typical example of this situation includes the non-additive route cost, where the sum of each link's cost comprising one route is different from the final route cost. This non-additivity is also found in the transit fare structure where the total fare is determined by the base fare and the additional mileage fare after a certain in-vehicle distance. Flow interactions are also observed among modes. In the field, varied types of vehicles share the road, influencing each other. This interaction can be characterized by classifying the mixed vehicles by operational characteristics of these varied vehicles on the road, such as cars and heavy trucks.

Adopting these flow interactions into the traffic situations where the asymmetric interactions are expected will significantly improve the model's ability to represent the real world. Conversely, insufficient consideration for flow interactions in the ATEP could result in incorrect travel cost estimation; which in the network equilibrium process is the basic measure of how users choose their routes. Generally, in the urban areas about 40% of delay is attributed to intersections, a typical location causing link interactions due to delay estimation (FHWA 2004). Thus, assigning traffic flows on a link with an

intersection without the consideration for the intersection delay could result in distorted route travel cost estimation, especially the routes that pass through the urbanized areas.

Although enhancement of the realistic model development is possible with the ATEP, allowing asymmetric interaction requires a more general mathematical formulation. Unlike the conventional NEP formulation that considers classical, simple form of travel cost, in which the cost is assumed to be independent of the flows on any other link, readily ensures the optimality and the uniqueness of the solution by taking derivative on the objective function, in the case of asymmetric interactions, an equivalent mathematical program cannot be derived by using the same argument as used in the symmetric case due to absence of the objective function. Alternatively, a gap function that measures the distance to the exact solution was utilized for substituting the objective function in the different formulation for the NEP (Gabriel and Bernstein 1997; Lo and Chen 2000). However, the general approach to deal with this challenge will be Variational Inequalities (VI) introduced by Smith (1979) and Dafermos (1980). Thanks to the special formulation that does not necessitate the objective function the VI can represent more relaxed conditions than the conventional function can.

The ATEP formulated with the VI allows more realistic modeling of the real world with asymmetric situations. This dissertation attends two aspects of the ATEP; possibilities for enhanced modeling of the real transportation system and the solution algorithms for the ATEP. Before investigating the reasonableness of the modeling through the ATEP, the study begins with research on the convergence of the ATEP with the efficient solution algorithm. For the ATEP, lack of research is found that investigates the influential factors imposed on the network data, or on the functional structure for the

conditions of convergence (Patriksson 1994). This research aims to investigate this with one projection method based on the extra-gradient theory (Khobotov 1987) owing to its favorable features for handling the ATEP. This method is appropriate for handling the varied situations of the ATEP because it does not require derivative information for seeking solutions and operates on the relaxed cost function condition.

The dissertation approaches the issues with the extra-gradient method for both traffic and transit assignment. The construction of the dissertation that applies the same solution algorithm to the traffic and passenger assignment is designed to provide a ground for the integrity of the extra-gradient based solution algorithm for the ATEP. The part that attends the asymmetric interactions in the traffic assignment focuses on identification of the influential factors on the convergence and on the enhancement of the algorithm in terms of computational efficiencies. The other part addresses the practical application of the ATEP in the transit assignment. In doing so, the objective of the research that address the potentials of the ATEP in terms of the solution algorithm and its application will be achieved. The sub-research objectives that will be dealt in each part of the dissertation are as follow:

1) Identify the influential factors that affect the convergence of the asymmetrically formulated traffic assignment problem
2) Develop an acceleration strategy for improving the computational speed of the extra gradient based-algorithm for the ATEP
3) Model the ATEP to provide insight on crowding effect and fare policy approaches for the transit system

The research begins with investigating the asymmetric situations in both traffic and public transportation systems. These asymmetric situations include situations where external flows influences on the given cost function:

- two-way streets
- signalized intersections
- un-signalized intersections
- varied vehicle types
- non-additive route fare in the transit fare structure

Each of these interactions is categorized by the external factor's source such as link, mode and route. Given the identified asymmetric interaction cases, the study investigates the influential factors of the ATEP's convergence. Specifically, the task is based on the Double Projection (DP) method introduced by Panicucci et al. (2007). Also, the convergence acceleration strategy for the DP method will be proposed for the computational enhancement.

The research attends the advanced ATEP modeling on the passenger assignment model, which accounts for the crowding effect and non-additive transit fare structure. The crowding effect influenced by the daily variation of passenger flows in a transit vehicle is modeled through the ATEP and expected outcomes will be investigated. In addition, the dissertation aims to provide an analysis of a proposed change to the Utah Transit Authority (UTA) fare policy by properly implementing the non-additive fare structure.

To achieve these objectives, a systematic approach has been organized that attends the following issues:

1. Identify the situations incurring asymmetric interactions through the extensive investigation and the literature review

2. Determine the factors that influence the convergence for known asymmetric modeling approaches
3. Develop an acceleration strategy for the DP method in such a way that computational efficiencies are achieved
4. Crowd modeling is constructed for the UTA network under the ATEP in the passenger assignment
5. Investigate the behavioral changes due to the crowd modeling and its consequences in the network level
6. Model the non-additive asymmetric transit fare system with the effective representation of the transit assignment modeling
7. Examine the potential transition to the distance-based fare structure of the UTA and the appropriate fare levels by modeling the non-additive fare structure.

CHAPTER 2

LITERATURE REVIEW

Through the algorithmic and the application perspectives, the dissertation examines the issues of the ATEP. The literature review addresses the research related to these two aspects of the ATEP. The review will first explain the basic concepts of asymmetric interactions with a simple example and identify a variety of situations in the transportation systems. Exploring various asymmetric situations intends to provide readers with insights to the ATEP's useful modeling practices for representing the real world. Following the descriptive explanations, the review presents the model formulation of the ATEP. In the section, the appropriateness of the VI formulation for the NEP is explained when it is associated with the asymmetric interactions, or the ATEP. Addressing the ATEP in the mathematical formulation is not limited only to the VI, but could go with the nonlinear complementarity problem (NCP), fixed point problem. As mentioned in the introduction, this study considers the VI because of it generality. The VI is equivalent to NCP when the feasible solution space is constrained by the positive real number, and is a special case of the fixed point problem which is normally used for establishing the existence of solutions to the VI (Patriksson 1994). Then, the review addresses how existing studies solve the ATEP. Various studies utilized existing solution algorithms developed for the NEP to solve the ATEP. The literature review also attends analytic approaches that dealt with the uniqueness issue associated with the ATEP. The following section touches the double projection method’s properties that make the method suitable for the ATEP. The final two sections of this chapter are about the reviews on the applications where the ATEP is modeled in the transit assignment. The

former one summarizes the researches that studied the crowd modeling in a transit vehicle. The literature that modeled the crowding effect in the transit assignment will be summarized in the section. The latter one pays attention particularly to the fare models associated with the non-additive fare structure. The research will describe how the fare structure is modeled differently in the literature.

A Simple Case of Asymmetric Interactions – Two Way Traffic Interaction

Consider a network represented by $\Phi = (N, A)$ consisting of a set of arcs A associated with a positive cost c_a ($a \in A$) and a set of nodes N . A simple example of link interaction involves pair-wise relationship between the two (opposite direction) links representing two-way streets (Sheffi 1985). As traffic builds up in one direction, the flow delay in the opposite direction increases due to the reduction in passing opportunities. This effect can become quite pronounced in moderate to heavy flow. Thus, the travel cost in a given direction is considered to be a function of the flow in both directional flows.

This interaction can be modeled as

$$c_a = c_a(v_a, v_{a'}) \qquad \forall a, a' \in A, a \neq a' \tag{2.1}$$

$$c_{a'} = c_{a'}(v_{a'}, v_a) \qquad \forall a, a' \in A, a \neq a' \tag{2.2}$$

where $c_{a'}, v_{a'}$ denotes the link cost and flows on link a' which is opposite to link a .

The interaction between the opposite directional flows can be assumed symmetric or asymmetric. Symmetric assumption implies that the effect of an additional flow unit along a particular link on the travel cost in the opposing direction equals the effect of an

additional flow unit in the opposing direction on the travel cost of the link under consideration. This symmetry condition can be expressed mathematically as

$$\frac{\partial c_{a'}(v_{a'}, v_a)}{\partial v_a} = \frac{\partial c_a(v_a, v_{a'})}{\partial v_{a'}} \qquad \forall a, a' \in A, a \neq a' \tag{2.3}$$

However, the interaction between the opposite directional flows is not always symmetric. For example, the addition of a flow unit into the congested directional flow on the opposite link cost is more significant than adding the flow unit to the opposite direction or less congested flow on the congested link cost simply due to passing ability.

$$\frac{\partial c_{a'}(v_{a'}, v_a)}{\partial v_a} \neq \frac{\partial c_a(v_a, v_{a'})}{\partial v_{a'}} \qquad \forall a, a' \in A, a \neq a' \tag{2.4}$$

This asymmetric interaction is a more relaxed condition than the symmetric one, which enables the flexible modeling of the real transportation system. Different asymmetrical interaction models can be developed based on this basic asymmetrical example.

The flow interactions of this simple example are based on the interactions among flows from different links. The case can be extended to the interactions among different mode's flows. Implicit asymmetric interactions are assumed when the sum of each link's cost comprising one route differs from the final route cost. The sources of these interactions are classified as links, modes and routes. This review explores the various cases of asymmetric interactions within these sources and elaborates on them by expanding the given example with more generalized formulations.

Asymmetric Situations in Transportation Systems

Link Interactions

Numerous traffic situations incur asymmetric link interactions. Flow interactions involving more than one link can be observed in (a) two-way streets, (b) the highway merging ramp, (c) un-signalized intersections due to turning and through movements and (d) at signalized intersections due to delay (Dafermos 1980; Sheffi 1985; Grange and Munoz 2009; Chen et al. 2011a). Simple sketches of these asymmetric traffic situations are illustrated in Figure 2.1. On a two-way street, one lane for each direction, vehicles consider oncoming traffic before passing (Figure 2.1.a). A similar situation is also applied to the on-ramp on a highway where the merging flows attempt to find a suitable gap from the adjacent highway traffic flows (Figure 2.1.b). At an un-signalized intersection, the ability to safely turn left or continue forward is affected by the opposing direction's flow (Figure 2.1.c). At the signalized intersection, incorporating the delay into the NEP also causes asymmetric link interactions because the delay estimation at an intersection considers all approaching traffic flows headed to the intersection (Figure 2.1.d). This delay estimation is usually modeled by Webster's delay model (1958). Readers interested in the equilibrium traffic assignment combined with the intersection delay may refer Meneguzzer (1995).

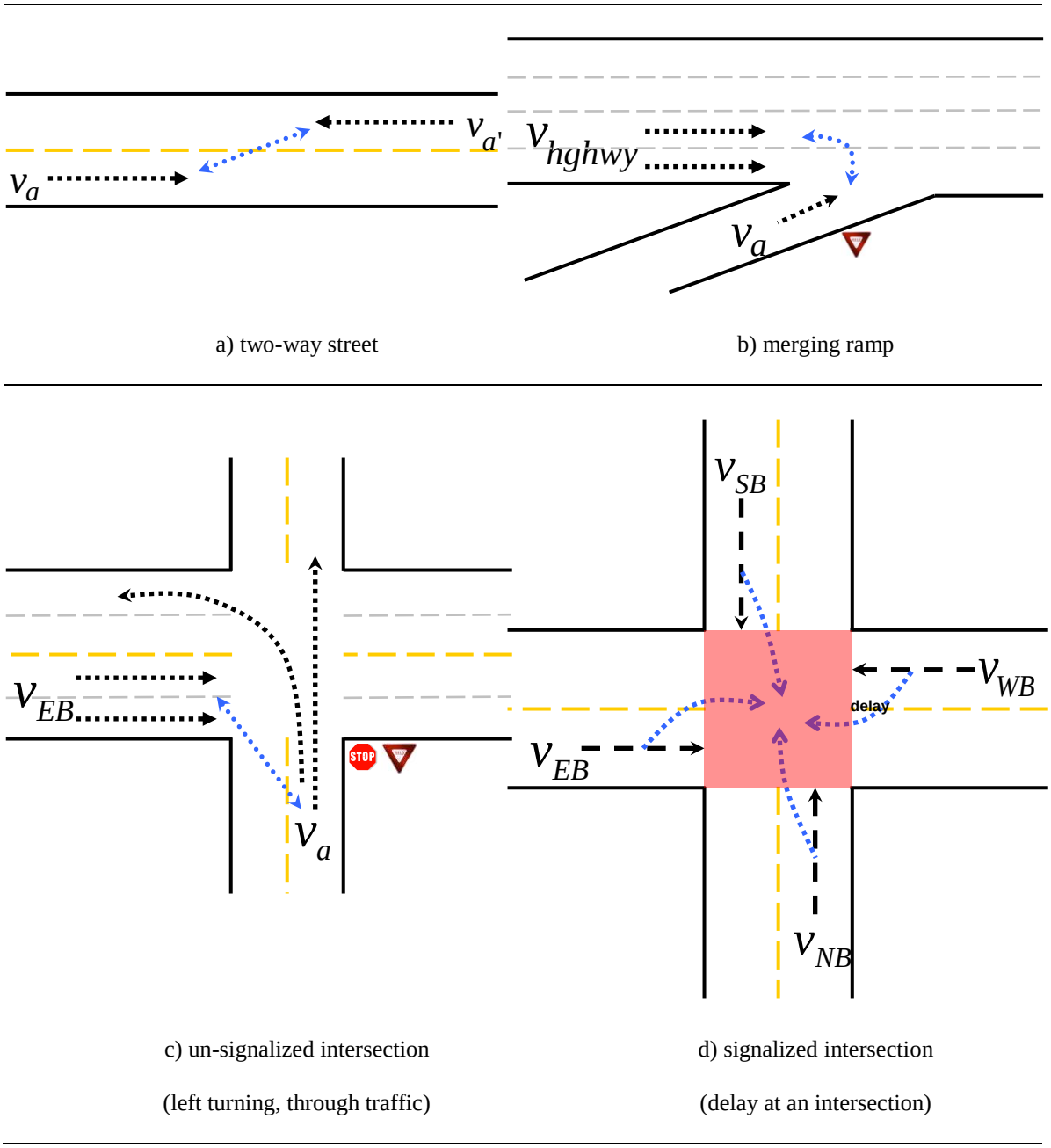


Figure 2.1 Situations bearing the potential of asymmetric link interactions.

The simple mathematical formulation for the two-way street interaction model can be extended to a general link interaction model as

Consider a network consisting of links with link flow vector V .

$$V = (v_1, v_2, v_3, ..., v_N) \tag{2.5}$$

In general, a vector of link costs C may include the whole link flow pattern V .

$$C = \{c_1(V), c_2(V), c_3(V), ..., c_N(V)\} \tag{2.6}$$

However, if the link cost functions are assumed to be independent and separable, then

$$C = \{c_1(v_1), c_2(v_2), c_3(v_3), ..., c_N(v_N)\} \tag{2.7}$$

Asymmetric link interactions occur when the link cost functions are non-separable and asymmetric as

$$\frac{\partial c_j(V)}{\partial v_i} \neq \frac{\partial c_i(V)}{\partial v_j} \quad \forall i, j \in A, i \neq j \tag{2.8}$$

Mode Interactions

Mode interactions are the situations where several vehicle types are influencing each other on the same transportation network. These interactions typically occur between vehicles with very different maneuver and operation characteristics such as between trucks, passenger cars, and buses. This could be further detailed as heavy trucks, light trucks, passenger cars, and emergency vehicles (Dafermos 1972).

When interactions among different types of vehicles are modeled, the dimension of the flow vectors and the cost functions are expanded to the number of vehicle types,

$$V^m = (v_1^m, v_2^m, v_3^m, ..., v_N^m), \ \ m = 1, 2, ..., M \tag{2.9}$$

where M is number of vehicle types. Also, a vector of the link cost for each vehicle type may include the whole flow pattern V^m of all vehicle types. Thus, each vehicle type l has an individual cost function while contributing to the cost function of other vehicles.

$$C = \{c_1^l(V^m), c_2^l(V^m), c_3^l(V^m), ..., c_N^l(V^m)\}, \quad m = 1, 2, ..., M \tag{2.10}$$

Asymmetric mode interactions refer the situation when the link cost function of a vehicle type is non-separable and asymmetric as

$$\frac{\partial c_a^q(V^M)}{\partial v_a^p} \neq \frac{\partial c_a^p(V^M)}{\partial v_a^q} \quad \forall p, q \in M, p \neq q \tag{2.11}$$

Route Interactions

Route interactions are usually represented by the non-additive route cost, where the sum of individual link's cost comprising one route is different from the final route cost. The non-additive route cost is incurred when an individual link's impact cannot be explicitly modeled or the allocation of its impact is unknown. An example is provided to illustrate the interaction. Consider one Origin-Destination (O-D) pair network with two links serially connected, as shown in Figure 2.2, and the route cost function used in the numerical test in Gabriel and Bernstein (1997).

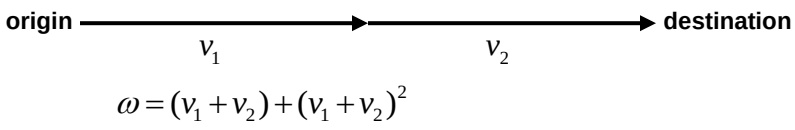


Figure 2.2 A small example for non-additive route cost.

With this route cost function, there is more than one way to define the individual link's impact on another link due to the term $v_1 v_2$. Furthermore, even if the term $v_1 v_2$ is evenly allocated to each link, the link cost function is asymmetric.

$$\begin{aligned} c_1 &= v_1^2 + v_1 + v_1 v_2 \\ c_2 &= v_2^2 + v_2 + v_1 v_2 \end{aligned} \tag{2.12}$$

Under the configuration, it is not possible to define a link cost symmetrically because there is more than one way to allocate $v_1 v_2$ to each link. This example shows that when the route cost is non-additive we cannot specify the link cost function but also the link cost function is not symmetric. Thus, this route-based function necessitates the domain of the solution space to be the route (Gabriel and Bernstein 1997). If we can specify each link's cost function, then it becomes the case of link interactions. For this example, the non-additive route cost is a more generalized form of the asymmetric link interaction.

The concept of non-additivity is widely used in the field of transportation studies such as valuation of the travel time, uncertainty, environmental issues, and the toll charging mechanism. The non-additivity in the time related concepts is implied when time evaluation does not linearly increase as the travel time gets longer. This non-linearly perceived travel time is applied to the uncertainty of the travel time. In order to arrive at their destinations with a predefined reliability threshold (Chen and Ji 2005; Lo et al. 2006; Shao et al. 2006; Siu and Lo 2006), people generally reserve a certain amount of buffer time (Travel Time Budget, TTB). This TTB representing reliability of on time arrival also increases non-linearly. Moreover, non-linearity of travel time evaluation

could be more concrete when the unreliability aspect of the travel time is considered (Chen and Zhou 2010) in addition to the TTB.

Unlike the features that implicitly affect the non-additive route cost, the toll charging mechanism explicitly characterizes the non-additivity. The toll fare mechanism is not additive, which necessitates the operation space of the NEP to be the route (Gabriel and Bernstein 1997; Lo and Chen 2000; Yang et al. 2004). Gabriel and Bernstein (1997) provide a good example. User1 and User2 in Figure 2.3 each pay a $2 toll per sub-section for using entire toll road used by User3 paying $3, which is not equal to the sum of sub-sections' tolls.

The non-linear relationship is also found between automobile operations and emission effects. Hydrocarbons and carbon monoxide gases from vehicle emissions are a non-linear function of the travel time (Gabriel and Bernstein 1997; Wallace et al. 1998).

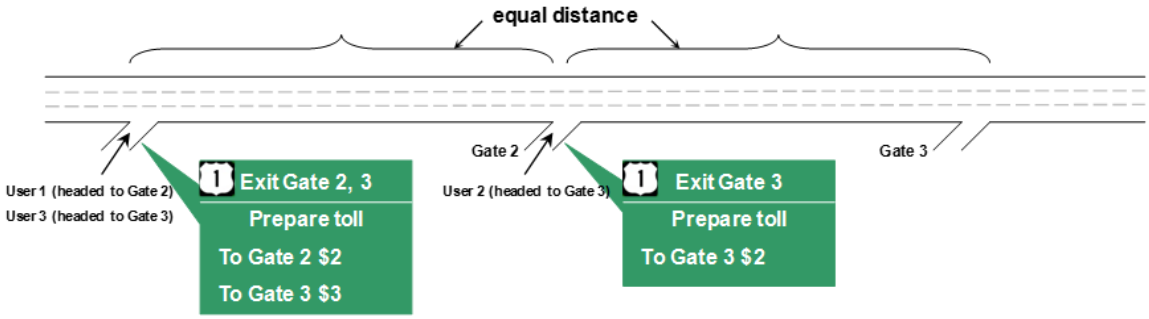


Figure 2.3 Non-additivity in the toll charging system.

The general formulation of the non-additive route cost function can be written as follows (Gabriel and Bernstein 1997):

$$\omega_{rs}^{p} = \theta \sum_{a \in A} \delta_{rs}^{p,a} c_a + g_p \left(\sum_{a \in A} \delta_{rs}^{p,a} c_a \right) + \gamma_{rs}^{p} \tag{2.13}$$

where θ is the operating cost per unit travel time (e.g., fuel consumption, vehicle rental cost), $g_p(\cdot)$ is a function that translates the value of travel time, which could be non-linear, and γ_{rs}^{p} is the out-of-pocket cost (such as tolls, transit fares). $\delta_{rs}^{p,a}$ is an incidence indicator which takes one of the route p connecting the origin r and the destination s passes through the link a; otherwise it is zero.

So far, the various asymmetric situations in transportation systems and its general mathematical formulation have been addressed. In the next section, the mathematical formulation of the NEP that holds identified asymmetrically formulated cost functions is addressed. Identified interactions are considered in the cost function that models the congestion impacts of flows in the NEP.

Formulation of the Network Equilibrium Problems with Asymmetric Cost Function

This section discusses the model formulation for the NEP that finds the User Equilibrium (UE) flow patterns with the asymmetric cost function. The demand associated with each O-D pair is assumed to be fixed throughout the chapter. The basic mathematical transformation of the NEP (Beckmann et al. 1956) that finds the flow pattern achieving the UE is presented first without consideration for the asymmetric interactions. When the cost function is asymmetrically formulated, the limit of the basic formulation for covering the ATEP is simply explained. This is followed by the VI formulation that accommodates the asymmetric cost functions into the NEP.

The network equilibrium problem is to find the flows that satisfy the UE condition which describes the steady-state travelers' behavioral patterns known as the Wardop's first principle (1952) as : "The journey time on all the routes actually used are

equal, and less than those which would be experienced by a single vehicle on any unused route."

The mathematical formulation equivalent to the definition of the UE can be presented as follows:

$$\begin{aligned} & f_{rs}^{p}(c_{rs}^{p}-\pi_{rs})=0, \quad p\in P_{rs}, r\in R, s\in S \\ & c_{rs}^{p}-\pi_{rs}\geq 0, \quad p\in P_{rs}, r\in R, s\in S \\ & \pi_{rs}\geq 0, \ r\in R, s\in S \\ & f_{rs}^{p}\geq 0, \ r\in R, s\in S \end{aligned} \tag{2.14}$$

where

f_{rs}^{p} : route flows of p connecting origin r and destination s

P_{rs} : set of routes connecting origin r and destination s

c_{rs}^{p} : route cost of p connecting origin r and destination s

π_{rs} : minimum cost between origin r and destination s

R, S : subset of N

Eq. (2.14) is the mathematical expression for the UE condition. Beckmann et al. (1956) formulated the equivalent minimization problem that finds the UE flow f^{*} satisfying eq. (2.14). In a network represented by graph $\Phi=(N,A)$, the Beckmann's transformation finds the traffic flow pattern by allocating the O-D demands to the network such that all used routes between a given O-D pair have equal travel cost, and no unused route has a lower travel cost.

$$\min \sum_{a\in A}\int_{0}^{v_a} c_a(w)dw \tag{2.15}$$

subject to

$$\sum_{p \in P_{rs}} f_{rs}^{p} = q_{rs} \quad \forall r \in R, s \in S \tag{2.16}$$

$$\begin{aligned} v_a &= \sum_r \sum_s \sum_p f_{rs}^{p} \delta_{rs}^{p,a} \quad \forall a \in A \\ f_{rs}^{p} &\geq 0 \qquad\qquad\qquad \forall p \in P_{rs}, r \in R, s \in S \end{aligned} \tag{2.17}$$

where

q_{rs} : fixed demand generated from origin r to destination s

$\delta_{rs}^{p,a}$: one if route p connecting r and s passes link a, zero otherwise

The above transformation to the equivalent optimization problem could be established because the cost function c is a gradient mapping. This gradient mapping is easily achieved by the assumption that the travel cost on a link is separable, which means the cost is independent of the flow of all other links in the network. This can be mathematically rewritten as:

$$\begin{aligned} \frac{\partial c_a(v_a)}{\partial v_b} &= 0 \qquad \forall a, b \in A, a \neq b \\ \frac{\partial c_a(v_a)}{\partial v_a} &\geq 0 \qquad \forall a \in A \end{aligned} \tag{2.18}$$

More generally, if the jacobian matrix $\nabla c(v)$ is symmetric by the symmetry condition (e.g. eq. (2.3)) that includes the separable case (eq. (2.18)), then the equivalent optimization problem is defined with the only change being that the property of an increasing function is replaced by the more general property of monotonicity (Patriksson 1994).

The following example of a symmetric two-way street cost function given in eq. (2.1) and eq. (2.2) and the symmetry condition of eq. (2.3), explains the construction of the equivalent mathematical programming.

Under the configuration, the line integral over the two link flows can be defined as two ways; one (z_1) goes through $v_a, v_{a'}$, the other one (z_2) with the opposite direction.

$$z_1 = \int_0^{v} c(w)dw = \int_0^{v_a} c_a(w,0)dw + \int_0^{v_{a'}} c_{a'}(v_a, w)dw \tag{2.19}$$

$$z_2 = \int_0^{v} c(w)dw = \int_0^{v_{a'}} c_{a'}(0,w)dw + \int_0^{v_a} c_a(w, v_{a'})dw \tag{2.20}$$

Using the Leibniz integral rule and the symmetry condition, the derivative of z_1 with respect to v_a is easily identified as $c_a(v_a, v_{a'})$.

$$\begin{aligned}
\frac{\partial z_1}{\partial v_a} &= c_a(v_a,0) + \frac{\partial}{\partial v_a}\int_0^{v_{a'}} c_{a'}(v_a,w)dw \\
&= c_a(v_a,0) + \int_0^{v_{a'}} \frac{\partial c_{a'}(v_a,w)}{\partial v_a} dw \\
&= c_a(v_a,0) + \int_0^{v_{a'}} \frac{\partial c_a(v_a,w)}{\partial v_{a'}} dw \\
&= c_a(v_a,0) + c_a(v_a,v_{a'}) - c_a(v_a,0) \\
&= c_a(v_a,v_{a'})
\end{aligned}$$

In the same way, the derivative of z_1 with respect to $v_{a'}$ is equal to $c_{a'}(v_a, v_{a'})$. Thus, the total cost results in $c_a(v_a, v_{a'}) + c_{a'}(v_a, v_{a'})$. The symmetry condition yields the same total cost when z_2 is differentiated with respect to $v_a, v_{a'}$, respectively.

The matched total cost with different paths of the line integral indicates that the construction of the objective function is possible, which expands to the equivalent mathematical programming for the NEP associated with the symmetry condition. In this case, the objective function may be expressed as follows:

$$\min z(V) = \frac{1}{2}[\sum_{a\in A}\int_0^{v_a} c_a(w, v_{a'})dw + \sum_{a\in A}\int_0^{v_a} c_a(w,0)dw] \tag{2.21}$$

Generally, the validity of the non-separable symmetry condition can be explained with the Green's theorem (1828). When c_a is non-separable and symmetrically formulated, the optimization formulation, eq. (2.15 ~ 2.17), is still valid for the UE because a line integral can be converted to a standard integral due to the Green's theorem (1828) where the line integral is independent on the path of integration on the symmetric jacobian of the cost function.

$$\oint_O (Ldx + Mdy) = \iint_D \left(\frac{\partial M}{\partial x} - \frac{\partial L}{\partial y}\right) dxdy \tag{2.22}$$

where O is a closed curve and D is a plane region bounded by O.

Apart from the valid objective function construction, the uniqueness of the equilibria is also guaranteed with the non-separable symmetric cost function when the jacobian is constructed in such a way that the diagonal dominance is maintained. By doing this, the each block of Hessian is positive definite due to the positive diagonal entries and the diagonally dominant condition. Note that this diagonally dominant condition also applies to the ATEP for ensuring the unique solution (Dafermos 1980).

However, when the cost function is asymmetrically formulated for better modeling of actual transportation systems by addressing the interactions in terms of link, mode and route, (2.6), (2.10), (2,13), respectively, the objective function, eq. (2.15), depends on the path of the integration. Thus, the objective function is not well defined to formulate the UE problem under the optimization framework.

One example of the asymmetric link cost functions that represent flow interactions on the two-way street (Grange and Munoz 2009) shows that the function cannot be properly integrated due to the asymmetric modeling,

$$c_a(v_a, v_{a'}) = t^{free}[1 + 0.15((v_a + 0.5v_{a'}) / (2 \cdot Cp_a))^4] \tag{2.23}$$

where t^{free} and Cp_a are the free-flow link cost and link capacity, respectively.

In order to handle this situation, Smith (1979) and Dafermos (1980) showed that the NEP with the asymmetric cost function can be formulated with a more general form, the VI.

The problem finding F^* in eq. (2.14) is equivalent to finding a solution $F^* \in \Omega$ of the following VI:

$$\begin{aligned} &\langle C(F^*), F - F^* \rangle \geq 0 \quad \forall F \in \Omega \\ &C_p(F) = \sum_{arc\ a \in p} c_a(\Delta F) \end{aligned} \tag{2.24}$$

where Ω is the set of feasible route flow vectors and Δ denotes the link-route incidence matrix

F : vector of route flows f_{rs}^p, $p \in P_{rs}, \forall r \in R, s \in S$

Ω $= \{ f_{rs}^p \geq 0; \sum_{p \in P_{rs}} f_{rs}^p = q_{rs} \ \forall r \in R, s \in S \}$

Δ $= \begin{Bmatrix} 1 & if\ link\ a \in p, \\ 0 & otherwise. \end{Bmatrix}$

The equivalency between the optimization problem and the VI formulation is presented in next section.

As mentioned earlier, the VI formulation contains the fixed point, non-linear complementarity problem and the optimization problem. This section addresses the general equivalency between the VI and the optimization problem by Kinderlehrer and

Stampacchia (1980). Readers are advised to refer Nagurney (2002) for the connection between the VI and the other mathematical formulations.

Proposition 1

Let F_0 be a solution to the optimization problem:

$$\min z(F) \tag{2.25}$$

subject to

$$F \in \Omega \tag{2.26}$$

where z is continuously differentiable and Ω is closed and convex. Then F_0 is a solution of the VI problem:

$$\nabla z(F_0)(F - F_0) \geq 0, \ \forall F \in \Omega \tag{2.27}$$

Proof

let $\phi(t) = z(F_0 + t(F - F_0))$, for $t \in [0,1]$. Since $\phi(t)$ achieves its minimum at $t = 0$, $0 \leq \phi'(0) = \nabla z(F_0)(F - F_0)$, that is F_0 is a solution of the VI.

Proposition 1-1

If $z(F)$ is a convex function and F_0 is a solution to VI, then F_0 is a solution to the optimization problem.

Proof

Since $z(F)$ is convex,

$$z(F) \geq z(F_0) + \nabla z(F_0)(F - F_0), \ \forall F \in \Omega \tag{2.28}$$

But the second part $\nabla z(F_0)(F - F_0) \geq 0$ since F_0 is a solution to the VI. Therefore, one concludes that

$$z(F) \geq z(F_0) \tag{2.29}$$

that is, F_0 is a minimum point of the mathematical programming problem of the optimization problems.

Solution Algorithms for Solving the ATEP

After Prager (1954) first proposed necessities of modeling for the unbalanced interactions between opposite directional flows in a two-way street, various studies investigated the asymmetric modeling under the traffic assignment framework. Before the VI formulation is known to be suitable for modeling the ATEP by Smith (1979) and Dafermos (1980), Netter and Sender (1970) identified the existence of the equilibrium solution in the asymmetrically formulated multi-class network equilibrium problem using the fixed point theorem. However, efforts for modeling the asymmetric interactions were limited to assume the asymmetric interactions to symmetric one and use a simple flow update strategy for the equilibrium process (Dafermos 1971, 1972).

The general formulation of the VI for the ATEP encouraged the researchers to examine the application of the existing solution techniques devised for the conventional Traffic Assignment Problems (TAP) and investigate the convergence conditions, behaviors due to the asymmetric modeling.

The most general solution approach to address the ATEP is to symmetrize the asymmetric interactions and solve the resulting separable sub-problems with conventional solution algorithms used for general traffic assignment problems (Patriksson 1994). Among the general solution algorithms for the TAP, the Jacobi method (known as the diagonalization, relaxation method) would be the most popular one, in which the off-diagonal parts are fixed in the construction of the sub-problem that can

be handled with Frank-Wolfe type solution algorithm. Mahmassani and Mouskos (1988) showed some numerical results on the diagonalization method applied to the case that models the asymmetric interactions between cars and trucks. The study identified the effective number, less than four, of internal Frank-Wolfe iterations for the best performance of the diagonalization method while Sheffi (1985) suggested the one iteration called the streamlined algorithm. A comprehensive research on the convergence trend by the diagonalization method for the ATEP associated with the explicit intersection modeling can be found in Meneguzzer (1995).

Many studies examined the performance of other solution algorithms against the diagonalization methods. These are characterized by cost approximation, especially, the linear approximation methods such as the projection, Newton type algorithms. The basic idea for the iterative linearization method applied to the VI formulation is to define a function $g:\Omega\times\Omega\to R$ which is continuously differentiable and monotone with respect to its first argument and continuous with respect to its second. The linearized approximation function at iteration k may be expressed as:

$$L(F)=\varphi(F,F^k)+(L(F)-\varphi(F,F^k)) \tag{2.30}$$

Thus, the sub-problem solved in iteration k for the VI is

$$[\varphi(F,F^k)+(L(F)-\varphi(F,F^k))]^T(F-F^k)\geq 0 \tag{2.31}$$

Dafermos (1983) defined a smooth function $g(F)$ in such a way that its gradient $\nabla_F g(F,Y)$ is symmetric and positive definite and $g(F,F)=L(F)$.

Then, the eq. (2.31) can be re-expressed as:

$$g(F^{k+1}, F^k)^T (F - F^{k+1}) \geq 0 \tag{2.32}$$

Here, the relaxation algorithm sequentially updates the decomposed problem $g_i(F^{k+1}, F^k)$.

$$g_i(F, Y) = c_i(Y_1, Y_2, ..., F_i, Y_{i+1}, ..., Y_{R \times S}) \tag{2.33}$$

When the sequential update is associated with immediate update of the previous decomposed problem, it is classified the Gauss-Seidel decomposition technique.

On the other hand, the projection method takes a fixed positive definite matrix G.

$$g(F, Y) = c(Y) + \frac{1}{\rho} G(F - Y), \qquad \rho > 0 \tag{2.34}$$

Different choice of the matrix G, an indicator of the level of similarity to the original function, results in different linearized solution algorithms (Pang and Chan 1982).

However, the level of resemblance is not consistently related to the efficient convergence. Fisk and Nguyen (1982) conducted an empirical study which compared the computational performance of the projection, and non-linear Jacobi method for the ATEP with link interactions. They identified that overall, the non-linear jacobi method is efficient for solving the ATEP. Nagurney (1984, 1986) questioned the experiment by Fisk and Nguyen (1982), in which only one network is utilized for the numerical tests. She examined the two algorithms, relaxation and projection methods, on the multimodal equilibrium problem with varied travel cost structures, then, the same experiments were applied to the traffic assignment with the asymmetric link flow interaction problem. The results of the experiments did not clearly declare the best one. In the experiment with the quadratic cost function, the efforts to resemble the original cost function by varying G in the projection method showed a slightly better performance than the relaxation method.

The modified Newton method was applied to the ATEP by Marcotte and Guelat (1988). They compared the performance of the proposed algorithm to the diagonalization, and the cutting plane method on the ATEP associated with different asymmetric complexities. The proposed algorithm outperforms others especially when the asymmetric complexities increase while other algorithms failed to converge.

The column generation method was also utilized for the ATEP. Lawphongpanich and Hearn (1984) applied the simplicial decomposition method where the solution is constructed with the convex combinations of all extreme aggregate flow vectors. Three numerical experiments conducted on the Nguyen-Dupuis network (Nguyen and Dupuis 1984) with different asymmetric complexities demonstrated that the simplicial decomposition is efficient in terms of computer time. The number of minimum tree paths produced by the simplicial decomposition is less than that of the Nguyen-Dupuis' algorithm, consequently the computer time is short. When link flow representation is not appropriate for modeling the route level-interactions, the NCP formulation associated with the column generation method is utilized. Gabriel and Bernstein, Bernstein and Gabriel (1997) solved the ATEP associated with non-additive route cost with the column generation method by minimizing the gap function that re-cast the original NCP. Lo and Chen (2000) also applied the column generation to the reformulated NCP, in which the route-specific toll is modeled.

Various solution approaches for the ATEP is well annotated by Patriksson (1994) where the solution techniques are described with unified manner through the concept of the partial linearization algorithm.

In addition to the algorithmic development for the ATEP, analytical approaches for the convergence condition for the ATEP constitute the other portion of the ATEP related studies. As mentioned earlier, under the asymmetric jacobian for the cost functions, in order for the VI to have the unique solution, a sufficient condition (Dafermos 1980) in which the flow is dominantly affected by the current flows must be satisfied. The main dependence on the current link cost is designed to maintain the monotonicity of the link cost function. However, the convergence of the problem is not always promised because verifying that a link cost is dominantly affected by the current link flows in the real network application is not easy. From the analytical perspective, the asymmetric effects which violate the sufficient condition are analytically investigated and sometimes it is relaxed for modeling the specific asymmetric traffic situations. Heydecker (1983) provided a weaker necessary condition for the convergence of the NEP that accounted un-controlled intersections including the priority junction and roundabouts. Marcotte and Wynter (2004) also presented a weaker convergence condition for the hierarchical nature of the cost interactions observed in the asymmetric mode interactions. They relaxed the monotonicity condition for mode interaction into the strong nested monotonicity condition. Florian and Spiess (1982) analytically derived the condition when the diagonalization method ensures the local convergence. Similarly, Dupuis and Darveau (1986) demonstrated the different convergence conditions between projection and diagonalization methods using degrees of asymmetric effects of the link cost function. Gabriel and Bernstein (1997) who firstly introduced the non-additive route cost function into the TAP showed the existence, and uniqueness condition for the problem by assuming that the function, which translates the time into money, is separable.

In addition, the other approaches for modeling the ATEP include the traffic assignment model combined with the independent simulation on the intersection delay (Wong et al. 2001), modeling the asymmetric interactions with side constraints (Chen et al. 2011a) and a methodology to equalize the different line integral paths for the linearly formulated asymmetric cost function (Grange and Munos 2009). Koutsopoulos and Habbal (1994) observed the improved modeling power of the TAP with different degrees of intersection representation.

In summary, nearly all of solution algorithms for the TAP are applicable to the ATEP because the basic algorithmic approach for the ATEP is to replace the asymmetric interactions with the symmetric one, which results in the symmetric sub-problem.

Even if a few studies report undesirable outcomes due to the ill-constructed the sub-problem (Bernstein and Gabriel 1997; Nguyen and Dupuis 1984), the potential to use the existing algorithms opens the wide range of solution algorithms for the ATEP. However, the asymmetric complexities inherent in the ATEP affect the monotonicity of the cost function, which determines the uniqueness condition of the ATEP. Moreover, normally, the fact that the condition is not easily verifiable for the real transportation network necessitates the judiciously chosen solution algorithm. The basic idea to use the DP method is to employ the solution algorithm that supports the asymmetric degrees until the monotonicity conditions hold. In light of this, the DP method is appropriate for the ATEP because it completely obviates the strong monotonicity assumption to the only monotonicity. The following section deals with these properties of the DP method; why these are advantageous for handling the ATEP. A listing of the literatures is presented in Table 2.1.

Table 2.1 Collection of literatures addressing asymmetric flow interactions.

Authors	Year	Approach	Summary and problem addressed
Prager	1954	Conceptual	Considered modeling for interactions between flows on opposite directions
Netter and Sender	1970	Analytical	Showed the existence of equilibrium solution in the asymmetric multiclass NEP using a fixed point theorem
Dafermos	1971	Analytical	Modeled the traveling cost on a link as a function of the entire flow pattern in the network
Dafermos	1972	Analytical	NEP for multiclass user in transportation networks
Smith	1979	Analytical	Introduced VI to formulate NEP
Dafermos	1980	Analytical	Applied VI and examined the solution properties for asymmetrically formulated NEP
Smith	1981	Analytical	Investigated the properties of a traffic control policy that ensure the solution of the ATEP
Dafermos	1982	Numerical	Applied relaxation algorithm for solving ATEP
Fisk and Nguyen	1982	Numerical	Compared solution algorithms for solving ATEP
Florian and Spiess	1982	Analytical	Derived the convergence condition for the diagonalization for ATEP formulated with link interactions
Pang and Chan	1982	Analytical	A general global convergence with different linearization method is provided
Heydecker	1983	Analytical	Relaxed the convergence condition for ATEP with signalized and un-signalized intersections
Lawphongpanich and Hearn	1984	Numerical	Approached the ATEP with simplicial decompositon method
Nagurney	1986	Numerical	Numerical comparison of algorithms for ATEP
Dupuis and Darveau	1986	Analytical	Investigated the relation between asymmetric complexities and convergence of the diagonalization and projection method
Mahmassani and Mouskos	1988	Numerical	Investigated the solution and convergence of mode interactions between cars and trucks using the diagonalization method
Marcotte and Guelat	1988	Methodological	Applied a modified Newton method and compared the performance to two other solution algorithms for ATEP
Koutsopoulos and Habbal	1994	Numerical	Investigated the different levels of detail of the intersection delay for improving the accuracy of the ATEP
Meneguzzer	1995	Numerical	Comprehensive studies on ATEP associated with intersection delay
Bernstein and Gabriel	1997	Numerical	Showed the different assignment results between additive and non-additive route cost
Gabriel and Bernstein	1997	Analytical	Introduced NEP formulated with the non-additive route cost
Wong et al.	2001	Methodological	Incorporated TRANSYT model into the NEP in order to more precisely account for the intersection delay
Lo and Chen	2000	Analytical	Using a new gap function, a nonlinear complementarity problem is converted to an equivalent unconstrained optimization for traffic equilibrim problem where the general route cost structure is modeled.
Marcotte and Wynter	2004	Analytical	Adapted the hierarchical interactions in two subspace into the NEP with mode interactions
Grange, Munos	2009	Analytical	Used the line integral by assuming the linear asymmetric cost
Chen et al.	2011a	Methodological	Modeled the asymmetric flow interactions using side constraint

Monotone Cost Function by the ATEP

When one of the asymmetric interactions described in the previous sections is modeled in the cost function, the monotone property of a cost function could be observed. In calculus, a function f defined on a subset of the real numbers with real values is called monotonic (also monotonically increasing, increasing or non-decreasing), if for all x and y such that $x \leq y$ one has $f(x) \leq f(y)$. One typical example of the monotone cost function can be depicted as in Figure 2.4 where function values are strictly increasing on the left and right while just non-decreasing in the middle.

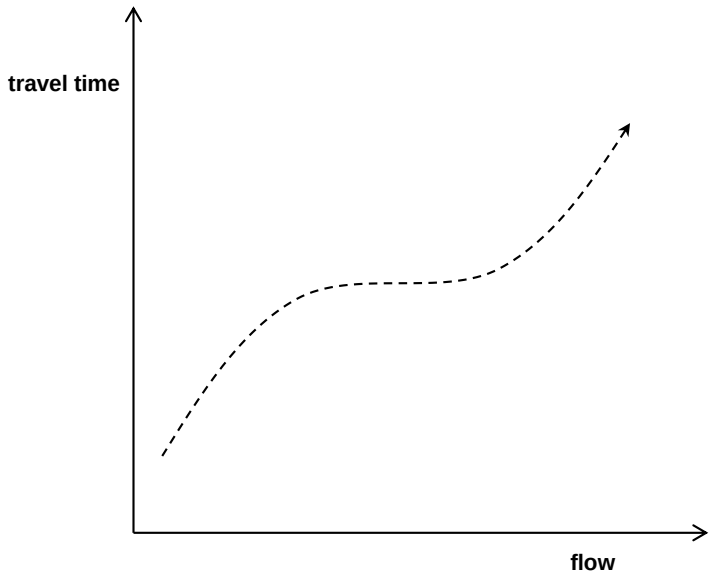


Figure 2.4 An example of a monotone cost function.

Normally, the travel cost increases as the flow on that link adds up. However, for instance, when a signalized intersection, especially an actuated signal system, is modeled in the TAP to improve the reality of the model, the cost function could have the non-

decreasing area because the actuated signal system provides more green time to the heavy flows. In this case, the cost function could be flat or close to be flat because the direction with the heavy flows will have more time to pass the link.

When the cost function could have the monotonicity property due to the asymmetric modeling and consequently the convergence is not verified the DP method is an appropriate solution algorithm because the method is designed to operate on the monotone mapping by performing another projection (Shao et al. 2006).

No Gradient Information

The fact that the DP method does not require the gradient for step-size calculation could be advantageous when the gradient information is not readily available. The characteristic of the DP method that obviates the gradient is useful when the gradient is not readily available due to the asymmetric modeling because the extra procedure is not necessary. As asymmetric modeling specification becomes sophisticated to closely represent the real world, it is often that the gradient is not easily obtainable, thus an extra procedure is integrated into the original program. The typical example is the TAP combined with a separate procedure for the traffic signal control that accounts queue spillback (Qian et al. 2012), coordination of signals (Wong et al. 2001). A few models that simplified the traffic signal can be easily integrated into the TAP whilst a special step such as approximating the gradient (Wong et al. 2001), analytical derivation (Wong 1995) for the gradient specific to the integrated procedure is necessary when a sophisticated separate module is utilized.

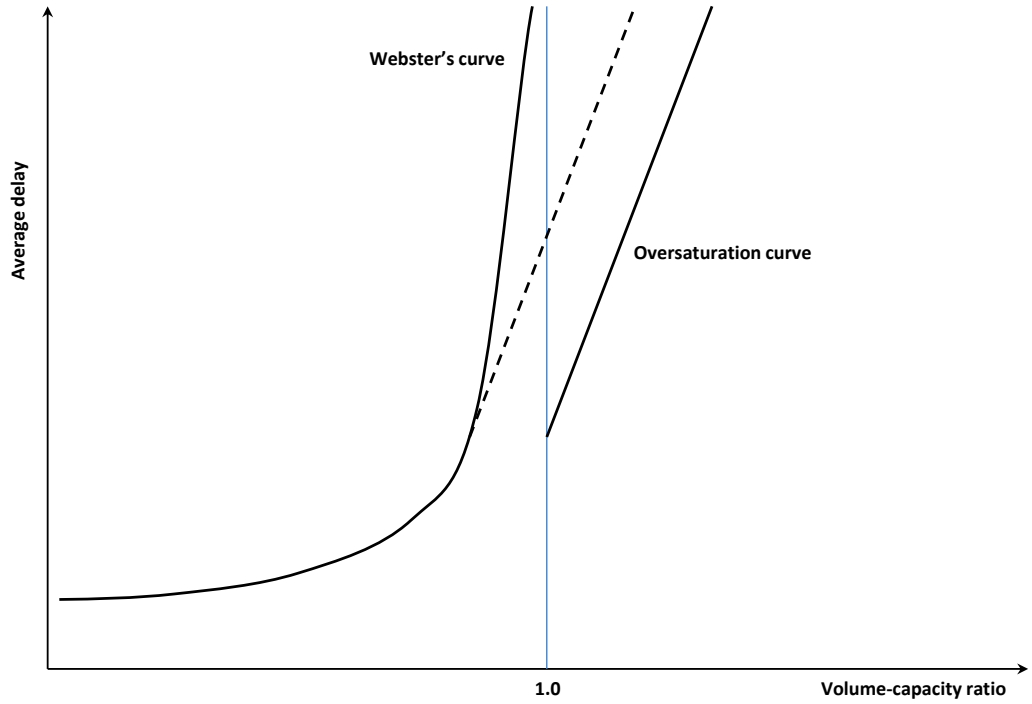


Figure 2.5 The idea of managing the non-differential region for Webster's curve.

A comprehensive reviews of the combined traffic signal control and the TAP can be found in Lee and Machemehl (2005). In addition, the non-differentiable region of the cost function is observed for the Webster's delay model when vehicle-capacity ratio is greater than 1.0. In order to estimate the gradient, the original Webster's curve after the saturation point is replaced with the oversaturation curve at certain point (Van Vuren and Van Vliet 1992).

Route-based Algorithm

Also, as mentioned in the case of the route interactions where the route-based solution algorithm is necessary, the DP method can properly handle the case because the operating space is designed to be at the route level. Before the column generation method is introduced, the solution approach based on the route flow variables were not

advantageous for solving the NEP because all route should be enumerated before the execution of the algorithm (Patriksson (1994) for summary of theses algorithms). Especially, it was not properly handled in terms of computer resources and efficiencies for large networks where the number of routes is much more than the number of links. However, by algorithmically generating the route flow variables (or columns), the marked advantages could be achieved for the route-based solution algorithm (Lo and Chen 2000). Normally, the column generation methods is associated with the column dropping scheme that removes the previously generated columns which are no longer believed to be necessary to express the optimal solutions.

Crowd Modeling in the Transit Assignment

As mentioned in the introduction of the chapter, the crowding effect is modeled in the transit assignment as one of the applications of the ATEP. This section reviews the previous studies that considered the crowding effect in the transit assignment. The review begins with the efforts to value the in-vehicle crowding of a transit vehicle through the statistical model and surveys.

The concerns on the crowding in transit vehicles are slowly gaining the attentions among researcher, consultancies and policy makers after Hensher and McLeod (1977) firstly reported the effect in the public transportation system. The efforts to identify the qualitative measure other than the time and cost associated with the transit travel begin to reasonably value the in-vehicle crowding. The basic idea for quantifying the in-vehicle crowding is based on the trade-off between the onerous crowding and the users' attitude to avoid the crowding. For example, as the crowding level increases in a vehicle, the

passengers are willing to pay more to reduce the crowding in the crowded vehicle than in the un-crowded vehicle. Generally, the valuation of the in-vehicle crowding is modeled as the multinomial logit model and is based the Stated Preference survey (SP). The excellent summary for the valuation of the in-vehicle crowding of the transit travel is presented by Whelan and Crockett (2009), and Li and Hensher (2011).

In the transit assignment modeling, crowding is represented by the in-vehicle congestion. This in-vehicle congestion sometimes in-directly represents the capacity constraint because the additional cost representing the limited capacity of a vehicle induces the users to take another route. In the meantime, when the strict capacity is modeled especially only in the boarding situation in order to prevent the passengers from boarding the fully occupied vehicles, normally, the capacity constraint is not represented by in-vehicle congestion.

The approach to deal with the limited capacity in a vehicle is very similar in both the schedule and frequency-based transit assignment model. The conventional method to model the in-vehicle congestion is to use the discomfort function increasing with the passenger flows. Spiess and Florian (1989) modeled the non-linear cost function in order to represent the dependence of the link cost on the transit passenger volume, namely, "discomfort" term which increases as the vehicles get crowded, in the optimal strategy transit assignment. Wu et al. (1994) formulated the asymmetric transit discomfort function affected by the continuing passengers and the newly boarding passengers in the frequency-based model. De Cea and Fernandez (1993) solved the capacity constraints in the frequency-based model with the "effective capacity" that eliminates the run with the fully loaded vehicle and counts effective frequency. Lam et al. (1999) pointed out that

the discomfort function used in those studies is actually the unbounded increasing function that does not strictly limit the capacity especially in the congested network with bottlenecks. The study explicitly considers the capacity constraints by endogenously determining the passenger overload delay under the stochastic user equilibrium.

A similar approach for modeling the discomfort in the crowded vehicle is found in the schedule-based model. Moller-Pedersen (1999) and Nielsen and Jovicic (1999) used a BPR type formula for calculating the passenger flow dependent transit link cost. Nguyen et al. (2001) inserted an asymmetric boarding penalty function to the link cost to prevent users from taking infeasible paths in terms of available capacities. The study only considered the capacity constraint. Since the available capacity of a stop is a function of the sum of passenger flows from different links going to the stop and the available capacity, the penalty function is constructed with several different flows of incoming links. Poon et al. (2004) modeled the absolute capacity of a vehicle as Lam et al. (1999), but in the schedule-based model. The successful convergence of the strictly constrained capacity was implemented by using the time-dependent network loading procedure (Tong and Richardson 1984). Also, Hamdouch et al. (2004) and Hamdouch and Lawphongpanich (2008) considered the rigid capacity of a vehicle by calculating appropriate probabilities by boarding priority in the time expanded network.

The majority of the studies do not consider the daily variation of the transit demand which can influence congestion at the stop or in the vehicle. Crowding is the direct experience by the transit users as the traffic congestions in the roads. When the level of crowding fluctuates by the daily variation of the transit demand, the variation is also easily recognized by the users. In this regard, it is desirable to consider the random

effect in the transit assignment model. Zhang et al. (2010) considered the both demand and supply uncertainty including the stochastic capacity constraint in the schedule-based model. The Passengers' Arriving, Boarding (PAB) and queuing behaviors are considered in order to model the vehicle's stochastic dwelling at a stop. They identified the different passengers' route choice according to different risk taking attitudes. However, its scope having only four lines focused on the route choice behaviors and the PAB process is modeled at a stop for the capacity constraint. Sumalee et al. (2009) explicitly modeled different discomfort levels of the passenger with and without seats with a stochastic seat allocation model and investigates its consequence on the temporal and spatial distribution of the demand. This chapter also models the random crowding effect under demand uncertainty in the schedule-based transit assignment model. However, it differentiates from Sumalee et al. (2009) in that the study models the different attitudes to the travel time variability. Szeto et al. (2011, 2013) modeled the reliability-based transit assignment that accounts for the variability of in-vehicle congestion and the risk-averse behaviors of passengers with the frequency-based transit assignment model. Later, Fu et al. (2014) also considered the in-vehicle crowding affected by the daily variation of the demand in the reliability-based user equilibrium traffic assignment model for congested multi-modal transport network. However, the frequency-based model formulation for the transit network has lack of capability for modeling the prioritized boarding at the transit stop.

This chapter applies the formulated problem into the large scale real transit network. Considering its realistic nature of the reliability-based assignment, it was not popularly applied to the real network due to the fact that the conventional single criterion

shortest path is not applicable. This study, as far as the study aware, the first one that applies the reliability based shortest path algorithm (Chen et al. 2011b) to the real transit network. The following Table 2.2 lists the literature that dealt with the crowding in the transit assignment modeling.

Table 2.2 A list of literatures that model the crowding effect in transit assignment.

Authors	Year	Modeling approach	Summary and problem addressed
Spiess and Florian	1989	Frequency-based	Non-linear link cost function responding to the passenger flows is modeled for the optimal strategy transit assignment
De Cea, Fernandez	1993	Frequency-based	Used effective capacity by eliminating the fully occupied vehicles
Wu et al.	1994	Frequency-based	Formulates the asymmetric passenger flows in the link discomfort function in the frequency-based transit assignment
Lam et al.	1999	Frequency-based	Explicitly model the hard top capacity of the transit vehicles using the overload passenger delay
Moller-Pedersen	1999	Frequency-based	Flow dependent cost is applied in order to model the capacity restraints
Nielsen and Jovicic	1999	Frequency-based	Applies the transit assignment model that accounts for the crowding cost to the joint route and mode choice assignment
Nguyen et al.	2001	Schedule-based	Asymmetrically models the capacity constraint in the boarding link in order to constraint the capacity
Hamdouch et al.	2004	Schedule-based	The probability of boarding a vehicle is estimated proportional to its residual capacity of a vehicle, which leads to the calculation of the expected cost
Hamdouch and Lawphongpanich	2008		
Poon et al.	2004	Schedule-based	Strictly constrained vehicle capacity modeling in the schedule-based model with time-dependent flow loading procedure
Sumalee et al.	2009	Schedule-based	An explicit and stand-alone seat allocation simulation that differentiates the seated and standing users in a vehicle is integrated into the stochastic transit assignment
Zhang et al.	2010	Schedule-based	A multi-class user reliability-based transit assignment model considers the stochastic features of arriving, boarding and queuing at a stop under the vehicle capacity constraint
Szeto et al.	2011	Frequency-based	The congestion in a vehicle is treated by the additional waiting time
Szeto et al.	2013	Frequency-based	The capacity constraints are developed through the effective capacity and chance constraints in the frequency-based model
Fu et al.	2014	Frequency-based	Crowding cost function is used in order to model the in-vehicle congestion of a transit vehicle in the multi-modal transport network

Fare Models with the Non-Additive Fare Structure

This section reviews the ATEP applications especially for the fare models associated with the non-additive fare structure. Before the review touches the fare models with non-additive fare structure, the following section describes the general fare structures used in practice.

Fare Structures in Use

Fare structures can be classified into two basic categories: flat and differentiated. Under the flat fare, users are charged the same fare, regardless of the travel distance, time of day, or quality of services while the differentiated fare vary according to one or more of those parameters. The followings are the summary of the fare structures used in practice from the TCRP report 10 (1996).

• Flat fare

Charging a flat fare is the simplest, most common fare strategy for transit agencies.

• Distance or zone-based pricing

The basic idea of the distance-based fares (zonal charges or surcharges beyond a certain distance) is to charge the fare for what the users use the services.

• Time-based (e.g., peak/off-peak) differential

Different fares are charged based on the time of day. The logic behind the time-based method of charging is that the peak period market is generally less sensitive to fare increases and the costs of providing service and accommodating high demands are significantly higher in peak than in off-peak hours.

• Service-based (e.g., bus or rail) differential

Fare differentiation by mode is often considered when a higher quality of service provided, the trip distance is relatively long, and the operating cost is higher than other transit modes. Typically, the fare for the rail service is the target of the differentiation because its service characteristics meet the conditions described above.

• Market-based, or consumer-based pricing

Market-based pricing is often seen as a way to price discrimination among different ridership markets (e.g., frequent versus infrequent users) and to reduce cash handling requirements by increasing prepayment.

In practice, each fare structure is used stand alone or combined with other structures. The market-based pricing is often included in the flat fare or other differentiated fare structures.

Fare Models with Non-additive Fare Structure

Among the fare structures used in practice, it is known that the distance-based and zone-based fare structure incurs the non-additivity (Gabriel and Bernstein 1997) like the non-additive toll charging scheme described earlier. The typical example is the distance-based fare structure where the fare is decomposed into the base fare and the additional fare per unit distance (i.e., per mile). Under the fare structure, a fixed unit fare for a unit distance is not available. For example, assume a fare policy that charges a base fare of $3 for 2 miles, in-vehicle travel distance and an $1 additional fee per unit distance. With the fare policy, a passenger traveling 3 miles pays $0.33 per mile while one traveling 4 miles pays $0.25 per mile. The fare is not constantly proportional to the distance a user travels;

therefore, from the modeling perspective this fare structure requires the non-additive route cost model.

The next section reviews the literatures that modeled distance and zone-based fare structures in the fare model. Since the optimal fare levels are dependent on the fare model where these fare structures are integrated, a brief description for the general forms of the fare models will precede the review.

General fare models

The optimal fare levels for fare structures are determined by the fare model where transit agencies' aims such as maximization of demand, welfare, revenue and profits are represented in the objective function. The general formulation of the objective function takes the form of the maximization problem affected by the fare μ and other decision related factors θ such as in-vehicle time, waiting time and number of transfers etc. Thus, it is typical that the fare structure model is integrated in the demand function $d(\cdot)$.

$$\begin{aligned} &\max\ z(\mu,\theta) \\ &s.t. \\ &G(\mu, d(\mu)) \le 0 \end{aligned} \tag{2.35}$$

where $z(\mu,\theta)$ is equal to one of the followings:

$d(\mu,\theta)$; demand maximization

$d(\mu,\theta)\cdot\mu$; revenue maximization

$d(\mu,\theta)\cdot\mu-\varphi(\cdot)$; profit maximization

$d(\mu,\theta)\cdot\mu-\varphi(\cdot)+\Delta\varpi$; welfare maximization

$\varphi, \Delta\varpi$ denotes the operating cost function and difference between willingness to pay by users and actual cost for using transit, respectively.

The following paragraphs describe the general formulation used in the literature for each objective.

Demand maximization

When the fare model's objective is to maximize the number of transported passengers, the demand function itself becomes the objective function (Borndörfer et al. 2012). Also, it is possible to represent the demand maximization as the passenger miles (Nash 1978; Glaister and Collings 1978). Generally, the demand maximization model is constrained by the budget and the vehicle capacity constraints because lowering fare easily maximizes the demand.

Revenue maximization

Revenue is equal to the number of passengers multiplied by the fare the users pay.

$$\max \sum_r \sum_s d_{rs}(\mu_{rs}) \cdot \mu_{rs} \tag{2.36}$$

Since it is known that the elasticity of the transit demand is less than one to the fare change the standalone objective function can have the solution that maximizes the revenue. Instead of using fare μ, it is possible to specify certain passenger interests or political goals (Borndörfer et al. 2012).

Profit maximization

Researchers commonly use profit maximizations to optimize fare levels in the non-additive fare structure. The profit is equal to revenue minus operating expense.

$$\sum_r \sum_s d_{rs}(\mu_{rs}) \cdot \mu_{rs} - \varphi_{rs}(\cdot) \tag{2.37}$$

In order to achieve the maximized profits, the revenue should be maximized while the operating cost minimized. Increasing revenue would decrease the demand and vice versa. Because the elasticity of the transit demand is less than one to the fare change, it is probable that the higher fare levels are the favorite condition for the profit maximization model. Typically, the operating cost is the transit vehicle's frequency function that responds to the demand intensity. Thus, the headway for the vehicle operations is subject to be optimized in the model as well.

Social welfare maximization

The social welfare is the sum of the producer benefit and the user benefit. The producer benefit is the same as the profit, which is revenue minus cost. The user benefit is the difference between the generalized price the users are willing to pay and the actual generalized price the users pay. Borndörfer et al. (2012) derived the expected user benefit for a given fare μ as the difference between the utilities of the best public transport alternative and the best non-public transport alternative. The difference produces the largest generalized utility that a passenger is willing to pay for any public transport alternative before switching to a non-public transport alternative.

$$E[\max\{[\max_{i\in I'} U_i - \max_{i\in I\neq I} U_i], 0\}] = \begin{pmatrix} \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n^2} \cdot e^{n\cdot\frac{\alpha}{\beta}} & \text{if } -\frac{\alpha}{\beta} \geq 0 \\ \frac{\pi^2}{4} + \frac{(\frac{\alpha}{\beta})^2}{2} - \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n^2} \cdot e^{n\cdot\frac{\alpha}{\beta}} & \text{otherwise,} \end{pmatrix} \quad (2.38)$$

It is assumed that the random utility of U_i follows the Gumbel distribution with parameters (η_i, ρ).

where

I : a set of alternatives

U_i : utility for choosing an alternative $i \in I$

$\alpha = \frac{1}{\rho}\ln\sum_{i\in I'} e^{\rho\eta_i} - \frac{1}{\rho}\ln\sum_{i\in I\neq I'} e^{\rho\eta_i}$

$\beta = \frac{1}{\rho}$

The social welfare is to add the producer benefit to the user benefit.

$$\sum_r\sum_s q_{rs}E[\max\{[\max_{a\in M'} U_{rs}^a(\mu_{rs}) - \max_{b\in M\neq M'} U_{rs}^b(\mu_{sr})],0\}] + d_{rs}^m(\mu_{rs})\cdot\mu_{rs} - \varphi_{rs}(\mu_{rs}) \quad (2.39)$$

The former part of the social welfare, which is the user benefit, could be the objective of the fare model.

Additional objectives other than the above mentioned four were also identified. Zhou and Lam (2001) took minimization of the total network travel cost by fixing the O-D demand. Minimization of the revenue decrease for agencies and the fare increase for passengers after the fare structure change into the integrated system was the objective by Pratelli (2004).

Fare models associated with non-additive fare structure

Generally, efforts for determining appropriate fare levels for the non-additive fare structures are conducted under the TRNDP (Transit Network Design Problem) framework. The following paragraphs review the research focused on their fare model and fare structure model specifications.

Daskin et al. (1988) formulated a quadratic model under the revenue maximization for optimizing the distance-based fare that consists of the base, mileage

and transfer charge. The simplified demand function based only on the fares is used for formulating the demand function.

$$d_{rs} = q_{rs}\left[\frac{e_{rs}\mu_{rs} + (1-e_{rs})\mu_{rs}^{0}}{\mu_{rs}^{0}}\right] \tag{2.40}$$

where

d_{rs} : estimated demand for an O-D pair between origin r and destination s

e_{rs} : elasticity of demand with respect to fare

μ_{rs}^{0} : initial fare

$$\mu_{rs} = b + m \cdot s_{rs} + \mu_{x} \cdot xf_{rs} \tag{2.41}$$

where

b : base fare

m : mileage fare

s_{rs} : total in-vehicle distance

$\mu_{x} \cdot xf_{rs}$: transfer fare & number of transfers

The sensitivity analysis suggested that the distance-based fare structure could be effective to raise the revenues because there are revenue potential for the long distance trips which are less elastic to the fare increase. A similar model formulation that reflects the demand decrease to the initial potential demand is observed in Chien and Spasovic (2001), Chien and Tsai (2007), Tsai et al. (2008), and Tsai et al. (2013).

The zone-based fare structure is further differentiated with time t by Chien and Tsai (2007) that formulated an optimization problem for locating the optimal fare and the headway of Newark city subway under profit maximization.

$$\mu_{rs} = b^t + \gamma^t \left| j - i \right| \quad i, j \in \{1, 2, ..., n\} \tag{2.42}$$

where

b^t : base fare at period t

γ^t : additional cost for crossing zones at period t

n : number of zones

$\left| j - i \right|$: number of zones traversed

The optimal fare for zone-based fare structure and the vehicle frequencies are nicely formulated by linearly formulating the elastic demand function. The study found that when the fare and the headway are differentiated by the time and the distance the maximized profit could be obtained.

Tsai et al. (2008) examined the optimal fare levels and the headways for the distance-based fare structure. In the study, the distance-based fare structure is modeled by the weight factor that adjusts the fares for different travel distances.

$$\mu_{rs} = u w_z s_{rs} \qquad \forall r, s; z \in \{short, middle, long\} \tag{2.43}$$

where

u : unit fare

w_z : weight factor which is unique for each travel range

Using the same methodology used by Chien and Tsai (2007), the unit fare per mile with the differentiated weight factor and the headway is optimized for Taiwan's high speed rail network. Tsai et al. (2013) further developed the previous model (Tsai et al. 2008) by tagging the time differentiated distance-based fare structure, which is the

Partition of Travel Distance (PTD). Introducing the variable enables to determine the optimal number of partitions for distance-based fare structure as well as the unit fare and the temporal headways. The study found that in order to maximize the profit, the PTD should be closely segmented as the demand decrease and vice versa. The additional variable necessitated the new solution algorithm for the optimization problem. The study identified the metaheuristic approach, the genetic process, is capable of finding the optimized solution of the given type of problem.

More about the TRNDP is well reviewed by Kepaptsoglou and Karlaftis (2009). Other than the TRNDP framework, Ling (1998) theoretically derived detailed conditions when the distance-based fare is attractive in terms of ridership, revenue, passenger-kilometer and consumer surplus. And Chien and Spasovic (2001) optimized fares with special consideration for the spatial characteristics of the demand in the urban area. A compact review of the literatures are presented in Table 2.3.

Table 2.3 Collection of literatures addressing non-additive fare structure.

Authors	Year	Demand model (maximization of)	Summary and problem addressed
Daskin et al.	1988	Revenue	Establish a quadratic model under the revenue maximization for optimizing the distance-based fare
Ling	1998	Demand, revenue Passenger-km	Theoretically observed more specific conditions upon the differentiated fare structure reform
Chien and Spasovic	2001	Profit and welfare	Optimized fares, route and station spacing and headway with special consideration for the spatial characteristics of the grid network.
Pratelli	2004	Revenue	Design an integrated zone-based transit system with optimal fare levels
Chien and Tsai	2007	Profit	Optimized the fare and the headway for zone-based fare structure
Borndörfer et al.	2008	Profit, revenue, demand, welfare	All models produce higher demand than the fixed fare for the network of Potsdam, Germany upon the distance-based fare structure reform
Tsai et al.	2008	Profit	Examined the optimal fare levels and the headways for the distance-based fare structure
Tsai et al.	2013	Profit	Optimized headways and time differentiated distance-based fare, which is the partition of travel distance

According to the table, we can see that majority of studies are focused on finding the optimal fare levels under profit or revenue maximization. One of the reasons for adopting the distance-based fare structure is to find the revenue potential of the long distance transit demand, which is less sensitive to fare increase. Thus, many studies attempted to find the optimized fare levels under profit maximization. However, profit oriented fare levels may not be desirable for the majority of the transit agencies operated by the government.

In addition, their scopes are not sufficiently large for analyzing travelers' route choice behaviors in the network. Most studies are designed to find the optimized fare levels for small transit routes that do not offer alternatives for the users. Thus, the demand model is about using the transit or not. Even if the several routes are considered the route choice is not modeled according to the fare change (Daskin et al. 1988). Nevertheless, reformulation of the fare structure induces the users' reconstruction of their routes according to the change. Recent fare reformulation in Seoul South Korea is the real example of the route change behaviors. When Seoul reformed its flat fare to the distance-based fare structure in 2004 it experienced 6% of transfer increases than before the reform. When the fare is charged based on the travel distance of the users their decisions on the routes would be inclined to choose the one that minimizes the travel distance. The practical change of the route was triggered by the free transfer policy.

The study touches these two issues in the last chapter of the dissertation.

Summary

Identification of the asymmetric interaction in the transportation system and the appropriate mathematical formulation for these interactions and applications is what this review seeks to embrace.

Various instances of the asymmetric interactions have been reviewed with the basic two-way street link interactions. The basic concepts and the general mathematical formulation of the cost function were also discussed for each case. There were several situations that incur asymmetric interactions. The review classified those interactions by flow source that asymmetrically interact as link, mode, and route. Link interactions were observed when the flows from different links intersect, such as highway ramps, signalized, and un-signalized intersections. Conversely, mode interactions are the flows of different modes having different operational characteristics influencing each other asymmetrically. Route interactions are identified when the route cost cannot be fully represented by the sum of the link cost comprising the route. The route interactions were generally represented by the non-additive route cost.

The next section discussed the necessary changes when the asymmetric effects were considered in the NEP's mathematical formulation. When the asymmetric interactions were modeled as the cost function in the NEP, because integration of the objective function is not independent on the path of the integration, the conventional minimization with an objective function was not appropriate for the ATEP. Instead, the VI properly formulated the ATEP.

Then, the efforts for solving the ATEP had been addressed. The most general solution approach to solve the ATEP is to symmetrize the asymmetric interactions and

solve the resulting separable sub-problems with the conventional solution algorithms used for general NEP. However, effectiveness of other solution algorithms such as the linearization method including the projection, Newton type algorithm and column generation approach also have been researched by various researchers.

In addition to that, the next section describes the advantages of the DP method for handling the ATEP. The usefulness of the DP method for the ATEP is supported by the fact that the operating space of the DP method is the route with relaxed condition on the cost function, and the method doe not requires for the derivative information.

The final section of the literature review pertained to the applications of the observed asymmetric situations in the transit assignment including the crowding effect and the non-additive fare structure. The review summarized the previous studies that consider the crowding effects in the transit assignment. The majority of the studies utilizes the congestion function for representing the crowding effect, however, is limited to the modeling, but lacks the practical application into the real transit networks. The last literature review concerns on the fare model in the transit system. The section went over the general fare structure used in the field and extended its review to the investigation on the fare model that determined the optimized fare levels for the non-additive transit fare structure.

In the next section, the DP method is briefly introduced followed by the characterization of the convergence of each type of the ATEP in terms of the influential factors by the DP method.

CHAPTER 3

ALGORITHM FOR ATEP: DOUBLE PROJECTION METHOD

Double Projection Method

One assumption of the VI formulation is that the mapping is non-expansive, which is a Lipschitz continuous and strongly monotone function (Ortega and Rheinboldt 1970). However, Korpelevich (1977) introduced the DP method that relaxed the strongly monotone mapping of C by performing another projection onto the set of feasible route flow vectors Ω. This study terms the first and the second projection as the predictor projection eq. (3.1) and the corrector projection eq. (3.2), respectively.

$$\bar{F}_k = proj_{\Omega}(F_k - \beta_k C(F_k)) \tag{3.1}$$

$$F_{k+1} = proj_{\Omega}(F_k - \beta_k C(\bar{F}_k)) \tag{3.2}$$

Later, Khobotov (1987) showed that the step-size conditioned with the following inequalities can be easily calculated and the step-size does not require the strongly monotone assumption.

$$0 < \hat{\beta} \le \beta_k \le \min\left\{\bar{\beta}, \varepsilon \frac{\left\|F_k - \bar{F}_k\right\|}{\left\|C(F_k) - C(\bar{F}_k)\right\|}\right\} \tag{3.3}$$

satisfying

$$\left\|F_{k+1} - F^*\right\|^2 \le \left\|F_k - F^*\right\|^2 - \left\|F_k - \bar{F}_k\right\|^2 \left\{1 - \beta_k^2 \frac{\left\|C(F_k) - C(\bar{F}_k)\right\|^2}{\left\|F_k - \bar{F}_k\right\|^2}\right\} \tag{3.4}$$

where $\hat{\beta}, \bar{\beta}$ are the next step-size and the upper limit of the step-size respectively.

The validity of the inequality is proven by either one of the following conditions. First, as $k \to \infty, F_{k+1} \to F^*$. Alternatively, as $k \to \infty, F_k \to \bar{F}_k$. The first condition is apparent as long as the appropriate sequence of step-sizes is selected, while the second one guarantees that the distance to the solution set F strictly decreases. The details of the proof of the convergence of the double projection method are presented in Appendix A.

In order to minimize the right-hand side of eq. (3.4), the following quantity has to be minimized:

$$\left\|F_k - \bar{F}_k\right\|^2 \left\{1 - \beta_k^2 \frac{\left\|C(F_k) - C(\bar{F}_k)\right\|^2}{\left\|F_k - \bar{F}_k\right\|^2}\right\} \tag{3.5}$$

However, there is no closed form for minimizing the quantity. Marcotte (1991) selected the value β_k as $\frac{1}{\sqrt{2}}\left(\frac{\left\|F_k - \bar{F}_k\right\|}{\left\|C(F_k) - C(\bar{F}_k)\right\|}\right)$ while Panicucci et al. (2007) choose $0.8\left(\frac{\left\|f_k - \bar{f}_k\right\|}{\left\|C(f_k) - C(\bar{f}_k)\right\|}\right)$ in their numerical experiments.

Later, Panicucci et al. (2007) applied the step-size rule (Khobotov 1987) to the ATEP associated with the two-way street link interactions with the following algorithmic procedure. The procedure conducts an iterative scheme that examines the validity of the temporary step-size and whether it satisfies condition eq. (3.3) for the suitable scale of the step-size. The sequence for estimation of the proper step-size is shown below:

Step 1. Initialization

set $k = 0, \beta_k = \bar{\beta}, F_k \in \Omega, \varepsilon(0 < \varepsilon < 1), \rho(0 < \rho < 1)$

Step 2. Find the appropriate step-size β_k

Perform the predictor projection $\bar{F}_k = proj_{\Omega}(F_k - \beta_k C(F_k))$

$$\text{if } \beta_k \geq \varepsilon \left[\left\| F_k - \bar{F}_k \right\| / \left\| C(F_k) - C(\bar{F}_k) \right\| \right] \text{ then reduce } \beta_k$$

$$\min \left\{ \rho\beta_k, \varepsilon \frac{\left\| F_k - \bar{F}_k \right\|}{\left\| C(F_k) - C(\bar{F}_k) \right\|} \right\}$$

Go to the top of step 2.

Step 3. Perform the corrector projection

$F_{k+1} = proj_{\Omega}(F_k - \beta_k C(\bar{F}_k))$

Step 4. Convergence test

If convergence criterion is met, then stop. Otherwise $k = k + 1$, go to step 2.

Factors Influencing the Convergence of the ATEP

After Smith (1979) and Dafermos (1980) introduced the VI for the ATEP, attention was given to the solution algorithm's convergence because the sufficient condition for the convergence is not easily verified for the real network. In order for the VI to have the unique solution under asymmetric interactions, a sufficient condition in which the flow is dominantly affected by the current flows must be satisfied (Dafermos

1980). The main dependence on the current link cost is designed to maintain the monotonicity of the link cost function. However, the problem convergence is not always promised because verifying that a link cost is dominantly affected by the current link flows in the real network application is difficult.

Previous studies show the lack of attention to this issue especially the problem with the large real networks and the tested networks used in those studies are small ones constructed for illustration purpose only (Patriksson 1993). Moreover, especially in the analytical approach, research tends to tailor models to one type of asymmetric interactions. This chapter intends to address these issues under the characterization on the convergence of the ATEP by embracing the large real networks and covering different types of interactions. The objective of this chapter is to investigate the factors influencing the convergence of the ATEP using the DP method. The properties of the DP method explained earlier justified the choice of the solution algorithm for this issue.

The basic experiment includes investigation on the effects caused by different asymmetric complexities on the cost function for each type of interactions because the asymmetric complexities are directly reflected on the cost function. In addition, an in-depth analysis will be conducted in order to identify other potential factors that influence the convergence. The experimental framework is designed on the small network and is extended to the large real networks that have different scales and demand intensities. These experiments naturally address the compatibility of the DP method for each type of asymmetric interactions.

The small network used for the following experiment was devised by Nguyen and Dupuis (1984) for testing a solution method for the NEP formulated with asymmetric link

cost. The shape of the network and the O-D demand is presented in Figure 3.1. The network consists of thirteen nodes and nineteen links. Additional details on the network characteristics can be found in Nguyen and Dupuis (1984).

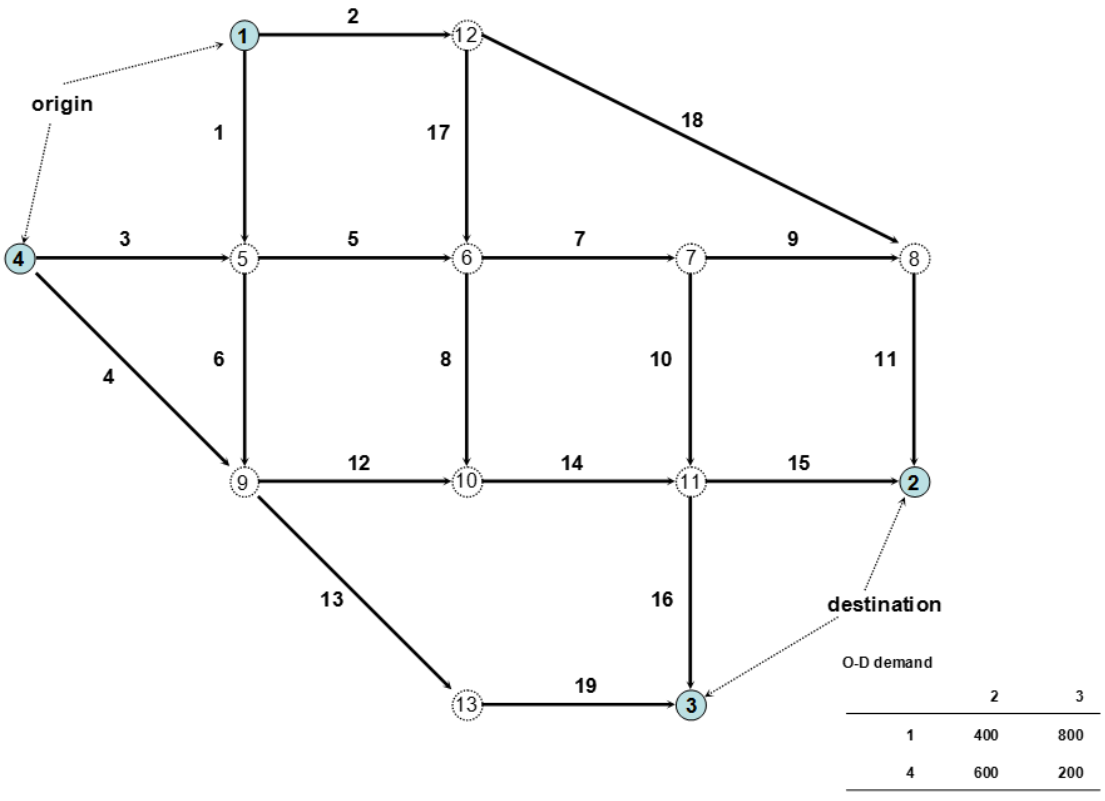


Figure 3.1Nguyen-Dupuis network.

The cost matrix *A* and the constant vector *h* associated with the N-D network are given in Appendix B. The selected networks for real network experiments and their characteristics are presented on Table 3.1. The first three are relatively small to medium sized networks with different levels of demand intensities. The larger scale network (Chicago Regional Network) experiment has a special data structure developed for the efficient use of the memory.

Table 3.1 Network characteristics of the test networks for real network experiments.

Network	#. nodes	#. links	#. O-D pairs	Total demand
Anaheim	416 (38)	914	1,406	104,694
Winnipeg	1,067 (135)	2,535	4,345	54,459
Barcelona	930 (97)	2,522	7,922	184,679
Chicago Regional	12,982 (1,790)	39,018	1,118,661	1,323,537

The number in parenthesis indicates the number of zones. The demand, which is greater than or equal to 0.1, is used for the Chicago Regional Network. It explains more than 97% of the original demand while the objective to apply the algorithm to the large scale network is not disturbed.

The numerical tests will be terminated by the Average Excess Cost (AEC) rule introduced by Boyce et al. (2004).

$$\max_{r.s} \sum_{p \in P^{rs}} \frac{f_p^{rs}}{q^{rs}} \left(\frac{C_p(f) - C_{\min}^{rs}(f)}{C_p(f)} \right) \leq \varepsilon \tag{3.6}$$

where $C_{\min}^{rs}(f)$ is the shortest route cost of an O-D pair between origin r and destination s. This study used $\varepsilon = 10^{-5}$. The computer specifications running the program coded with the C# language are Intel Core I5 680 processor with 3.59 GHz speed, and 4.00GB of RAM.

Link Interactions

In order to model the link interactions with different asymmetric complexities, the following link cost function that decomposes cost matrix A into the diagonal and the off-diagonal parts tagged with the asymmetric controller ρ_{lk} is utilized.

$$C(V) = [A_{diag} + \rho_{lk} A_{offdiag}]V10^{-3} + h \tag{3.7}$$

where A_{diag} and $A_{offdiag}$ are the diagonal and off-diagonal parts of the asymmetric link cost matrix of the N-D network.

In this experiment, the first order Taylor expansion at V_0 of $C(V)$ is approximated so that the relative ratio between the asymmetric and the symmetric effects can be estimated. This measure is advantageous over the asymmetric controller because the measure indicates the relative proportion of the asymmetric effect to the symmetric effect by breaking down the evaluation at V_0 into the symmetric and skew-symmetric parts.

$$C(V) \approx C(V_0) + C'(V_0)(V - V_0) \tag{3.8}$$

Here, decompose the derivative $C'(V_0)$ into its symmetric and skew-symmetric parts as:

$$C(V) \approx C(V_0) - C'(V_0)V_0 + X(V_0)\Delta V + Y(V_0)\Delta V \tag{3.9}$$

where

$X = (1/2)\{C'(V_0) + C'(V_0)^t\}$, symmetric part

$Y = (1/2)\{C'(V_0) - C'(V_0)^t\}$, skew-symmetric part

Then, the influence of the skew-symmetric part to the symmetric part is as:

$$R = \frac{V_0^t |Y| V_0}{V_0^t |X| V_0} \tag{3.10}$$

It is worth noting that the current flows V_0 enter into the definition of R and influence the degrees of effects caused by the skew-symmetric part. The cost function arranges the asymmetric configurations while the network flows, arguments to the cost function, determines the actual asymmetric effects. The irregularity of the network flows is not accounted in the analytical approach that assumes matrix A to have consistent

effects throughout the solution procedure. However, the net effects of the matrix depend on the flows in the network. Simply, if a link flow is zero, then the influence caused by the zero flow link is canceled and overall asymmetric effect will diminish than the originally intended.

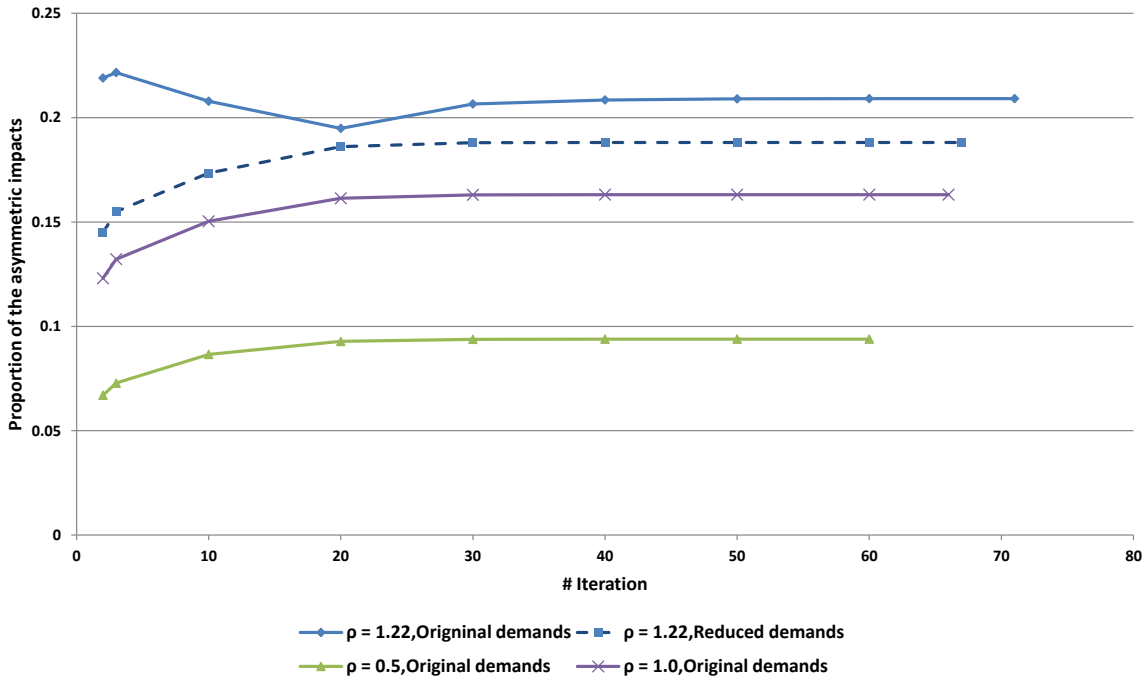


Figure 3.2 Actual asymmetric effects throughout the convergence.

In order to demonstrate this, different demands have been applied to the same asymmetric complexities, $\rho_{lk} = 1.22$, and observed R. Figure 3.2 presents the results. The top two lines share the same asymmetric effects but different demand levels while other lines are with the original demand. We can see that the asymmetric complexities are orderly configured according to the scale of ρ_{lk}. The figure demonstrates that when the demand level is low the practical asymmetric complexities reflected in the solution procedure is less than the one with higher demand intensity. The fact that $\rho_{lk} = 1.22$ is the

maximum asymmetric complexities that can be covered by the diagonalization method due to Florian and Spiess (1982), the actual asymmetric complexities could be higher or lower than the value. In the real networks, it is possible that the theoretical configurations would not work as the effective upper limit because the real network is associated with varied demand intensities.

In this section, a cost function that models the different asymmetric link interactions has been established and used for the experiment designed to analyze the actual asymmetric effects with the given function. The experiment results identified that the asymmetric effects of the ATEP are not solely the function of the cost function but are the results of the combined effects between the cost function and the network flows.

The section examined the realization of the asymmetric effects due to the variability of the network flows. It may possible that the same analogy can be applied to the cost function in such a way that the intended asymmetric cost functions are ideally constructed. This will be dealt in the following real network experiments.

For the real network experiment, the asymmetric link interactions have been modeled with two way street interactions by the following link cost function.

$$c_a = t_a^{free}\{1+\alpha([v_a + \rho_{lk} \cdot v_a']/2Cp_a)^{\beta}\} \tag{3.11}$$

Assuming that the demand intensities affect the effective asymmetric complexities, ρ_{lk} will be different for different networks because the demand intensities are vary in the real networks. This is demonstrated by identifying ρ_{lk} that breaks the diagonal dominance for different networks. In order to test this, the study observes the solution while increasing ρ_{lk} until the network outputs different solutions due to the violation of the uniqueness condition.

The uniqueness of a solution is verified by the following test of similarity of link flow pattern proposed by Meneguzzer (1995).

$F_l^{B,i}$ = initial all-or-nothing flow on link l in the base solution

$F_l^{k,i}$ = initial all-or-nothing flow on link l in the k_{th} alternative solution,

$k = 1, 2, 3, ..., 10$

$F_l^{B,E}$ = equilibrium flow on link l in the base solution

$F_l^{k,E}$ = equilibrium flow on link l in the k_{th} alternative solution

$k = 1, 2, 3, ..., 10$

For each link l and for each alternative solution k , calculate the degree of the deviate from the base and equilibrium link flow solution

$$\left|F_l^{B,i} - F_l^{k,i}\right| \le 0.1F_l^{B,i} \tag{3.12}$$

$$\left|F_l^{B,E} - F_l^{k,E}\right| \le 0.1F_l^{B,E} \tag{3.13}$$

Next compute the

$$P_k^i = \frac{N_k^i}{N} \tag{3.14}$$

$$P_k^E = \frac{N_k^E}{N} \tag{3.15}$$

where

N_k^i = total number of links that satisfy eq. (3.12)

N_k^E = total number of links that satisfy eq. (3.13)

N = total number of links in the network

The ratio in eq. (3.14) and eq. (3.15) are called the Proportion of Similar Initial Link Flows (PSILF) and the Proportion of Similar Equilibrium Link Flows (PSELF),

respectively. The idea of the method is to examine if the different initial solutions will have the same equilibrium solution. The degree of difference is measured by PSELF and PSILF. The successful test would yield a small PSILF and a large PSELF. In addition, the study added one more measure that calculates the average amount of flows that obviate from eq. (3.15).

$$ADLF = \frac{\sum_{l} \left| F_l^{B,E} - F_l^{k,E} \right| > 0.1 F_l^{B,E}}{(N - N_s^E)} \tag{3.16}$$

Table 3.2 shows the values of each variable for the tested networks introduced above. The reason for presenting the assigned results when $\rho_{lk} = 0.0$ is to provide a reference point for comparing the other results comes from different asymmetric complexities.

At the reference point, for three small and medium sized networks, the similarity of the initial flow pattern to the base solution is kept low while almost the unique UE solution (99% of PSELF) is obtained from different initial points. In the case of Chicago regional network, 83% of PSELF is achieved. This value is not sufficiently strong to indicate that an identical link flow solutions are obtained but considering its enormously huge network size and the low value of ADLF suggest that the value is acceptable.

At a glance on Table 3.2, the high value of PSELF, meaning that it is very close to the reference case, is maintained even for the high asymmetric degrees. Theoretically, when ρ_{lk} is equal to 0.5 or higher the diagonal dominance of the current link's flow would not be expected. However, all networks yield almost the same levels of PSELF to the separable case even with $\rho_{lk} = 0.8$.

Table 3.2 PSILF and PSELF with different asymmetric complexities.

Anaheim					Barcelona			
solutions	#. iterations	PSILF	PSELF	ADLF	#. iterations	PSILF	PSELF	ADLF
$\rho_{lk} = 0.0$								
1	86	0.18	0.99	83.96	262	0.05	1.00	23.15
2	66	0.19	0.98	142.21	188	0.05	1.00	15.10
3	85	0.17	0.99	69.58	204	0.03	1.00	18.55
4	75	0.20	0.99	47.20	151	0.05	0.99	14.16
5	79	0.19	1.00	26.34	217	0.04	1.00	75.70
6	99	0.18	1.00	0.00	183	0.04	1.00	26.98
7	71	0.17	0.99	155.37	178	0.04	1.00	15.58
8	67	0.18	0.99	111.71	308	0.04	1.00	49.53
9	56	0.20	0.98	114.10	208	0.06	1.00	21.96
10	73	0.15	1.00	37.32	241	0.05	1.00	23.99
average	76.0	0.180	0.991	78.778	214.0	0.045	0.997	28.472
$\rho_{lk} = 0.5$								
1	62	0.16	0.98	117.79	493	0.05	1.00	61.80
2	61	0.15	0.99	32.19	568	0.05	1.00	52.24
3	59	0.17	0.99	46.79	455	0.03	1.00	72.52
4	54	0.18	0.98	45.84	462	0.05	1.00	62.87
5	64	0.19	0.99	54.33	555	0.04	1.00	85.61
6	72	0.16	0.99	40.76	474	0.07	1.00	57.65
7	64	0.17	0.99	32.79	369	0.05	1.00	49.56
8	57	0.15	0.99	38.90	453	0.05	1.00	53.52
9	60	0.17	0.98	90.06	513	0.05	1.00	89.98
10	65	0.18	0.99	62.36	524	0.04	1.00	88.51
average	61.8	0.168	0.987	56.181	486.6	0.048	0.998	67.425
$\rho_{lk} = 0.8$								
1	87	0.17	1.00	24.91	514	0.06	0.99	62.34
2	55	0.17	0.99	69.95	563	0.07	1.00	82.60
3	93	0.19	0.99	37.39	589	0.06	0.99	72.76
4	76	0.15	1.00	43.40	833	0.05	0.99	70.14
5	70	0.20	0.97	102.88	864	0.06	1.00	72.09
6	73	0.17	0.99	115.44	961	0.06	1.00	81.04
7	79	0.19	0.99	94.11	715	0.05	1.00	78.89
8	98	0.19	0.98	75.74	601	0.03	1.00	84.18
9	65	0.19	0.98	122.59	620	0.05	1.00	82.79
10	90	0.19	0.99	65.41	863	0.04	1.00	60.95
average	78.6	0.180	0.989	75.181	712.3	0.051	0.995	74.778
$\rho_{lk} = 1.0$								
1	115	0.15	0.96	140.2	583	0.05	0.99	23.4
2	71	0.17	0.96	118.2	779	0.05	0.98	110.2
3	100	0.16	0.97	96.1	613	0.06	0.99	26.4
4	78	0.15	0.96	97.6	688	0.05	0.98	73.0
5	72	0.17	0.97	100.8	690	0.04	0.98	97.4
6	63	0.17	0.97	0.0	707	0.04	0.99	26.1
7	110	0.18	0.98	53.0	688	0.05	0.99	28.8
8	70	0.18	0.97	125.0	951	0.05	0.99	26.8
9	93	0.20	0.97	107.7	771	0.06	0.99	75.2
10	71	0.17	0.96	157.3	640	0.05	0.98	69.2
average	84.3	0.170	0.965	99.585	711	0.050	0.986	55.660

Table 3.2 continues

Winnipeg					Chicago regional			
Solutions	# Iterations	PSILF	PSELF	ADLF	# Iterations	PSILF	PSELF	ADLF
$\rho_{lk} = 0.0$								
1	992	0.14	0.99	11.71	621	0.18	0.83	69.60
2	1211	0.17	0.99	10.10	877	0.18	0.83	69.33
3	794	0.14	0.99	9.45	648	0.19	0.83	69.89
4	868	0.15	0.99	10.35	583	0.18	0.83	71.07
5	1324	0.15	0.99	11.17	672	0.18	0.82	68.83
6	794	0.16	0.99	8.47	529	0.18	0.82	68.23
7	442	0.16	0.99	8.62	669	0.18	0.83	69.61
8	786	0.16	0.99	8.08	543	0.18	0.83	66.10
9	407	0.14	0.99	8.60	582	0.18	0.83	65.07
10	582	0.15	0.99	10.83	544	0.18	0.83	72.13
average	820	0.153	0.991	9.739	626.8	0.183	0.826	68.986
$\rho_{lk} = 0.5$								
1	554	0.18	1.00	19.57	819	0.18	0.83	59.24
2	587	0.16	1.00	8.04	909	0.18	0.83	59.93
3	548	0.16	1.00	11.84	1054	0.18	0.82	59.01
4	554	0.15	1.00	16.81	877	0.18	0.82	57.47
5	565	0.17	1.00	8.07	807	0.18	0.83	57.54
6	525	0.17	1.00	10.14	1897	0.18	0.83	56.80
7	489	0.14	1.00	8.61	1144	0.18	0.82	56.76
8	599	0.17	1.00	16.51	1149	0.18	0.82	57.30
9	548	0.14	1.00	14.20	805	0.18	0.83	57.68
10	500	0.15	1.00	19.24	3525	0.18	0.82	57.28
average	546.9	0.159	0.997	13.303	1298.6	0.182	0.826	57.901
$\rho_{lk} = 0.8$								
1	590	0.16	1.00	10.48	1196	0.18	0.81	58.44
2	805	0.16	1.00	7.54	1706	0.18	0.82	57.13
3	644	0.15	1.00	3.14	1210	0.18	0.82	57.80
4	497	0.16	1.00	7.02	1350	0.18	0.81	58.45
5	730	0.14	1.00	4.84	1218	0.18	0.81	57.62
6	692	0.16	1.00	9.83	887	0.18	0.82	58.98
7	931	0.14	1.00	7.14	1107	0.18	0.82	56.82
8	872	0.14	1.00	12.03	991	0.18	0.81	56.31
9	620	0.15	1.00	2.82	802	0.18	0.82	58.54
10	651	0.15	1.00	9.70	3525	0.18	0.82	57.28
average	703.2	0.151	0.998	7.454	1399.2	0.180	0.816	57.736
$\rho_{lk} = 1.0$								
1	563	0.16	0.98	50.09	1493	0.18	0.78	71.09
2	692	0.16	0.97	32.25	1408	0.19	0.79	71.05
3	717	0.20	0.97	23.96	1029	0.18	0.79	70.57
4	615	0.17	0.97	52.58	1986	0.18	0.78	71.44
5	722	0.18	0.98	17.68	1026	0.18	0.78	70.83
6	642	0.17	0.97	31.51	1600	0.18	0.78	70.81
7	742	0.17	0.96	45.58	1279	0.18	0.78	69.42
8	769	0.18	0.97	26.30	2248	0.18	0.79	69.56
9	727	0.19	0.97	18.94	972	0.18	0.78	71.81
10	789	0.18	0.98	27.58	1379	0.18	0.79	70.88
average	697.8	0.175	0.971	32.647	1442.0	0.181	0.784	70.746

Seemingly, the violation of the unique solution is observed when $\rho_{lk} = 1.0$ because a significant difference on PSELF appears. However, in the case of the Barcelona network, the diagonal dominance seems to be maintained until ρ_{lk} is very close to one.

The study investigates the factors that contribute to the relatively high value of the asymmetric complexities incurring the unique solution. For the two-way street cost function, if the both directional links share the same congestion parameters, then the problem becomes the symmetric case by arranging the derivative of the current directional flow next to the opposite one for constructing the jacobian. When both the congestion functions are the same, eq. (3.7) models the symmetric effects. Table 3.3 presents the proportion of the symmetric two-way street pair in the test networks.

Table 3.3 Percentage of symmetric two-way street pairs in the real networks.

Networks	number of two-way street pairs with the same performance function	number of two-way street pairs	% of symmetric pairs
Anaheim	271	280	97
Barcelona	963	1151	84
Winnipeg	661	724	91
Chicago regional	8776	18391	48

Very high percentage of symmetric pairs is observed except for the Chicago regional network. The high percentage of symmetric pairs implies that the asymmetric effects designed by the cost function will not be realized as originally planned due to the incomplete asymmetric configuration in the real networks. Similar to the small network experiment that identified the dependency of the practical asymmetric effects on the demand intensities, the real network experiment identified incomplete asymmetric effects due to the network configuration.

Through the experiments, it has been identified that possibility exists in that the actual asymmetric effects are lower than originally planned because the symmetric configurations might be hidden in the cost function due to the network configuration and because of the demand intensities. This might be the supporting evidence for the suggestion by Patriksson (1993) in which the theoretical convergence conditions are too strong and that it may be possible to weaken them significantly.

Mode Interactions

In order to model asymmetric mode interactions, the link interactions formulated by the off-diagonal part of A is eliminated by taking only the diagonal part of A. Limiting asymmetric effects to one link is typical for mode interactions, since interactions between modes are designed to model flow reactions between modes sharing the same physical infrastructure (Mahmassani and Mouskos 1988). The link cost function is expressed simply as:

$$c_a^m(v_a^1, v_a^2, ..., v_a^M) = [A_{diag}]s_a^m(v_a^1, v_a^2, ..., v_a^M)10^{-3} + h \tag{3.17}$$

where $s_s^m(\cdot)$ is a function that models the interaction among modes in link a.

Even if the cost function is symmetric with respect to the link flows, asymmetric effects among modes could be modeled by function $s(\cdot)$, which represents the interactions among modes. In this experiment, flow interactions between two modes, cars and trucks, are modeled using a simple form as:

$$\begin{aligned} s_a^{car} &= v_a^{car} + \varepsilon v_a^{truck} \\ s_a^{truck} &= \rho_{md} \cdot v_a^{car} + \varepsilon v_a^{truck} \end{aligned} \tag{3.18}$$

where ε and ρ_{md} are the factors that transform the effect of trucks into passenger car equivalent (PCE) unit and the asymmetric measure which controls the impact of the

auxiliary mode. Here, ρ_{md} plays the same role in the link interactions for posing the asymmetric effects between modes.

With the defined cost function for each mode, an experiment is prepared to analyze the situations where the interaction between modes gradually increases. Both cost functions are influenced by flows of the cars and trucks; however, the program has a unique solution as long as ρ_{md} is equal to zero. This is because the strong nested monotonicity, a weaker form of monotonicity, is maintained over two cost functions of cars and trucks. Cohen and Chaplais (1988) introduced nested strong monotonicity providing a weaker form of the monotonicity condition and the process of interactions for the Multiclass Network Equilibrium Problem (MNEP).

The consequences of violation of the monotonicity condition controlled in different cost functions directly touch the uniqueness issue. When ρ_{md} is zero, strong nested monotonicity is easily confirmed since the NEP is operated with the strongly monotone function for C^{car} with fixed trucks' flows while the cost function for the truck with the parametric solution obtained before is strongly monotone due to $\rho_{md} = 0$. This is demonstrated in the numerical experiment that produces the same link flows on the individual class with several different initial solutions when $\rho_{md} = 0$ (Table 3.4). On the other hand, the uniqueness of the MNEP is no longer valid when $\rho_{md} = 1$ and higher. When ρ_{md} becomes one, it seems that the balance of power between two cost functions is tipped and consequently the condition for the unique solution is violated.

Table 3.4 Equilibrium solutions with the violated monotonicity conditions.

ρ_{md}	0			1					
Link number	Link flows from different initial solutions			Link flows from initial solution 1			Link flows from initial solution 2		
	Cars	Trucks	sum	Cars	Trucks	sum	Cars	Trucks	sum
1	470	176	646	600	132	732	382	200	582
2	430	124	554	300	168	468	518	100	618
3	216	82	298	0	150	150	450	7	457
4	384	118	502	600	50	650	150	193	343
5	216	104	320	0	168	168	526	7	534
6	470	153	623	600	114	714	305	200	505
7	346	128	474	0	236	236	745	7	752
8	0	0	0	0	0	0	0	0	0
9	216	82	298	0	150	150	450	7	457
10	130	47	177	0	86	86	295	0	295
11	516	182	698	300	250	550	750	107	857
12	234	68	302	450	0	450	0	143	143
13	620	203	823	750	164	914	455	250	705
14	234	68	302	450	0	450	0	143	143
15	234	68	302	450	0	450	0	143	143
16	130	47	177	0	86	86	295	0	295
17	130	24	154	0	68	68	218	0	218
18	300	100	400	300	100	400	300	100	400
19	620	203	823	750	164	914	455	250	705

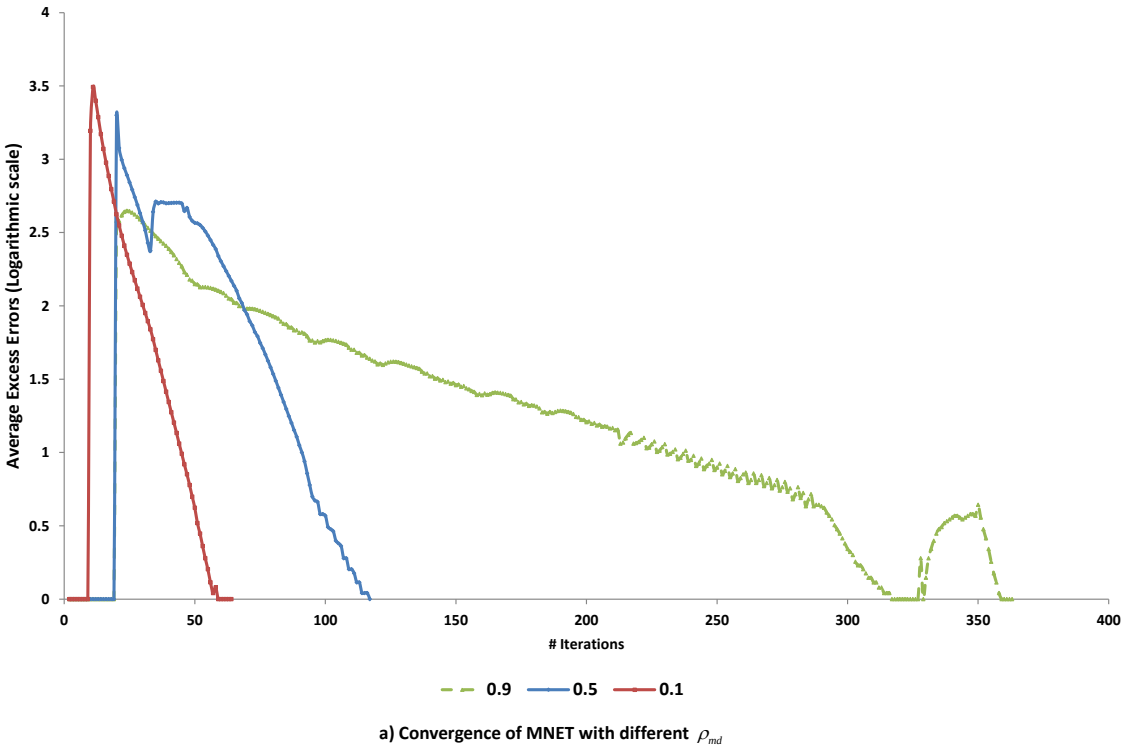
In Table 3.4, the link flow solutions are presented when ρ_{md} is equal to zero and one. As explained above, different equilibrium link flows were identified when $\rho_{md} = 1$ due to the violation. Conversely, the unique equilibrium solutions are identified when ρ_{md} is lower than one by the method (Table 3.5). The solutions between the presented asymmetric complexities can be found in Appendix C.

Meanwhile, convergence behaviors represented by the AEC for each O-D pair have been compared when asymmetric influences grow on the truck's cost function (Figure 3.3 a).

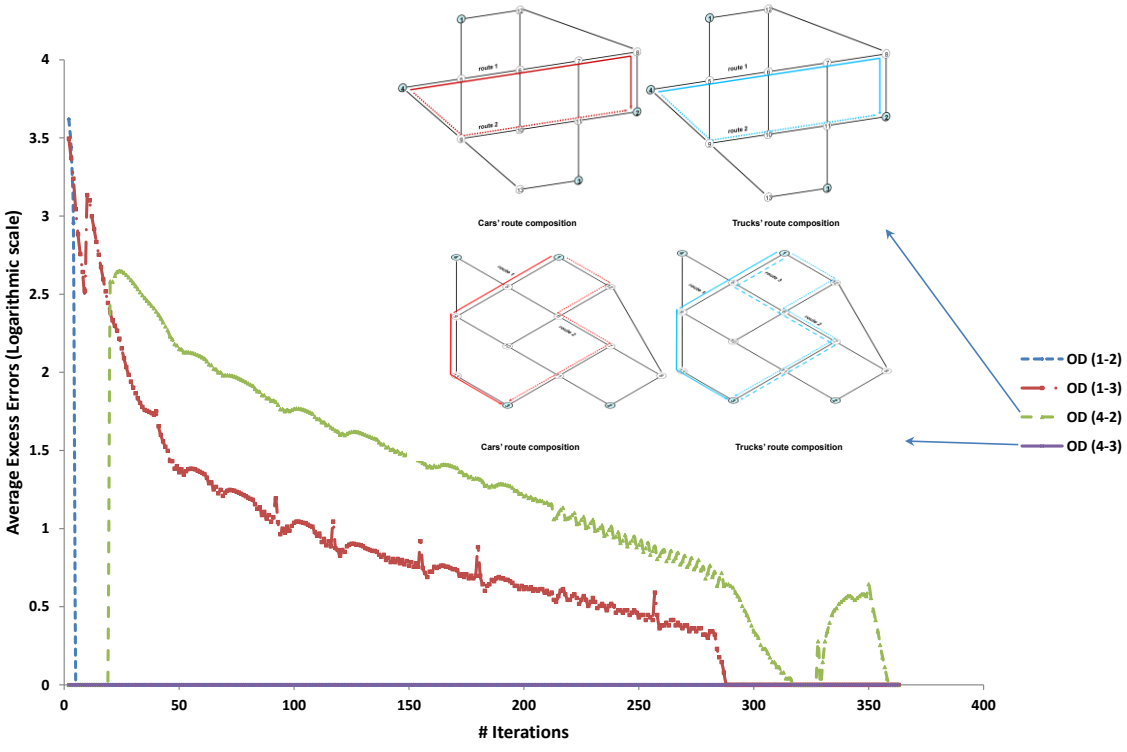
Table 3.5 Equilibrium solutions caused by the asymmetric mode interactions.

direction	Link sequence	Route costs and route flows							
		$\rho_{md} = 0.0$				$\rho_{md} = 0.9$			
		Cars		Trucks		Cars		Trucks	
Origin 1 to destination 2	2 18 11	70.00	300.00	53.87	100.00	70.13	300.00	68.34	100.00
	2 17 7 9 11	86.87	0.00	64.84	0.00	87.12	0.00	84.68	0.00
	2 17 8 14 15	91.78	0.00	70.07	0.00	91.97	0.00	89.55	0.00
	1 5 7 9 11	87.40	0.00	64.84	0.00	87.18	0.00	84.68	0.00
	1 6 12 14 15	109.77	0.00	80.69	0.00	109.63	0.00	106.40	0.00
	1 5 8 14 15	92.31	0.00	70.07	0.00	92.02	0.00	89.55	0.00
	2 17 7 10 15	88.94	0.00	67.22	0.00	89.04	0.00	86.63	0.00
	1 5 7 10 15	89.47	0.00	67.22	0.00	89.10	0.00	86.63	0.00
Origin 1 to destination 3	2 17 8 14 16	95.96	0.00	72.59	0.00	96.01	0.00	93.41	0.00
	2 17 7 10 16	93.11	130.34	69.74	24.55	93.09	130.39	90.50	26.34
	1 5 7 10 16	93.64	0.00	69.74	21.91	93.15	0.00	90.50	19.27
	1 5 8 14 16	96.48	0.00	72.59	0.00	96.07	0.00	93.41	0.00
	1 6 12 14 16	113.94	0.00	83.21	0.00	113.68	0.00	110.26	0.00
	1 6 13 19	93.11	469.66	69.74	153.55	93.09	469.61	90.50	154.40
Origin 4 to destination 2	3 5 7 9 11	77.41	215.94	59.55	81.55	77.32	214.89	75.35	82.16
	3 5 7 10 15	79.48	0.00	61.93	0.00	79.24	0.00	77.30	0.00
	3 5 8 14 15	82.32	0.00	64.78	0.00	82.17	0.00	80.21	0.00
	3 6 12 14 15	99.78	0.00	75.40	0.00	99.78	0.00	97.06	0.00
	4 12 14 15	77.42	234.06	59.55	68.45	77.34	235.11	75.35	67.84
Origin 4 to destination 3	3 5 7 10 16	83.65	0.00	64.45	0.00	83.29	0.00	81.16	0.00
	3 5 8 14 16	86.50	0.00	67.30	0.00	86.21	0.00	84.08	0.00
	3 6 12 14 16	103.96	0.00	77.92	0.00	103.83	0.00	100.93	0.00
	3 6 13 19	83.13	0.00	64.45	0.00	83.24	0.00	81.16	0.00
	4 12 14 16	81.60	0.00	62.07	0.00	81.39	0.00	79.21	0.00
	4 13 19	60.77	150.00	48.60	50.00	60.80	150.00	59.44	50.00

In Figure 3.3, the AEC have been computed for each O-D pair level and the routes used for each mode are presented in order for the in-depth analysis for the convergence of the MNET (Figure 3.3 b)). The figure shows that as the influences of interactions between modes increase the more efforts are required for the convergence and the prolonged convergence is caused by the O-D pairs that overlap the routes with another mode's routes. This implies that the convergence of the MNET is dependent on controlled asymmetric interactions and the route composition between modes.



a) Convergence of MNET with different ρ_{md}



b) Convergence of each O-D pair when $\rho_{md} = 0.9$

Figure 3.3 Convergences of mode interactions with different asymmetric complexities.

The figure also shows that the O-D pair that fully shares the route composition determines the convergence of the problem because the sheer asymmetric impacts intended by ρ_{md} are realized in the O-D pair.

In the case of the mode interactions, the N-D network experiment identified that the convergence of the mode interactions is affected by the asymmetric complexities among modes and the route composition. Especially its impact is significant when each mode fully shares their routes. The following paragraphs address the problem on the real networks.

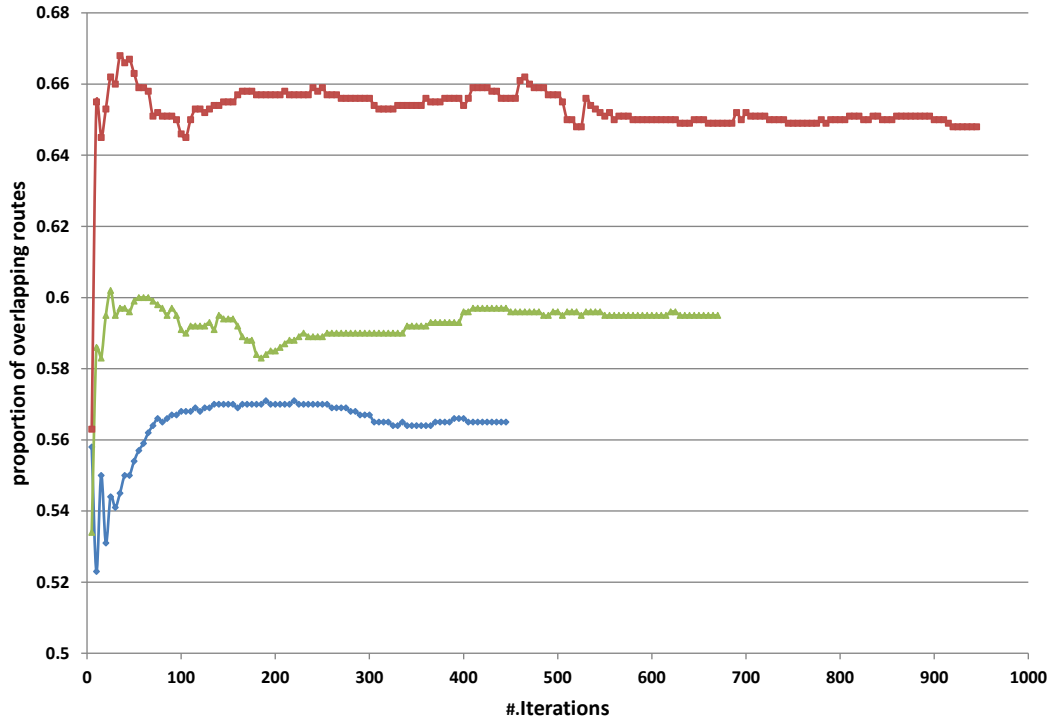
We have observed that the convergence of the MNEP is determined by the route composition of each vehicle type in the small network experiment. The flows assigned on the overlapped routes need more efforts to be stabilized for the UE condition because they are influencing each other on the same route. Also the convergence was affected by the asymmetric complexities. This relation is sought in the real networks too. In the real network implementation, the convergence behaviors and the route composition are analyzed with different asymmetric complexities.

Precisely as the small example does, asymmetric mode interactions are modeled with two modes.

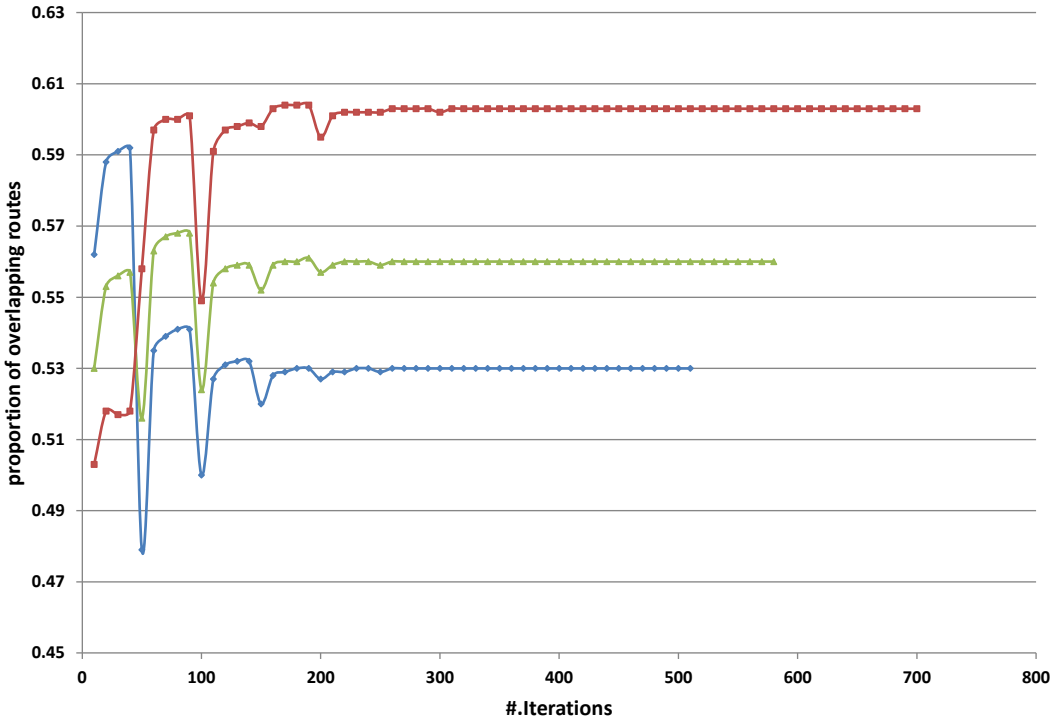
$$c_a^{car} = t_a^{car,\,free}\{1+\alpha([v_a^{car} + \eta v_a^{truck}]/2Cp_a)^{\beta}\} \tag{3.19}$$

$$c_a^{truck} = t_a^{truck,\,free}\{1+\alpha([\rho_{md}\cdot v_a^{car} + \eta v_a^{truck}]/2Cp_a)^{\beta}\} \tag{3.20}$$

Figure 3.4 shows the proportion of the overlapping routes to the total enumerated routes throughout the convergence on Winnipeg and Chicago regional networks.



a) Winnipeg network



b) Chicago regional network

Figure 3.4 Proportion of the overlapping routes throughout the convergence.

It also simultaneously presents the convergence pattern due to the different asymmetric complexities. The experiments use an aggregate measure for analyzing the route composition by counting the overlapping routes rather than pointing out the route that determines the convergence.

The hypothesis of the real network experiment is that O-D pairs with overlapping routes influences the convergence as we identified in the small network experiment which have shown that the O-D pair determines the convergence, thus the large number of overlapping routes is associated with the possibility of the longer convergence. The results demonstrate this in that the proportion of the overlapping routes and the amount of efforts to converge increase as the asymmetric complexities grow.

Increasing pattern of the route composition can be attributed to the fact that the truck's cost function is getting similar to the function of the car as the asymmetric complexities increase. The large proportion of the route share between modes creates the environment where the sheer effects configured by the asymmetric complexities in the cost function can be realized.

The same experiments are conducted on the other networks where the correlation is not strictly observed. As presented in Table 3.6, it seems that a significant difference on the number of iterations is not found in Anaheim network and an increasing pattern is not observed in Barcelona network. However, overall, it is not hard to conclude that the general trend for the lengthy convergence is in the direct proportion to the asymmetric complexities.

Table 3.6 Asymmetric complexities and convergence in terms of #.iterations.

Networks	Asymmetric complexities		
	0.4	0.6	0.8
Anaheim	23	25	29
Barcelona	96	125	110
Winnipeg	449	670	946
Chicago regional	519	589	701

Route Interactions

In order to model route interactions, the following non-additive route cost function is applied to the N-D network.

$$\omega_{rs}^{p} = A_{diag} \cdot V \cdot \Delta_{rs}^{p} 10^{-3} + [A_{diag} \cdot V \cdot \Delta_{rs}^{p} 10^{-3}]^{\rho_{rt}} \tag{3.21}$$

where Δ denotes a column vector representing link-route incidence relationship.

The asymmetric complexity is controlled by ρ_{rt} that translates the non-additive interactions resulting from the route to an appropriate unit for this example. By adopting a function c_a, which is differentiable and monotone, the problem can have a unique solution (Gabriel and Bernstein 1997). Associating c_a with the diagonal part of matrix A easily satisfies the conditions.

In order to handle the non-additive cost function in the route-based assignment technique such as the DP method, it is necessary that all routes should be enumerated before implementing the algorithm because the conventional shortest path algorithm based on the Bellman's principle of optimality is not applicable. However, when the non-additive cost function has the monotone increasing property the problem can be solved with the conventional shortest path algorithm with the appropriate additive term. The proof of the equivalency between two problems can be found in Gabriel and

Bernstein (1997) that showed the feasible regions for two problems are identical. In this experiment, the solutions from enumerated routes and ones obtained by using the additive term taking only the first term of eq. (3.21) with the Dijkstra's algorithm (1959) are compared. The DP method produces the same equilibrium solutions.

In Table 3.7, solutions from three cases of asymmetric complexities are presented when ρ_{rt} is 0.5,1.0 and 1.5. The route flow solutions in Table 3.7 are identical with different asymmetric complexities while only the route costs increased. This is because the route costs' relative scale is preserved for different asymmetric complexities.

Table 3.7 Equilibrium solutions with different degrees of asymmetric route interactions.

direction	Link sequence	Route costs and route flows					
		$\rho_{rt}=0.5$		$\rho_{rt}=1.0$		$\rho_{rt}=1.5$	
Origin 1 to destination 2	2 18 11	62.49	400.00	110.15	400.00	463.78	400.00
	2 17 7 9 11	74.65	0.00	132.99	0.00	608.74	0.00
	2 17 8 14 15	80.16	0.00	143.39	0.00	678.74	0.00
	1 5 7 9 11	74.65	0.00	132.99	0.00	608.74	0.00
	1 6 12 14 15	91.94	0.00	165.68	0.00	836.78	0.00
	1 5 8 14 15	80.16	0.00	143.39	0.00	678.74	0.00
	2 17 7 10 15	77.13	0.00	137.67	0.00	639.97	0.00
	1 5 7 10 15	77.13	0.00	137.68	0.00	639.97	0.00
Origin 1 to destination 3	2 17 8 14 16	82.95	0.00	148.65	0.00	715.12	0.00
	2 17 7 10 16	79.92	108.15	142.94	108.12	675.68	108.16
	1 5 7 10 16	79.92	75.92	142.94	75.94	675.68	75.92
	1 5 8 14 16	82.95	0.00	148.66	0.00	715.12	0.00
	1 6 12 14 16	94.72	0.00	170.94	0.00	875.64	0.00
	1 6 13 19	79.92	615.93	142.94	615.93	675.69	615.93
Origin 4 to destination 2	3 5 7 9 11	68.67	321.46	121.73	321.47	535.71	321.48
	3 5 7 10 15	71.16	0.00	126.41	0.00	565.71	0.00
	3 5 8 14 15	74.19	0.00	132.13	0.00	603.01	0.00
	3 6 12 14 15	85.99	0.00	154.41	0.00	755.59	0.00
	4 12 14 15	68.67	278.54	121.74	278.53	535.76	278.52
Origin 4 to destination 3	3 5 7 10 16	73.95	0.00	131.68	0.00	600.07	0.00
	3 5 8 14 16	76.98	0.00	137.39	0.00	638.07	0.00
	3 6 12 14 16	88.78	0.00	159.68	0.00	793.23	0.00
	3 6 13 19	73.95	0.00	131.68	0.00	600.08	0.00
	4 12 14 16	71.47	0.00	127.00	0.00	569.54	0.00
	4 13 19	56.54	200.00	99.01	200.00	397.79	200.00

Also it numerically demonstrates that both problems with additive and non-additive cost functions used in the experiment have the same feasible region. For different asymmetric complexities, the convergence rate represented by the required number of iterations does not show significant difference. This also might be attributed to the same solutions with just different scales of route costs.

Instead, the study focused on the polynomial type of the route cost function for identifying the influential factors on the convergence. Since the equilibrium process is to settle down the costs of routes carrying the positive flows for each O-D pair into one, it is highly possible that the positive exponent on the route cost could be the influential factor because it determines sensitivity of the route costs that consequently affect the amount of efforts to adjust route costs. Assuming that the influential factors for the convergence of the route interactions are ρ_{rt} and the route flows because it determines the route costs, a measure is devised taking these arguments in its functional form.

$$vblt^{rs} = \sqrt{\sum_{p \in P^{rs}} \left\| C(f_p^{k-1}) - C(f_p^{k}) \right\|} \tag{3.22}$$

The study terms this measure as the route cost variability because it quantifies the fluctuation of the route cost.

The assumption behind the measure is that the unstable solution will highly fluctuate at the beginning but it will finally be stabilized at the vicinity of the UE solution. Thus, the O-D pair that maintains the high variability still has a room to be stabilized and it is highly likely that it determines the final convergence.

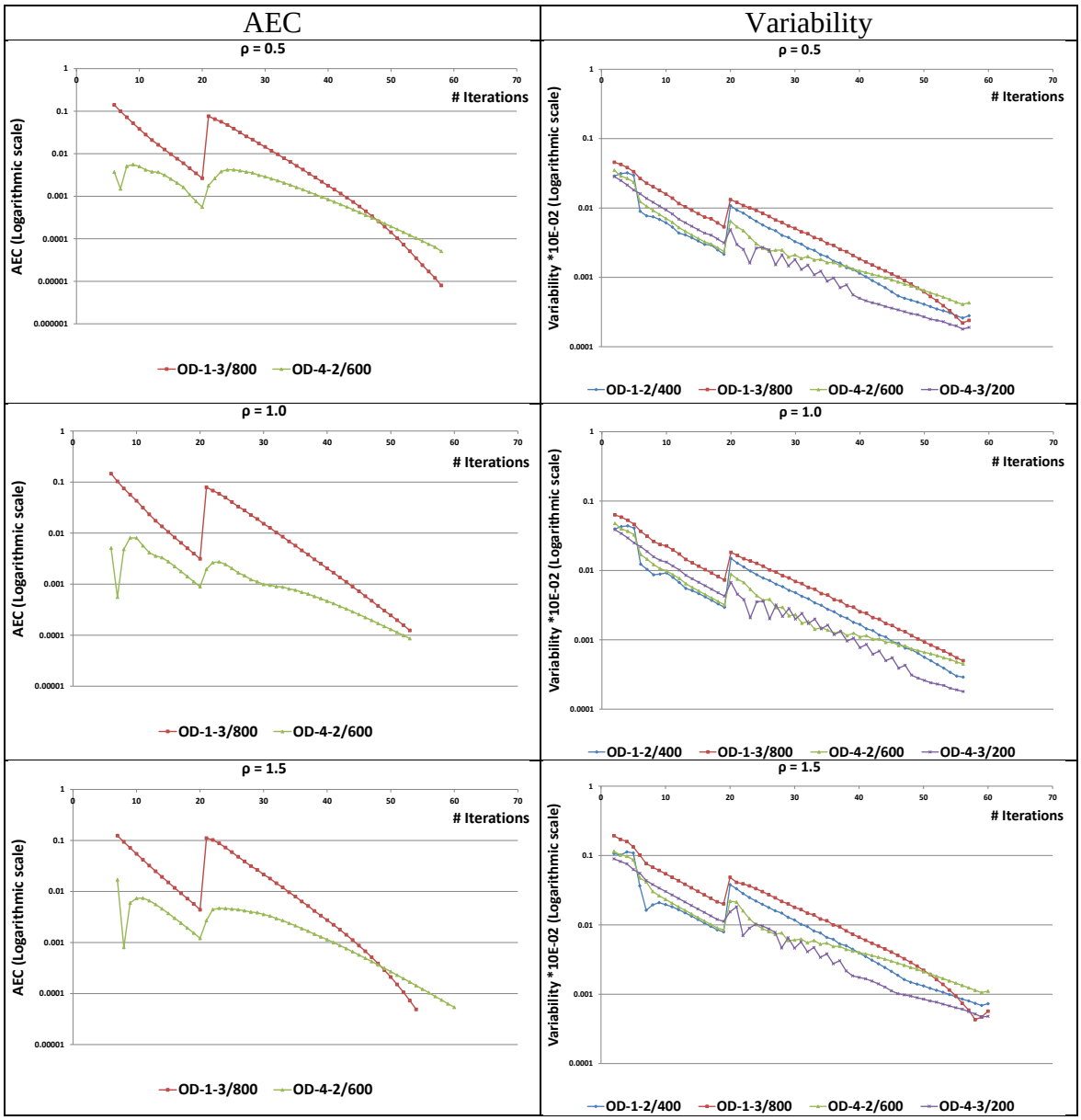


Figure 3.5 AEC and variability for each O-D pair with different ρ_{rt}.

Comparison between the variability for each O-D pair to disaggregated AEC demonstrates that the one with the largest variability exactly coincides with the lastly converged pair (Figure 3.5). The O-D pair associated with the higher variability affected by ρ_{rt} determines the actual convergence. In the figure, the other two O-D pairs are omitted because the demands of the pairs are assigned into one route only for each.

When ρ_{rt} =0.5 and 1.5, the O-D pair connecting zone 4 and zone 2 determines the convergence while when $\rho_{rt} = 1.0$, the one connecting zone 1 and zone 3 determines it.

The same pattern is represented in the right column where the same colored O-D pair keeps the highest variability among all O-D pairs. Based on the experiments, the study makes a conclusion that the convergence of the route interactions is dependent on the route flows and ρ_{rt}'s impact on the route cost.

Real network experiments for the route interactions are conducted with the non-additive route cost function. As in the small network experiment, the following scale-downed non-additive route cost function is utilized on Winnipeg network where the total 174,491 routes are identified by Bekhor et al. (2006).

$$\omega_{rs}^{p} = \sum_{a \in A} c_a \delta_{rs}^{p,a} + (\sum_{a \in A} c_a \delta_{rs}^{p,a} / 10)^{\rho_{rt}} \tag{3.23}$$

As discussed in the small network experiment, the relation between the convergence and the variability is examined in the real network experiments where the link cost function is non-linear. First of all, the variability is plotted with different ρ_{rt}s on the Winnipeg network (Figure 3.6).

We can observe that the bigger exponent incurs the higher variability and consequently the convergence is slow down. As the small network experiment that indirectly associated the variability with the contributing factor on the convergence, the relation between the asymmetric complexities and the convergence is identified in the real network experiment where the higher asymmetric complexities incur the lengthy convergence.

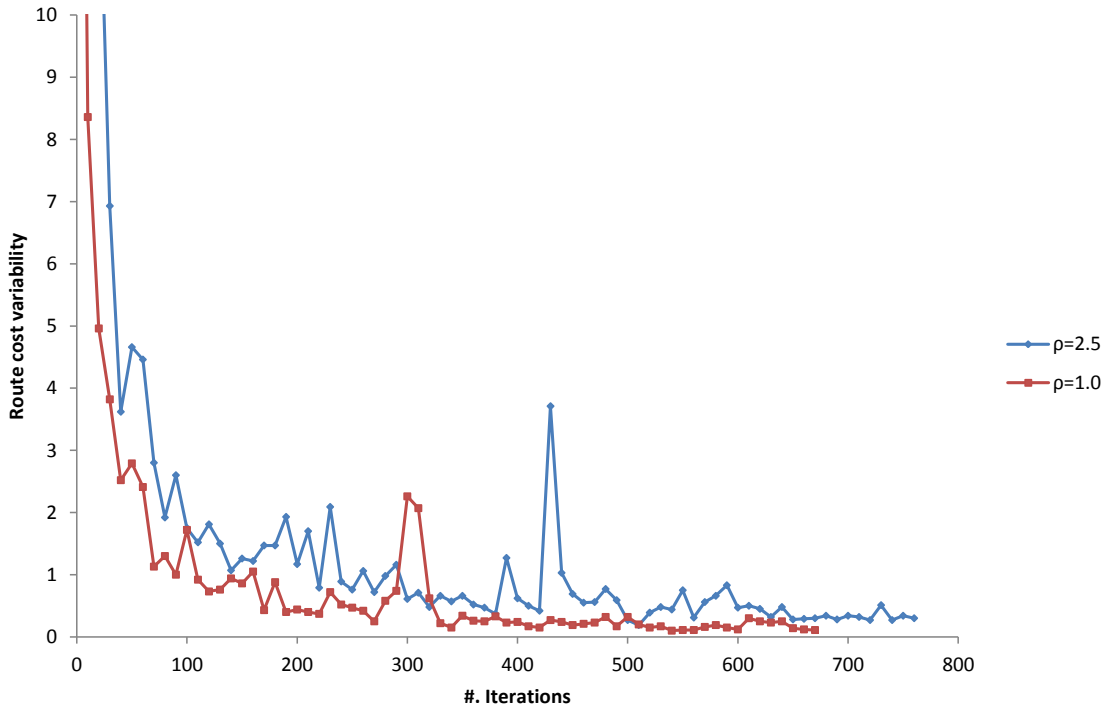


Figure 3.6 Different variability by the different asymmetric complexities.

The experiments extend the variability test to the other types of interactions addressed in the chapter in order to see if there exists a consistent pattern between them. Those applications of the DP method on real networks result in the following computational performance (Table 3.8). The implementation of the route interactions is conducted with the Dijkstra's type shortest path finding method (1959) without enumerating paths by tagging $\rho_{rt} = 1.0$ because it does not violate the Bellman's optimality.

According to Table 3.8, a pattern is found that the problem associated with the mode interactions requires relatively higher computations than others while the cases of link interactions are less burdensome than the problem associated with the route interactions.

Table 3.8 Implementation of the DP method on three types of asymmetric interactions.

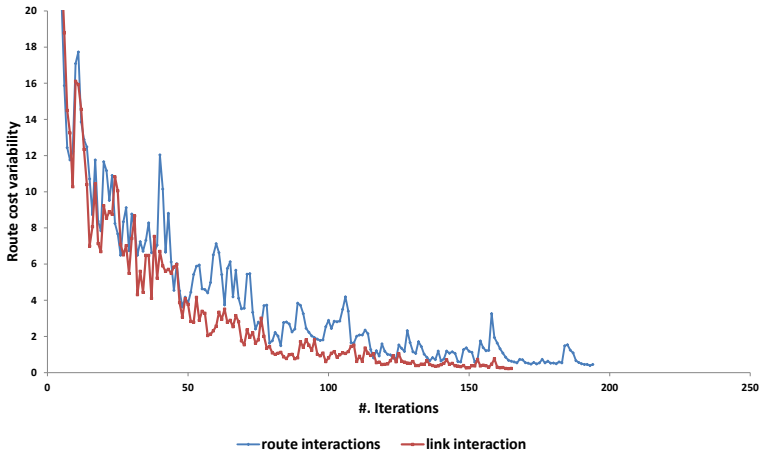
Network	Link interactions		Mode interactions		Route interactions	
	#. iteration	CPU time (sec.)	#. iteration	CPU time (sec.)	#. iteration	CPU time (sec.)
Anaheim	38	0.21	131	1.38	36	0.20
Winnipeg	159	4.22	875	34.26	598	21.56
Barcelona	160	5.89	223	16.53	179	6.91
Chicago	975	56mins	2024	4hr 32mins	1424	1hr 48mins

Even if the relative difference is not strictly consistent because of the Anaheim network experiment, overall, the higher computational works are associated with the route interactions than the link interactions.

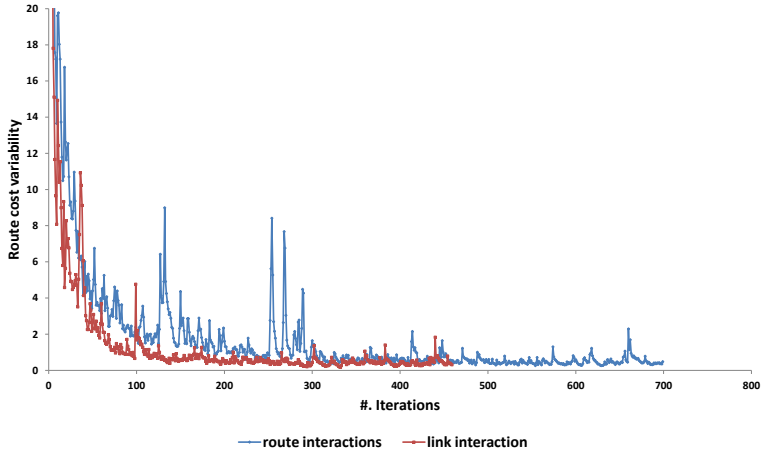
Leaving the mode interactions aside that can be attributed to the simultaneous equilibrium process for more than a single mode, the pattern that the solution finding process for the route interactions requires more computations than the link interactions would be attributed to the structural difference on the cost function.

The route cost function in the route interactions has a complicated form that includes the additive term of the link cost and the exponent term on the sum of link cost. This complexity might be attributed to the lower computational efficiency. When more terms are added into the cost function its sensitivity responding to levels of flows would be higher than the case of the single term as the convergence is affected by the demand intensity.

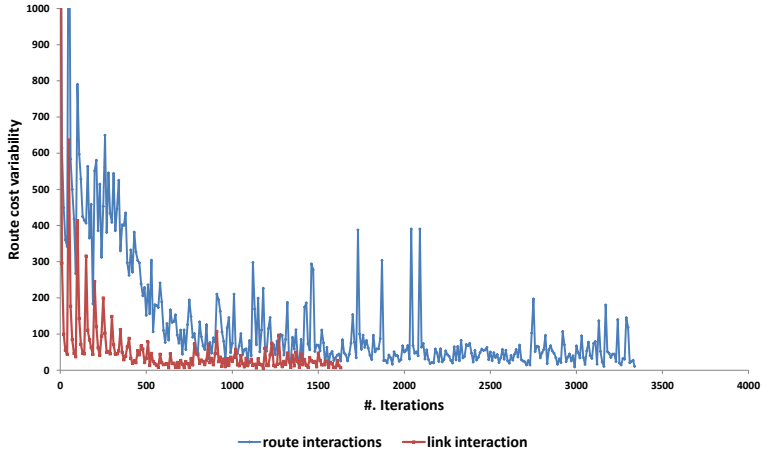
This section analyzes the issue with the variability that measures the levels of fluctuations of the route costs. The basic idea is that if the route cost function is sensitive then route flows would not be easily settled down to the stable condition.



a) Barcelona



b) Winnipeg



c) Chicago regional

Figure 3.7 Comparison of variability between the route and link interactions.

Figure 3.7 shows that the comparison of variability between the route and link interactions for tested networks. The variability of the route cost fluctuates widely and is relatively higher in the route interactions than the one from the link interactions. Because the UE is defined on the route costs, the frequent and wide fluctuations due to the additional term associated with the asymmetric complexities incur the different variability and convergence patterns.

Summary

This chapter investigated the contributing factors that exist behind the established cost function on the convergence of the ATEP through the direct application of the DP method to each case of identified asymmetric interactions. The experimental design is featured with different asymmetric complexities on the N-D network and its extension to the real networks where the experiments are designed to examine and to confirm the new findings.

As the results of the experiments, contributing factors on constructing practical asymmetric effects and convergences for different types of interactions have been identified. In the case of the link interactions, the actual asymmetric complexities were affected by both the supply and demand sides of the problem. The originally established asymmetric effects on the two-way cost function have the symmetric feature due to the network configuration while the demand also affects the realization of the asymmetric effects. The convergence of the mode interactions was dependent on the route composition for each vehicle type. It has been identified that the resemblance of the route composition among modes creates the favorable conditions for the realization of the mode interactions, consequently the convergence is also affected by the situation. Unlike

the other types of interactions, in the case of the route interactions, the factor is indirectly identified using the variability measure in which the fluctuation of the route cost due to the asymmetric complexities associated with the route flows is quantified. In the section, it has been identified that the route costs' sensitivity due to the complexities of the route cost function in the route interactions incurred a patterned computational efficiencies between link and route interactions.

In addition to identifying the contributing factors on the convergence, the acceleration strategy for the DP method to solve the ATEP's computational burdens is developed in the next chapter. The simple theory of the strategy is to properly fit the solution algorithm to the ATEP structure.

CHAPTER 4

ACCELERATION STRATEGY FOR THE CONVERGENCE OF THE DOUBLE PROJECTION METHOD

Basic Concept of the Acceleration Strategy

This chapter is dedicated to developing a strategy for the speedy convergence by fitting of the original DP method and the ATEP. The strategy utilizes the structural characteristics of the ATEP where the problem is broken into each O-D pair. Decomposing the problem and estimating the suitable step-size for each decomposed problem will achieve the efficiencies. The strategy employs another scheme that collectively selects the appropriate step-size for the large leap to the solution between two projections of the DP method. Prior to the introduction of the proposed strategy, previous studies that concerned on the efficient step-size estimation for the projection method are briefly described in the following section.

Efforts for Improving the Projection Method Focused on the Appropriate Step-Size

Despite its simple structure and implementation scheme of the projection operator, some practical challenges are identified because of the absence of an objective function in the VI formulation. This unique characteristic of the VI forces analysts to observe the convergence of the iterative process to a solution. The update equation of the projection method shows that the efficiency depends on two parameters, the step-size β_k and the profitable direction $C(F_k)$.

An adequately estimated step-size is one of the important factors in determining the efficiency of the projection method. However, determining a suitable step-size is challenging because the non-negative step-size is bounded by the unknown Lipschitz constant L and a uniform modulus α. In order to guarantee the convergence, the step-size needs to satisfy the following condition.

$$0 < \beta^l \le \beta_i \le \beta^u \le \frac{2\alpha}{L^2} \tag{4.1}$$

where β^l and β^u are the lower and upper bounds of the step-size, L is the Lipschitz constant of the mapping C such that

$$\left\| C(F_1) - C(F_2) \right\| \le L \left\| F_1 - F_2 \right\| \tag{4.2}$$

and α is the uniform modulus in the strongly monotone mapping of C such that

$$(F_1 - F_2)^T \left[C(F_1) - C(F_2) \right] \ge \alpha \left\| F_1 - F_2 \right\|^2 \tag{4.3}$$

To overcome the difficulty of finding a suitable step-size, various studies have developed step-size rules. In this section, research that developed step-size rules are categorized according to their source of information they referred to estimate the step-sizes.

First, the monotonically decreasing property of the step-size is utilized in developing a sequence of the step-size, which decreases monotonically to the initial step-size β_0. A generalized form of the method of successive average is used by selecting the step-size as follows (Nagurney and Zhang 1996):

$$\{\beta_i\} = \beta_0 \{s_i\} = \beta_0 \{1, \frac{1}{2}, \frac{1}{2}, \frac{1}{3}, \frac{1}{3}, \frac{1}{3}, ...\} \tag{4.4}$$

A similar predetermined step-size sequence conditioned with:

$$\beta_i > 0, \lim_{i\to\infty} \beta_i = 0, \sum_{i=1}^{\infty} \beta_i = \infty \tag{4.5}$$

is used in Dupuis and Nagurney's work (Dupuis and Nagurney 1993).

These simple methods are easy to implement, but they suffer slow convergence due to the strictly decreasing step-sizes. He et al. (2002) identified that Nagurney and Zhang's (1996) method failed to achieve the convergence with 50,000 iterations on the artificially devised test problems.

Second, gradient information is utilized for estimating the step-size (Bertsekas 1976). In the Goldstein-Levitin-Polyak's projection method, Armijo's step-size rule (Amijo 1966) utilizes the gradient information to estimate the step-size. Later, a modified Armijo's rule (Han and He 2001) is proposed in such a way that it does not require gradient information. The sequence above describing step-size rules is strictly decreasing. Typically, those step-size rules monotonically reduce the step-size as the algorithm approaches the optimal solution. The convergence rate of the projection method can be very slow once the step-size becomes small at certain iterations. This is because it remains small at all successive iterations. In order to tackle the diminishing feature of the step-size, a self-adaptive step-size rule that directly controls the step-size is introduced (He et al. 2002; Han and Sun 2004). Under this rule, a self-adaptive scaling parameter iteratively adjusts itself to satisfy the Lipschitz condition and strongly monotone assumption without the prior knowledge of the constants. The step-size calculated with the rule does not always decreases but increases when necessary.

Similar efforts are underway for the extra gradient method (Korpelevich 1977). Khobotov (1987) proposed a step-size rule for the extra gradient method. In the extra gradient method, the strongly monotone assumption on the mapping is relaxed by

performing another projection. This is called the double projection. Marcotte (1991) proposed the primal-dual implementation of the DP method into the NEP. Panicucci et al. (2007) indirectly dealt with the step-size issue in their DP method using a re-initialization scheme. This prevents the step-size from getting smaller by re-initializing it.

The enhancement strategy for the DP method will be introduced in the next section and aims for the convergence improvement through the appropriate step-size rule for the ATEP.

Acceleration Strategy

When the TAP is formulated in the space of route flows, a Cartesian product of all possible route flows represents the feasible set of route flows. However, because each O-D pair is entangled with the corresponding O-D demand, the Cartesian product is constrained by the O-D demand for each O-D pair. Using the update equation, this allows the flow update to decompose into the small sized projections for each O-D pair. This study focuses on the fact that when the TAP is decomposed into each O-D pair a customized step-size for each O-D can be estimated in addition to the single step-size which is based on the whole route flow vectors from all O-D pairs. As long as those step-sizes from different route flow space sizes are properly estimated in order to achieve the convergence, it is advantageous for devising a strategy to utilize available step-sizes to maximize convergence enhancement.

In short, enhancement of the DP method is achieved by two strategies: customizing step-sizes and selecting a better step-size. The former intends to find the appropriate step-size, while the latter finds the step-size for a large leap toward the

solution. The following paragraphs explain what each strategy intends to and provide proof how the new step-size strategy produces a valid result.

Customized Step-Size for Decomposed Transport Network

The proposed methodology utilizes the decomposable structure of the ATEP. The demand constraint associated with each O-D pair naturally decomposes the problem into the projection for each O-D pair route flows. This scheme is very intuitive since the concept of the UE is defined and achieved within each O-D pair.

Step-Size Selection Strategy

The step-size for the space of whole route flow vectors is also available. In other words, the problem can be solved with either of the step-size accounting for the whole route flow vectors β_k, or the step-sizes for smaller sets of route flow vectors comes from each O-D pair β_k^{rs}. The availability of step-sizes from different route flow spaces enables a devised strategy to enhance the convergence. This study exchanges the step-sizes between the predictor projection and the corrector projection in order to improve the efficiency of the algorithm. The basic idea of the methodology is to use the bigger step-size between two step-sizes; β_k, β_k^{rs} Even if an appropriate step-size is estimated for the projection for the route flow vectors of a smaller space, it can be replaced with another step-size if it enhances the convergence as long as the new one is valid. Also, the DP method provides a chance for the replacement between the predictor and the corrector projections.

Valid Step-Size for Entire Route Flow Space

In order to satisfy the inequality condition eq. (3.3) for two different sets of route flow vectors, this study estimates β_k, β_k^{rs} sequentially. This sequential method guarantees suitable β_k that satisfies eq. (3.3) with a few number of iterative procedures due to the non-expansive mapping $C(F)$. Namely, for certain number of iterative implementations for reducing O-D specific step-size $\beta_k^{rs} \le \varepsilon \frac{\left\| F_k^{rs} - \bar{F}_k^{rs} \right\|}{\left\| C(F_k^{rs}) - C(\bar{F}_k^{rs}) \right\|}$ for each O-D pair, $\beta_k \le \varepsilon \frac{\left\| F_k - \bar{F}_k \right\|}{\left\| C(F_k) - C(\bar{F}_k) \right\|}$ is guaranteed because $F_k^{rs} \subseteq F_k$ and mapping $C(F)$ is non-expansive.

$$\left\| C(F_k) - C(\bar{F}_k) \right\| \le \left\| F_k - \bar{F}_k \right\|, \quad F_k, \bar{F}_k \in \Omega \tag{4.6}$$

Benefits of Using the Bigger Step-Size

Figure 4.1 depicts the benefits of using the bigger step-size for the corrector projection. At iteration k, the predictor projection finds the suitable step-size β_k^{rs} that satisfies the inequality condition eq. (3.3) for its O-D pair. When computation for all step-sizes for all O-D pairs is completed additional estimation β_k based on the whole route flows can be accomplished. Now, we have two available step-sizes for each corrector projection for each O-D pair; β_k^{rs} and β_k. The larger one is chosen for the corrector projection since it leads the projection closer to $\bar{F}_k^{rs}$ than does the smaller one. Figure 4.1 where the larger one is depicted by $\beta_k^{'}$ illustrates that choosing a larger step-size in the corrector projection facilitates the second condition for the convergence.

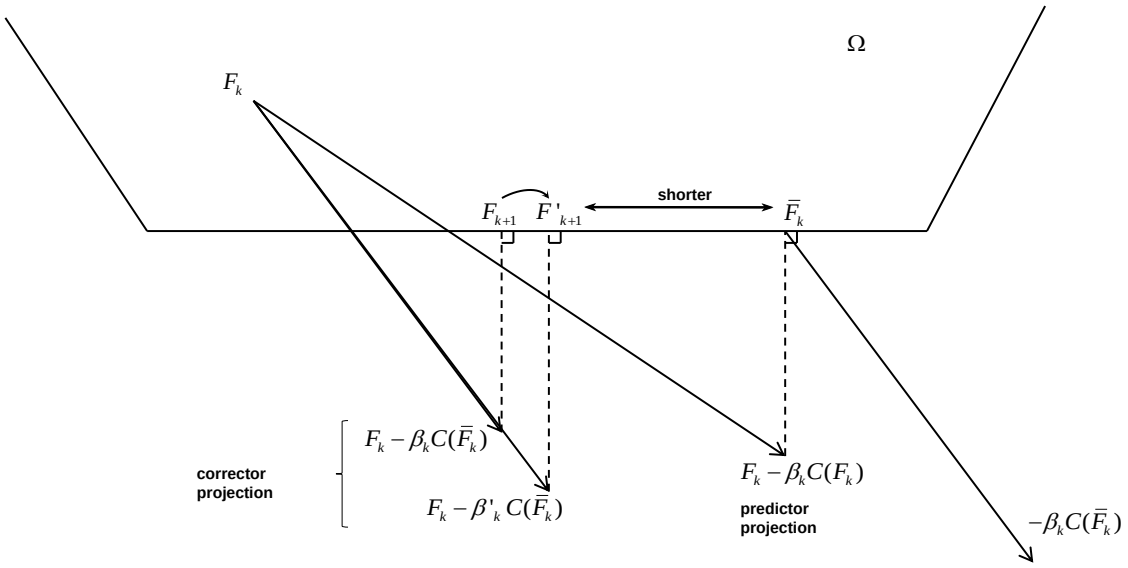


Figure 4.1Graphical representation of benefits using the bigger step-size.

Algorithmic Implementation

The sequential step-size estimation method described above is implemented in a straightforward manner. Satisfying the conditions for smaller spaces β_k^{rs} first, then checking the condition for the whole route flow space β_k is the process for completion. When the condition is not fulfilled for β_k, an additional reducing procedure for β_k is implemented until it satisfies eq. (3.3) for β_k. However, this procedure does not slow down the overall convergence since the condition for β_k is easily achieved by implementing a few reducing procedures for β_k^{rs}. This convergence issue will be discussed in the numerical test section. The following are details of the new algorithm. To assist in understanding the proposed algorithm, a flow chart has been provided in Figure 4.2. The flow chart shows the process for determining the step-size and the escaping route from the predictor projection to the corrector projection for each O-D pair.

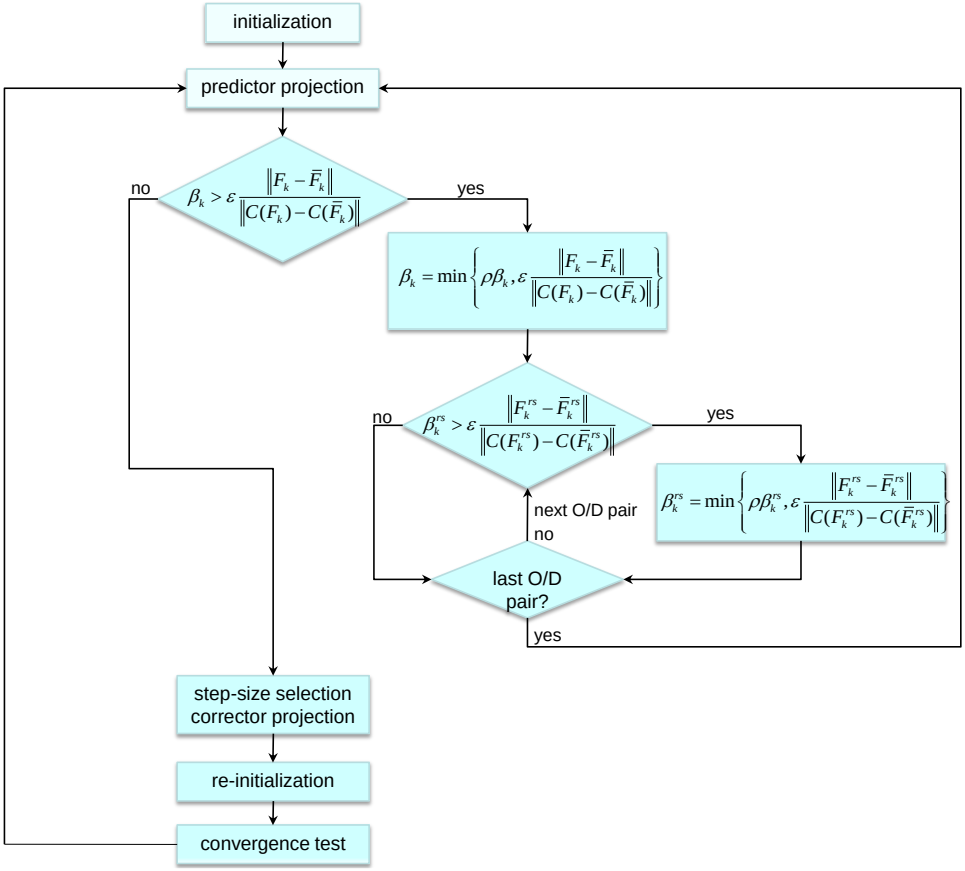


Figure 4.2 Flow chart for the proposed algorithm.

Step 1. set $k = 0$, and select parameters $\varepsilon(0 < \varepsilon < 1), \rho(0 < \rho < 1)$, and initial vector $F_k \in \Omega$, $\beta_k^{rs} = \bar{\beta}$ for each O-D pair with origin $r \in R$ and destination $s \in S$.

Step 2. Find the appropriate step-size β_k, β_k^{rs}

For each O-D pair

$$\bar{F}_k^{rs} = proj_{\Omega}[F_k^{rs} - \beta_k^{rs} C(F_k^{rs})]$$

If $\beta_k > \varepsilon\left[\left\|F_k - \bar{F}_k\right\| / \left\|C(F_k) - C(\bar{F}_k)\right\|\right]$ then reducing β_k

$$\beta_k = \min\left\{\rho\beta_k, \varepsilon\frac{\left\|F_k - \bar{F}_k\right\|}{\left\|C(F_k) - C(\bar{F}_k)\right\|}\right\}$$

Do while $\beta_k > \varepsilon\left[\left\|F_k - \bar{F}_k\right\| / \left\|C(F_k) - C(\bar{F}_k)\right\|\right]$

For each O-D pair

if $\beta_k^{rs} > \varepsilon\left[\left\|F_k^{rs} - \bar{F}_k^{rs}\right\| / \left\|C(F_k^{rs}) - C(\bar{F}_k^{rs})\right\|\right]$ then

$$\beta_k^{rs} = \min\left\{\rho\beta_k^{rs}, \varepsilon\frac{\left\|F_k^{rs} - \bar{F}_k^{rs}\right\|}{\left\|C(F_k^{rs}) - C(\bar{F}_k^{rs})\right\|}\right\}$$

end if

For each O-D pair

$$\bar{F}_k^{rs} = proj_{\Omega}[F_k^{rs} - \beta_k^{rs} C(F_k^{rs})]$$

End do

Step 3. Comparing the step-sizes and perform the outer projection

For each O-D pair

If $\beta_k^{rs} < \beta_k$ then $\beta_k^{rs} = \beta_k$

$$F_{k+1}^{rs} = proj_{\Omega}[F_k^{rs} - \beta_k^{rs} C(\bar{F}_k^{rs})]$$

Step 4. Re-initialize the step-sizes.

For each O-D pair

$$\beta_k^{rs} = \left[\beta_k^{rs} > 0 \,|\, \min\left\{\bar{\beta}, \varepsilon\frac{\left\|F_k^{rs} - \bar{F}_k^{rs}\right\|}{\left\|C(F_k^{rs}) - C(\bar{F}_k^{rs})\right\|}, \beta_k\right\}\right],$$

For entire O-D pairs

$$\beta_k = \min\left\{\bar{\beta}, \varepsilon\frac{\left\|F_k - \bar{F}_k\right\|}{\left\|C(F_k) - C(\bar{F}_k)\right\|}\right\}$$

Step 5. Convergence test.

If convergence criterion such as eq. (3.6) is met then stop. Otherwise, $k = k + 1$ go to step 2.

Numerical Examples

Even if the step-size selection strategy is utilizing the valid step-size, choosing the bigger step-size is not a mathematically proven technique but close to the heuristic variations, thus, several numerical tests for demonstrating adaptability and performance of the proposed methodology have been set up on different network sizes. The computer specifications are the same as the one used in the previous chapter. Before presenting the numerical tests' results, experimental conditions such as the route enumeration strategy, initial parameter setting, and stopping criterion need to be addressed.

When the TAP is formulated with the route flows, the main concern is to manage the number of routes, which grow exponentially as the network size increases. This study utilizes the column generation approach that creates a new route only when a new route is necessary to achieve the UE status (Bertsekas and Gafni 1982). However, it is not always necessary to implement the column generation for every iteration since the efficiency of the DP method can be achieved by the alternating process between the corrector and approximated predictor projection which is not necessarily precise for each iteration. According to the numerical test that examined the appropriate interval of iterations for the column generation (Panicucci et al. 2007), in terms of CPU time, it is most effective for the DP method when implemented every ten iterations. The stopping criterion used for the numerical example is the same as the previously used for examining

applicability of the DP method on the real transportation network. The network used in this example is the same in Chapter 3.

Performance Overview

The performance of the proposed algorithm is applied to the identified three cases of asymmetric interactions and the computational results from both proposed algorithm and the original algorithm (Panicucci et al. 2007) are presented in Table 4.1.

Table 4.1 Computational results of the proposed algorithm.

Network		Proposed algorithm		Basic algorithm	
		#. iterations	CPU time (sec.)	#. iterations	CPU time (sec.)
Link interactions	N-D	30	0.00	29	0.00
	Anaheim	38	0.21	59	0.17
	Winnipeg	159	4.22	678	14.17
	Barcelona	160	5.89	4916	127.23
	Chicago Regional	975	56mins	5000+	6hrs+
Mode interactions	N-D	58	0.02	95	0.10
	Anaheim	131	1.38	376	1.32
	Winnipeg	875	34.26	10465	324.12
	Barcelona	223	16.53	6004	215.47
	Chicago Regional	2024	4hr 32min	5000+	13hrs+
Route interactions	N-D	43	0.00	50	0.01
	Anaheim	36	0.200	278	0.99
	Winnipeg	598	21.56	3007	108.59
	Barcelona	179	6.91	5630	173.28
	Chicago Regional	1424	1hr 48mins	5000+	6hrs+

There are noticeable benefits from the specified step-sizes for each O-D pair and step-size selection strategy for all aspects of the computational performances with the exception being for small-sized networks. For the N-D network, the basic algorithm needs less effort than the proposed one. In the Chicago network example, a few special engineering techniques (such as discarding column carrying zero flows, applying the

proposed step-size selection strategy only for the un-converged O-D pairs) are utilized in order for the efficient implementation of the algorithm. The column generation is also implemented with a longer interval (50 iterations).

Mechanism of the Proposed Algorithm

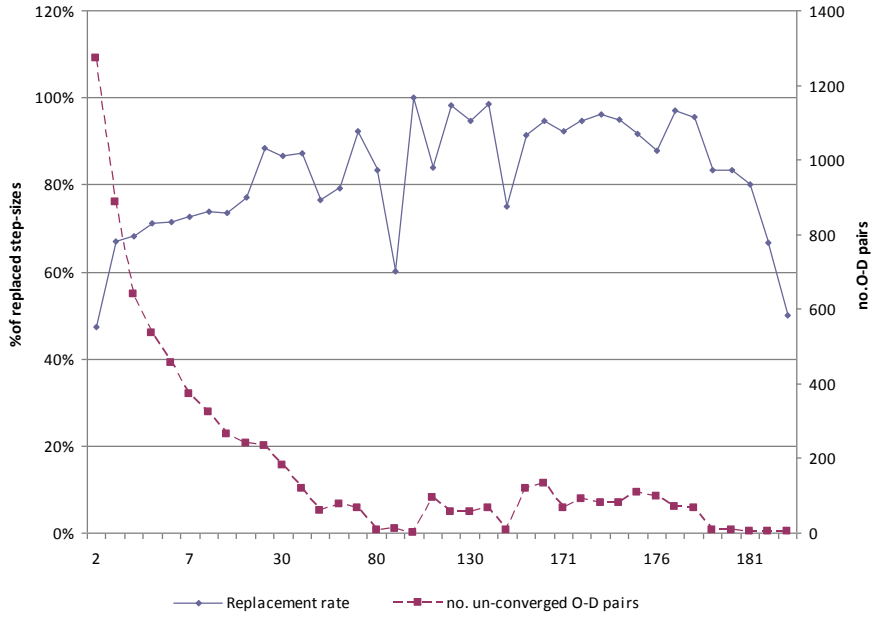
The Winnipeg Network is utilized for analyzing two strategies how the proposed algorithm preformed better than the basic algorithm. This section provides the relevant information on how that was accomplished. Since the proposed algorithm is designed to facilitate the convergence, a convergence profile for O-D pairs having more than one route is constructed in order to easily capture the convergence trend. The following analysis is based on the constructed route information on those routes.

Utilization of step-size selection strategy

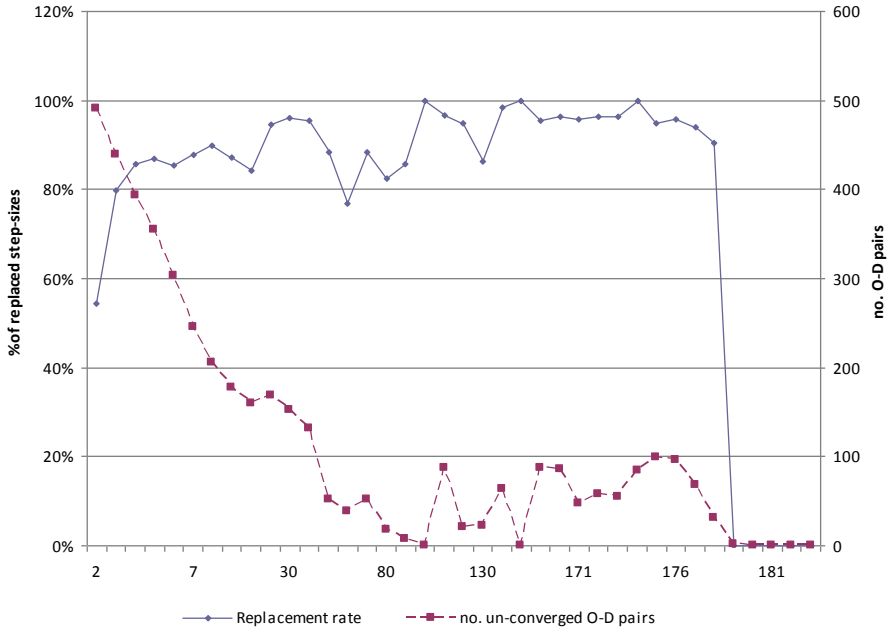
An analysis is conducted on whether during the corrector projection, the step-size selection strategy is actively engaged in replacing the smaller units with larger units. This is examined by counting the number of replacements only for the not-converged O-D pairs. Figure 4.3 shows the step-size selection usage during the corrector projection as the algorithm continues. It illustrates only two cases of O-D pairs producing two (upper graph) paths and three (lower graph) path routes. The solid line of each graph indicates the step-size replacement rate while the dotted line is the number of un-converged O-D pairs. It is easily identified that a larger step-size from the whole O-D pairs is highly utilized for the corrector projection during the whole iterative process.

Illustration of benefits for using the bigger step-size in the real network

Figure 4.4 shows how the larger step-size effectively improves the solutions in the real network implementation.

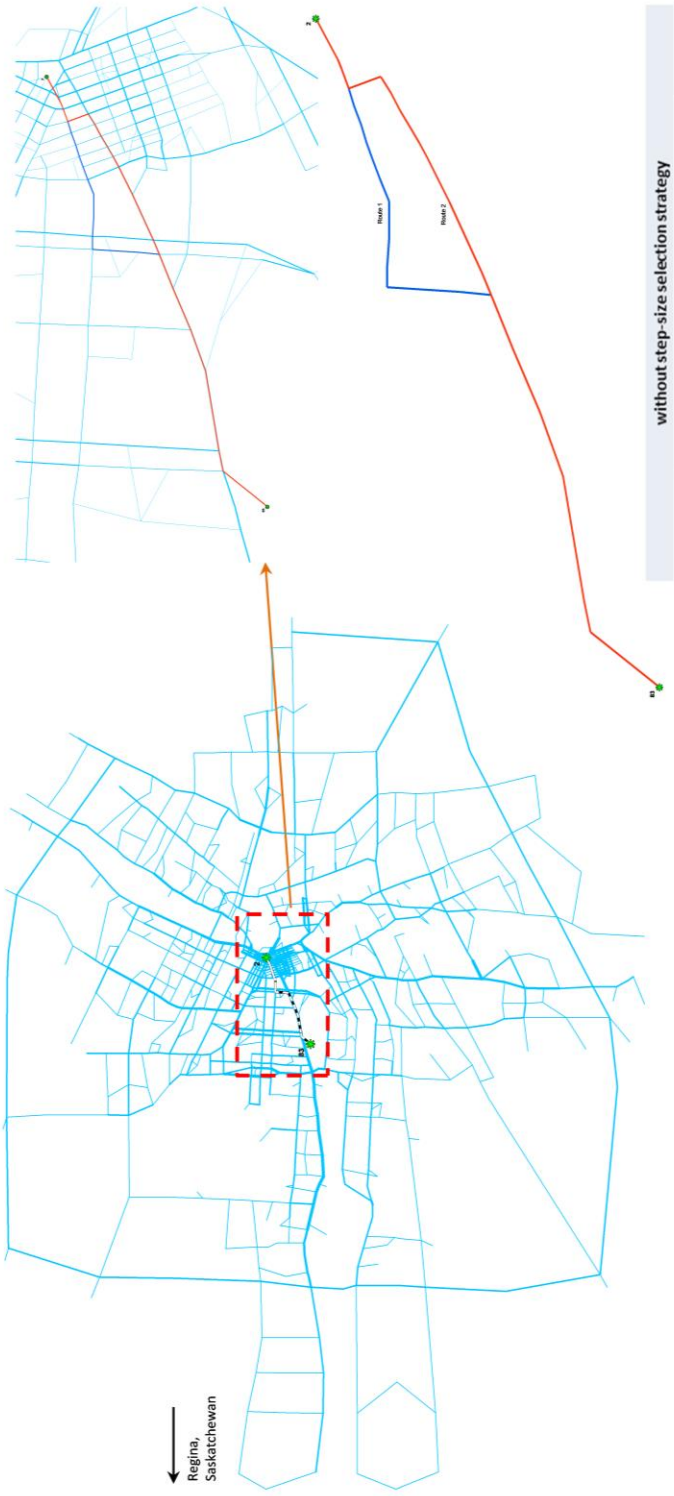


a. A case of O-D pairs producing two routes



b. A case of O-D pairs producing three routes

Figure 4.3 Usage of bigger step-sizes for the corrector projection.



without step-size selection strategy

#.Iterations	Flows		Travel time		UE
	route1	route2	route1	route2	status
2	11.82	0.18	8.37	8.36	x
3	11.82	0.18	8.32	8.33	x
4	11.82	0.18	8.29	8.31	x
5	11.82	0.18	8.27	8.3	x
6	11.82	0.18	8.26	8.3	x
7	11.82	0.18	8.25	8.3	x
:	:	:	:	:	:

with step-size selection strategy

#.Iterations	Flows		Travel time		UE
	route1	route2	route1	route2	status
2	11.76	0.24	8.37	8.36	x
3	11.78	0.22	8.32	8.33	x
4	11.89	0.11	8.29	8.31	x
5	12	0	8.27	8.3	O

Figure 4.4 Illustration of benefits from the step-size selection strategy.

In the experiment that compares the route flow adjustment using the proposed strategy, dumping route flows from the unattractive route (Route 2) to the shortest route (Route 1) is rapidly conducted due to the larger step-size (See the lower left corner of Figure 4.4). In Figure 4.4, an iterative process of equilibrium between two routes has been illustrated on a selected O-D pair that created two routes in the algorithm's beginning. The former utilizes proposed step-size strategy while the latter implements the O-D pair without the strategy.

Implementation of the sequential method

In addition, this section analyzes how many implementations for the reducing procedure are required for satisfying the inequality condition for eq. (3.3). Table 4.2 represents the reducing procedure average when counted every ten iterations.

Table 4.2 Number of required reducing procedure implementation for each O-D Pair.

Iteration/network	Anaheim	Winnipeg
10	1.8	1
20	1.2	1.2
30	1.4	1
40	1	1.1
50	1	1.2
60	-	1.1
70	-	1.1
80	-	1.3
90	-	1.2
100	-	1.0
110	-	1.2
120	-	1.1
130	-	1.1
140	-	1.0
150	-	1.1
160	-	1.3
170	-	1.2
180	-	1.2

Quick analysis indicates that the procedure does not hinder the overall process. Generalization of the required number of reducing procedures' pattern cannot be done at this stage. However, more projection needs to be implemented at a relatively earlier stage.

Performance with different flow update-strategies

This section compares the performance comparison from different flow-update strategies in the ATEP. As described earlier, the high-utilization of the step-size selection strategy contributed to the significant performance enhancement of the DP method. This section attempts to explore the factor that influences the high-replacement rate in the flow update strategy.

As presented in the previous section, the proposed algorithm is based on the all-at-once flow update. In other words, the flow update is conducted for all O-D pair at once. This flow update method gains its computational efficiency for each iteration because previous fixed cost information is utilized without update while more iterations are required for the convergence. However, if the cost update is conducted immediately after each O-D pair's flow update, then, the computational burden will increase because extra calculation is needed for the next O-D flow update. This flow update strategy is useful because it accelerates the convergence with relatively new information on the link cost despite of the price for cost update at each iteration.

This section compares the performance of these two flow update strategies and identifies the factor that causes different performance. The convergence behaviors represented by the AEC is plotted in Figure 4.5. The x-axis is the computation time in seconds while the y-axis is the relative gap in the logarithmic scale. The figure clearly shows that the advantages of using the proposed algorithm for implementing the ATEP

problem. Even if the convergence has been successfully achieved by the two methods, its rate to the given relative gap ($\varepsilon = 10^{-8}$) is significantly different. The proposed algorithm takes less than 12 seconds for the given stopping criterion while the one-at-a-time flow update takes more than 4 minutes. The fluctuation of the relative gap is attributed to the intermittent column generation of the DP method.

Table 4.3 summarizes the computational results in terms of the number of iterations, projections for O-D pair, link cost update, and CPU time as the convergence criterion gets strict. It is important to note that the total number of iterations for proposed algorithm is significantly lower than those of other flow update strategies. In addition, the table shows that step-size selection strategy also accelerates the one-at-a-time flow update strategy.

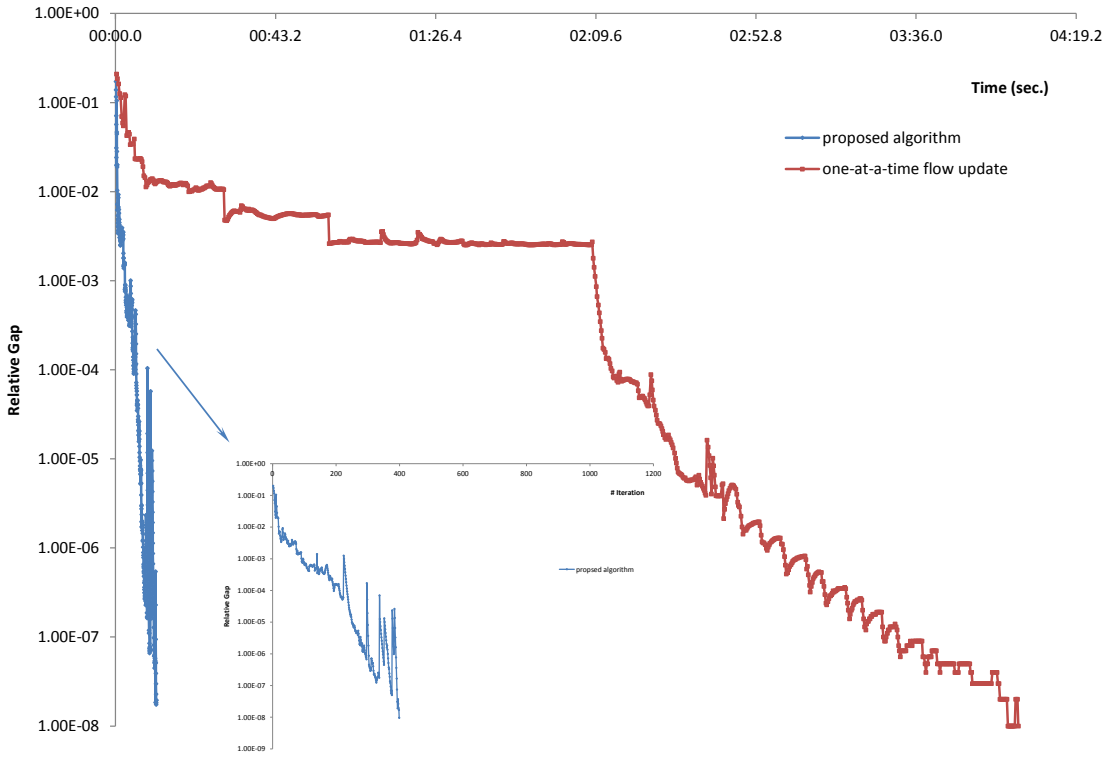


Figure 4.5 Diminishing relative gaps by proposed and one-at-a-time flow update.

Table 4.3 Comparisons of computational results from different flow update strategies.

update strategies/ $\zeta(f)$		10^{-5}	10^{-6}	10^{-7}
proposed algorithm	# Iterations	159	239	260
	# projections	323160	429841	456385
	# link cost update	1.69E+06	2.05E+06	2.18E+06
	CPU time (sec.)	4.22	7.12	7.60
one-at-a-time with step-size selection	# Iterations	368	428	477
	# projections	680280	836029	919781
	# link cost update	9.16E+07	1.10E+08	1.23E+08
	CPU time	1min 37secs	1min 54secs	2min 10secs
one-at-a-time w/o step-size selection	# Iterations	598	883	1053
	# projections	1169091	1737417	2083998
	# link cost update	1.55E+08	2.31E+08	2.74E+08
	CPU time	2min 46secs	3min 58secs	4min 47secs

Even if the one-at-a-time flow update strategy is known to be better than all-at-once flow update strategy (Chen 2001) in terms of the required number of iterations, the table shows different results. The following section addresses this issue focused on the utilization of the step-size selection strategy. The previous section illustrated that the step-size selection strategy is highly utilized by the proposed algorithm. Figure 4.6 shows how the step-size selection strategy is frequently utilized in the one-at-a-time flow update strategy. It is noticeable that the level of utilization of the whole O-D pair's step-size between the two flow update strategies is significantly different. The bigger step-size from the whole O-D pair is actively adopted by the proposed algorithm while its utilization is minimal by the one-at-a-time flow update strategy. This implies that the step-size from the whole O-D pair is not bigger than the one from the one-at-a-time flow update. Generally, at the beginning of the solution process of the TAP, the step-size is relatively big because many O-D pairs are not converged yet while the step-size becomes small for the fine tuning of the route flows for the UE condition under stopping criterion. Thus, at certain iteration, if there are many un-converged O-D pairs, then the step-size for the whole O-D pair would be big. In the all-at-once flow update, this big step-size is

directly represented in the whole O-D pair's step-size while in the one-at-a-time flow update, the whole O-D pair's step-size gets smaller whenever each O-D pair's route cost is updated because the update enhances the UE condition of the O-D pair. The diminishing step-size lowered the effective use of the step-size selection strategy and consequently the convergence also has slowed. Moreover, the computation for the whole O-D pair's step-size after each O-D pair's cost update would deteriorate the computational efficiency.

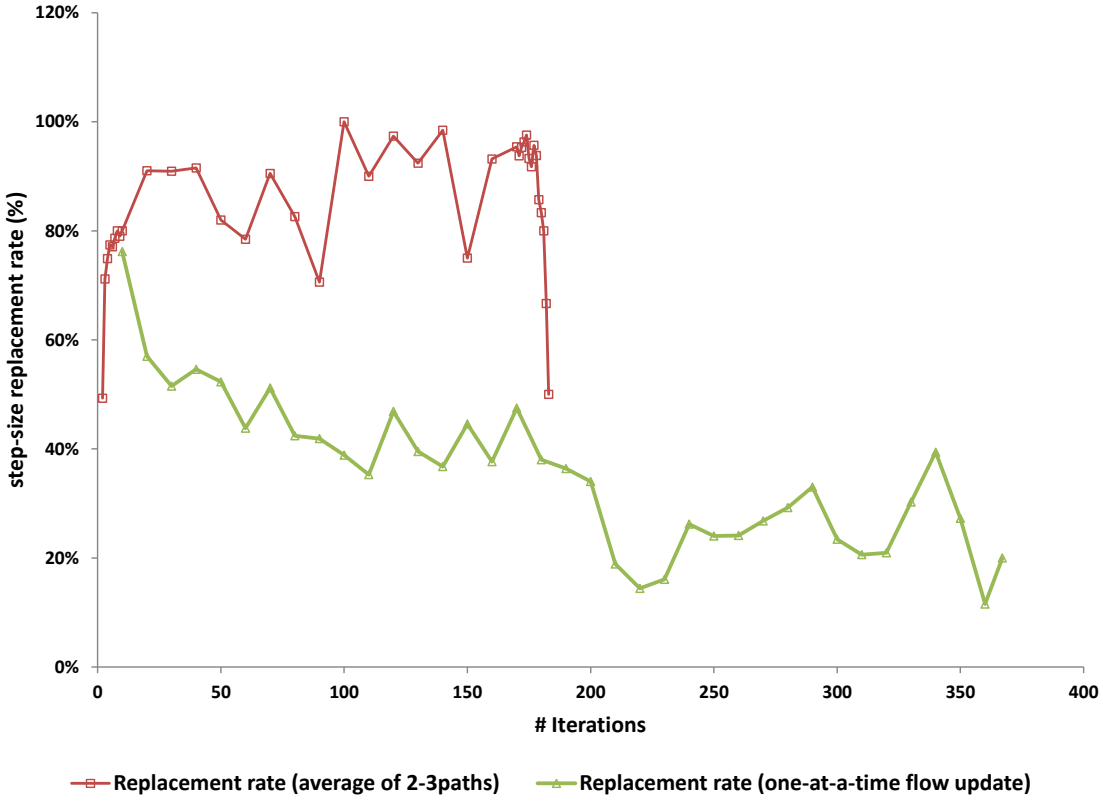


Figure 4.6 Utilization of the step-size selection strategy by the different flow update.

Summary

This chapter introduced a heuristic solution algorithm acceleration strategy for the DP method designed for the ATEP. The efficiency of the new algorithm was achieved by effectively treating the step-size. Formulating the ATEP in the space of route flows, both customized step-sizes for each O-D pair and for the entire route flow space were produced. The production of these step-sizes devised a selection strategy for allowing the corrector projection to choose the larger step-size. The strategy was successfully implemented because the DP method provides an opportunity to select an alternative step-size between the two projections.

The numerical tests demonstrated that the new algorithm significantly relieved the computational burdens of the basic algorithm. In addition, the details of the mechanisms of the proposed algorithm have been identified. The O-D specific step-size associated with the step-size selection strategy sufficiently evolved the current solution. This achievement allowed for the immediate utilization in the corrector projection for the step-size selection strategy. The problem was quickly converged due to the close connection of the effects. Lastly, the analysis on the flow update strategy identified that the all-at-once flow update contributed to the larger step-size from the whole O-D pair that consequently increased the utilization of the step-size selection strategy.

The next chapter of the dissertation addresses the practical application of the ATEP in the transit assignment using the DP method. Especially, the chapter seeks to answer the question; Can the ATEP be modeled to provide insight on crowding effect for the transit system?

CHAPTER 5

INVESTIGATIONS ON MODELING RANDOMNESS OF CROWDING IN THE LARGE SCALE PUBLIC TRANSPORTATION SYSTEM

In this chapter, the crowding effect in a transit vehicle is modeled in the time expanded network considering the daily variation of the passenger flows. Traditionally, the two major attributes, time and cost, have been extensively utilized for explaining the disutility associated with travelling (Small and Verhoef 2007). However, as the analytical tools and the survey method develop, various qualitative attributes that influence the experience of travelling are identified. One of the attributes linked to disutility of travelling from using the public transportation would be crowding in a transit vehicle. People traveling on public transportation have been held to have a diminished or reduced expectation of privacy due to crowding. It is expected that more emphasis will be placed on the crowding by users as the economies are matured enough to be called the developing and developed countries (Tirachini et al. 2013). The expectation is based on the socio-economic trend that people would concern on the comfort and quality of the transit service as their income increase. The actual use of the crowding cost function in practice in Australian Transport Council guidelines (2006) supports the trend that attaches the qualitative aspect to the traditional time and cost based disutility.

The objective of this chapter is to model crowding effect considering the day-do-day variation of the transit users in the time expanded network and to examine its modeling effect. Crowding is the direct experience by the transit users as the traffic congestions in the roads. When the level of crowding fluctuates by the daily variation of

the transit demand, the variation is also easily recognized by the users. In light of this, it is desirable to consider the random effect in the transit assignment model. A few studies approach the issue from the modeling perspective, however, the explicit consideration for the effect in the schedule-based model has not been done in the real large scale network. Especially, lack of studies is found that utilized the real network because the non-additivity property of the route cost due to the uncertainty modeling.

The rest of the chapter is organized as follow. The next section describes the UTA system which will be used as the case study network for following two application studies of the dissertation followed by the concept of the time expanded network and its construction for the UTA system. Then, the study formulates the problem with the crowding cost function and introduces the solution algorithm focused on the reliable shortest path algorithm. In the numerical example, the section investigates the expected outcomes by applying the crowd modeling into the transit assignment in the time expanded network. The section attends the behavioral change due to the crowd modeling and extends the investigation into the details of the cost and the variability caused by the users' attitudes and the daily fluctuation of passenger flows. The last section concludes the chapter.

Case Study Network – UTA System

The study area for this research is based on the UTA transit network covering Weber, Davis, Salt Lake, Tooele, and Utah counties in the state of Utah. The major transit corridor covers the Wasatch Front where roughly 80% (2.1 million) of Utah's population resides. Several downtown and commercial areas in the region are connected

by UTA's system that operates buses, light rail (TRAX) and FrontRunner commuter rail. The UTA transit system can be characterized by their services connecting the Salt Lake City area to all other major suburban areas. The majority of bus lines and TRAX pass through the Salt Lake City area. There are only five express bus lines and FrontRunner service provided to passengers residing in other major population centers along the Wasatch Front, specifically the Ogden-Clearfield and Provo regions. The following paragraphs briefly overview the UTA system in terms of its transit vehicles, fare structure and ridership pattern.

Transit Vehicles

Major vehicle types of the UTA consist of buses, a light rail and a commuter rail. UTA's local bus is the same as the traditional bus system. However, the unique offering in UTA's bus system is MAX, UTA's Bus Rapid Transit (BRT) service. As of 2012, there is only one BRT line in operation. In addition, there are five express bus lines making long-distance trips from the Ogden and Toole areas, with limited number of stops during morning and afternoon peak hours. The bus system map for the core of the Wasatch Front where the Salt Lake County is located is presented in Figure 5.1.

In addition, three TRAX lines support downtown Salt Lake City and several of its suburbs. TRAX is the light rail system powered by overhead electrical wires and runs on a steel-tracked fixed guideway installed on the shared and exclusive right of way.

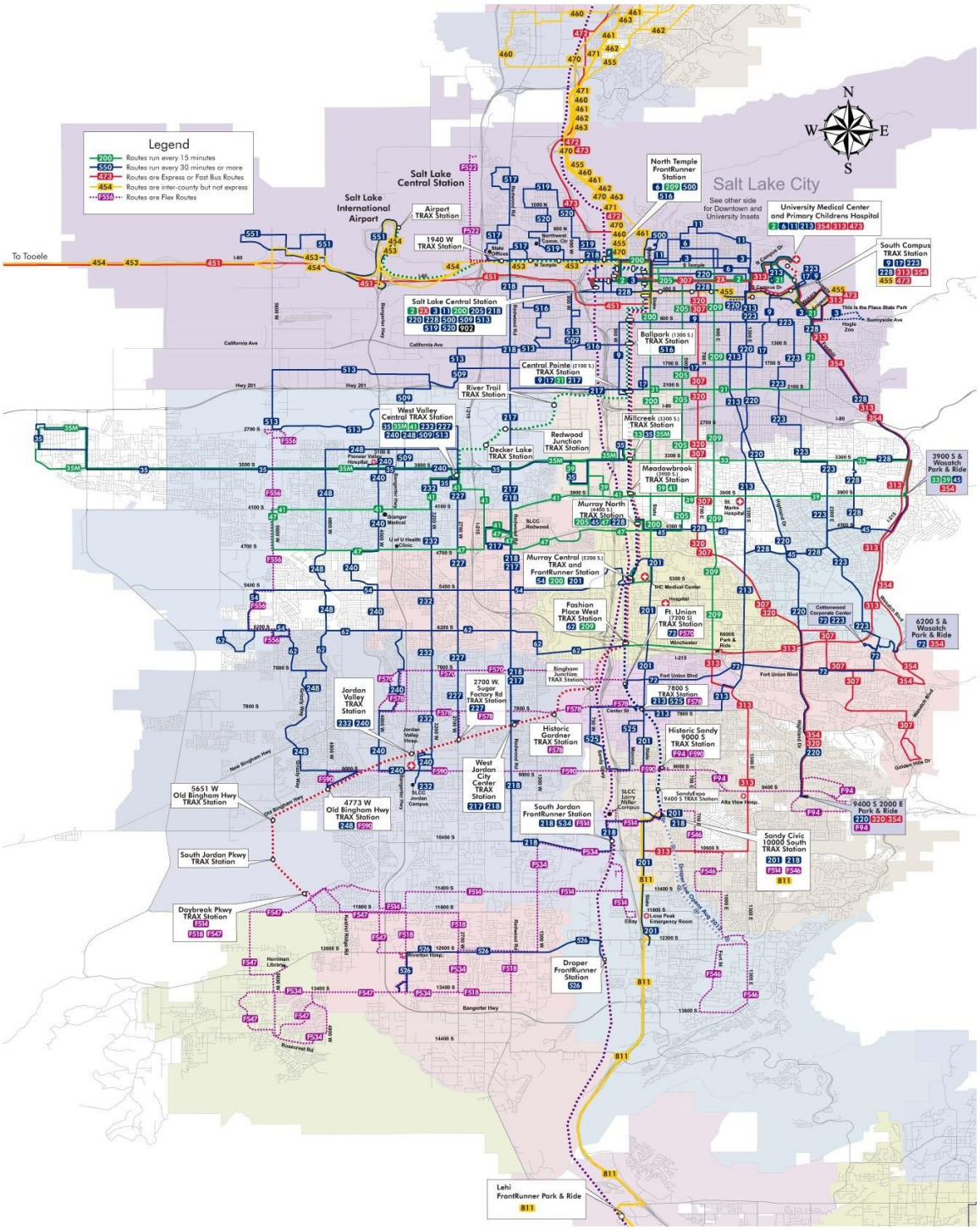


Figure 5.1 Transit system map of Salt Lake County (RideUTA, 2013).

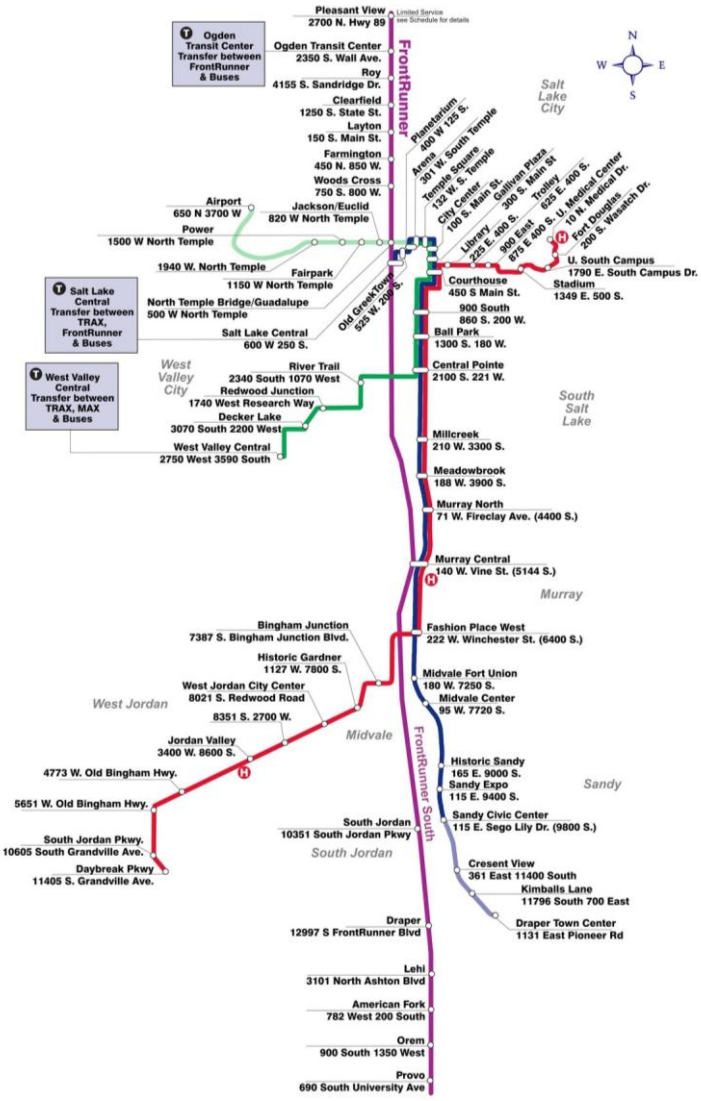


Figure 5.2 Rail system map of the UTA (RideUTA, 2013).

Despite the maximum speed of 60 mph, the train's operation speed in the downtown area is only 10 mph and as fast at 40 mph in the suburban areas. The un-scaled rail system map is presented in Figure 5.2. The planned extensions of the green and blue line are drawn faintly.

Unlike TRAX, FrontRunner is a push/pull diesel locomotive system. FrontRunner runs from Pleasant View in the north to Provo in the south. This is to meet

UTA's goal of increasing mobility to those living along the Wasatch Front. Additionally, UTA seeks to decrease traffic congestion on Interstate-15 of which a large proportion of the corridor is overlapped.

Fare Structure

The fare structure of the UTA network consists of both the local fare and the premium fare systems. As of 2012, the local fare is a one-time payment of $2.35 for the one-way local buses and TRAX. The premium fare is good on all buses, including the express buses and TRAX with $5.25 for one-way (transfer permitted) service. Also, the premium fare includes the FrontRunner payment which provides a free transfer to any bus or TRAX. This fare is distance based with a base fare of $2.35 for travel to one station plus $0.55 for each additional station. The maximum fare from Provo to Pleasant View is $9.50 one way.

Ridership

The average weekday's daily ridership of the UTA is about 145.3 thousand trips (APTA, 2013) including trips for FrontRunner, TRAX and express, MAX, and local buses. The ridership pattern over past 12 years has been plotted in Figure 5.3. Even if there are some up-and-downs, overall, the UTA's transit demand grows with 3.7 percent annual growth rate during the period. After the commuter rail is introduced in 2008, the ridership is a bit shifted up and stabilized around 140 thousand trips. The dotted line indicates the moving average with 2 year periods.

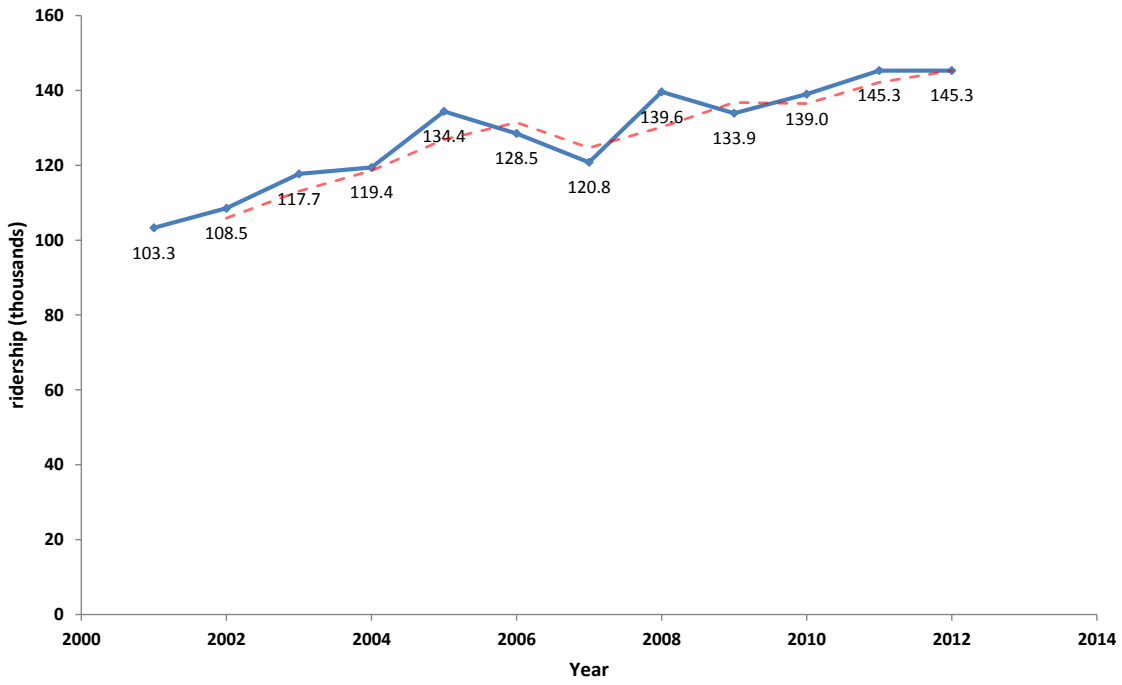


Figure 5.3 Ridership trend of the UTA.

Time Expanded Network and Construction

Time Expanded Network

To be able to describe the transit network associated with the timetable, a time expanded network $G = (TN, TA)$ is constructed with a set of transit nodes $i \in TN$ tagged with a particular time $t(i)$ and a set of transit links. Thus a transit link $(i, j) \in TA$ represents both the spatial movement starting at $t(i)$ and ending at $t(j)$, and time spending $t(j) - t(i)$. These transit links are divided into two classes; one associated with spatial movement and time spending, and the other one associated with time spending only.

The first class includes:

In-vehicle links represent the movement of trips by transit vehicles, where the node i, j of a link (i, j) does not include the set of origin and destination nodes.

Boarding links bridges the travel demand in the origin into the adjacent transit vehicle run, where i is the origin and j is the head node of the transit link.

Egress links connect the in-vehicle link into the final destination node, where i is the end node of a transit node and j is the final destination.

While the first class describes the typical physical movement of trips, the second class is used for describing the time spending at a stop or transfer.

Transfer links describe the transfer behavior at a stop, where i, j are the transit nodes representing different runs of transit vehicles.

Stationary links are used for modeling the on-board passengers continuing to the next stop, where i, j are the transit nodes representing the same run of a transit vehicle.

A trip is described on the time expanded network by an alternating sequence of these links. A passenger's route from an origin to a destination is starting at the origin with a boarding link and the combination of in-vehicle, stationary, and transfer links and reaching the destination via the egress link. The following Figure 5.4 depicts a conceptual route to the destination using the links and nodes described above.

When a passenger wants to transfer into another run of a transit vehicle, or continue the journey as on-board passengers, a FIFO (First In First Out) rule is modeled. Thus, an ordered set of links $S^-(j)$ that incident to every transit node j is prepared as:

$$S^-(j) = \{(i_0, j), (i_1, j), (i_2, j), (i_3, j), ..., (i_n, j)\}$$

where (i_0, j) is the stationary link that carries the on-board passenger continuing to the next stop.

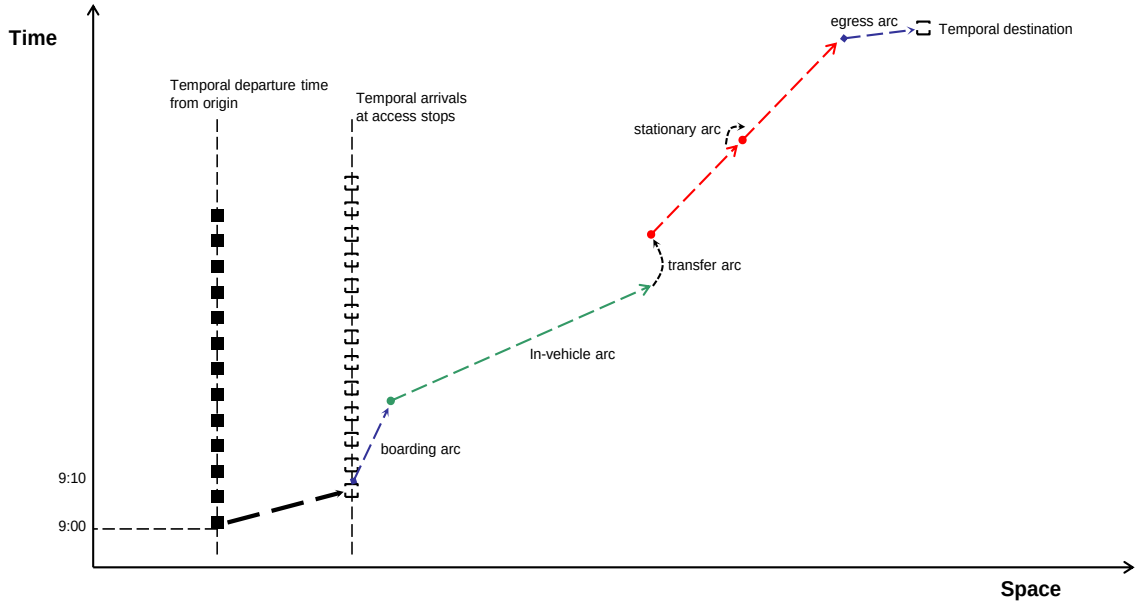


Figure 5.4 Route construction in the time expanded network.

Under the FIFO rule, the on-board passengers have the priority for occupying the transit vehicle while the other passengers who want to board the vehicle need to share the residual capacity according to their respective arrival time at the stop. The passenger flows in $S^{-}(j)$ yields the asymmetric feature in the cost function.

Construction of the Time Expanded Network

In the planning stage, the operational details of future transit lines are not provided. That is one of the reasons that the frequency-based transit assignment model is preferred to the technique based on the time-expanded network. However, with the same level of information given to the frequency-based assignment technique, the time-expanded network also can be easily constructed.

Figure 5.5 presents typical transit route information for the frequency-based assignment technique for a future line in the planning stage. It contains the name of the

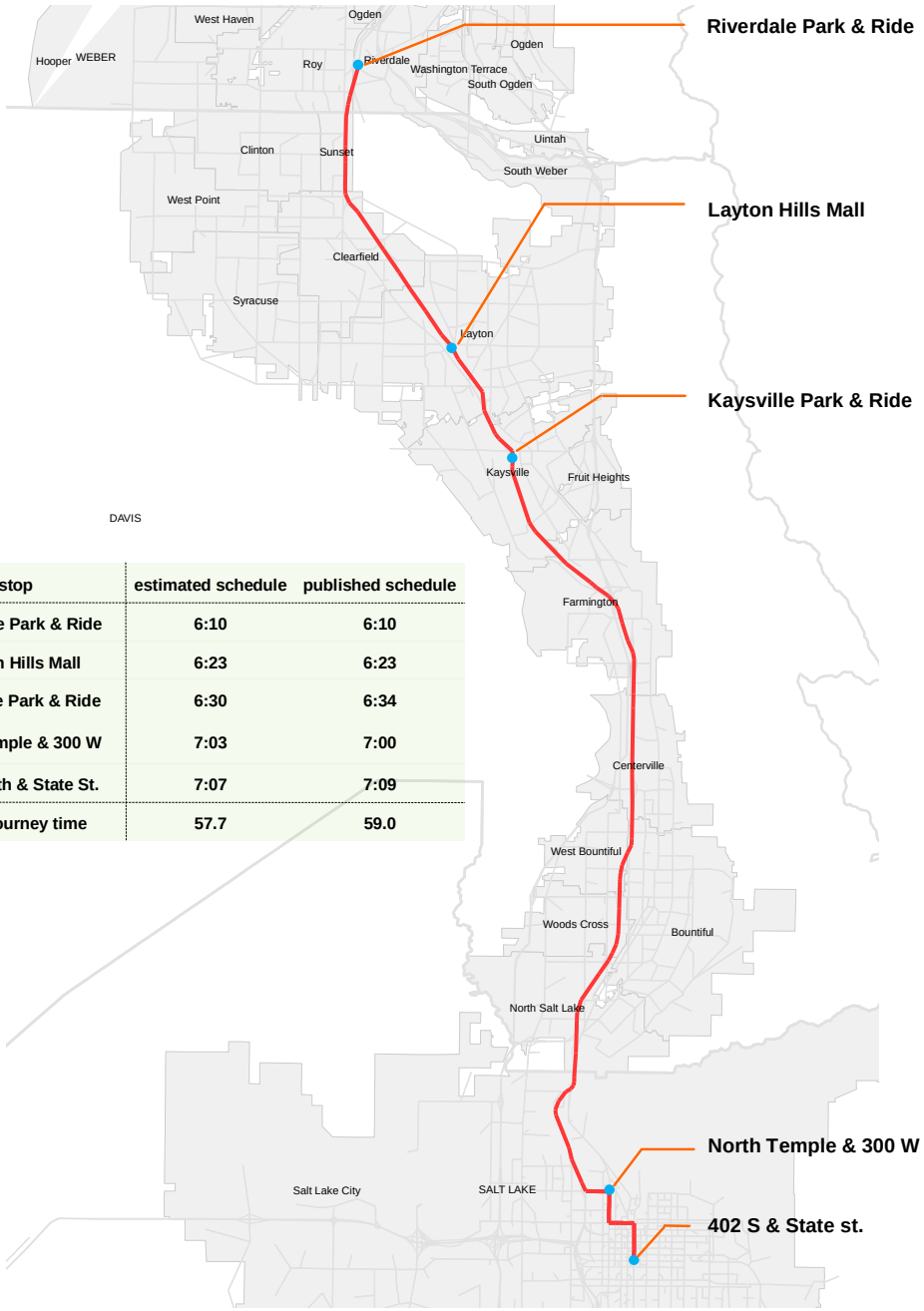
bus line (“route 456”), mode (“6” means the express bus), one-way operation (“T” means true), frequency of the vehicle (30 minute interval), operation speed (25 miles per hour), and the route represented by the node sequence.

```
LINE NAME="ROUTE456", MODE=6, ONEWAY=T, FREQ[1]=30,N=2108,
-1761, -2104, -2102, -2101, -2136,  -2098,  12642,  -2098,  -2020,  -2018,
-2017,  -2016,  -2014,  -1990, -2011,  -1957,  -2458,  1956, -1882,  -1881,
-1879,  -1849,  -1853,  -1839,  -1952, -1837,  -1835,  -1986,   1744,  -2554,
-1809,  -1797,  -1807,  -9251,  -9389, -9409,  -9417,  -9252,  -9434,  -9432,
-9253,  -9254,  -9255,  -9256, -9258,  -9257,  -1598,  -1600,   2522,  -1600,
-1598,  -9259,  -9260,  -9436, -9438,  -9261,  -1526,  -1528,   1530,  -1528,
-1526,  -9262,  -9446,  -9447, -9448,  -9451,  -9453,  -9201,  -10003, -2607,
-2623,  -9329,  -2625,  -10377, -10341, -2373, -7746, -10351,  -13483,  -13485,
-10353, -10354, -7272, -7523, -7525, -10358, -10360, -10362, -10357, -7535,
-10365, -7526, -10506,-10507, -10374, -10512, -10369, -10246, -3611,  -3503,
-3663, 3531, -3625, -3680, -3688, -3717, 3629, -2496, -4861,3631, -3710,
-5445, -3638, -5470, -3646, -3651, -3759, -3653, -3654, 3655
```

Figure 5.5 A sample route information used for frequency-based assignment.

With the information shown in Figure 5.5, the timetable for the time-expanded network can be easily approximated. By estimating the travel time between nodes, the timetable at each stop of the future line can be approximated.

The point of the time approximation to the frequency-based assignment technique is simply dividing the distance referred from the highway network by the given operation speed. Tagging the start time of the bus operation finalizes timetable estimation at each stop. A sample construction of a timetable has been approximated with a current bus route in the UTA. The comparison between the approximated timetable and the published one is presented in Figure 5.6 with bus-line number 472.



stop	estimated schedule	published schedule
Riverdale Park & Ride	6:10	6:10
Layton Hills Mall	6:23	6:23
Kaysville Park & Ride	6:30	6:34
North temple & 300 W	7:03	7:00
402 South & State St.	7:07	7:09
Total journey time	57.7	59.0

Figure 5.6 Comparison of estimated schedule to the published one for Route #472.

Some deviations are observed from the published timetable at each stop; however, overall, the estimated schedule is well fitted to the published one as well as the total

journey time. The same degree of representation of a future transit line to the frequency-based transit assignment technique in the time expanded network can be constructed.

Notation and Problem Formulation

Crowding Cost Function

The asymmetric crowding effect in boarding, and in-vehicle links is modeled with the BPR function with different parameter values. The additional cost due to the crowded vehicle on link a is

$$Cw_a = \xi \cdot \{t_0 \times (1 + \alpha(\frac{v_a}{vc_a})^{\beta}\} \tag{5.1}$$

where α, β, t_0 are positive scalars. vc_a denotes the capacity of the vehicle running link a and ξ is the conversion parameter of the travel time (minutes) equivalent to one dollar value. The crowding cost function is calibrated as $\alpha = 0.13, t_0 = 15, \beta = 6.0$ by the following travel time saving curve (Figure 5.7) developed by Tirachini et al. (2013).

Because the study examines the crowding effects, it is not necessary to apply the strict model in order to avoid the overcrowding. However, we can still expect the capacity constrain effect with the cost function (Schmocker 2006).

With the crowding cost function, the travel cost c_{rs}^{p} of a transit route $p \in P_{rs}$ is the sum of the line segment travel time l_a that follows the published time table with the fare μ and the time cost due to the crowding.

$$c_{rs}^{p} = \sum_{a \in TA} (l_a + Cw_a) \cdot \delta_{rs}^{p,a} + \mu \tag{5.2}$$

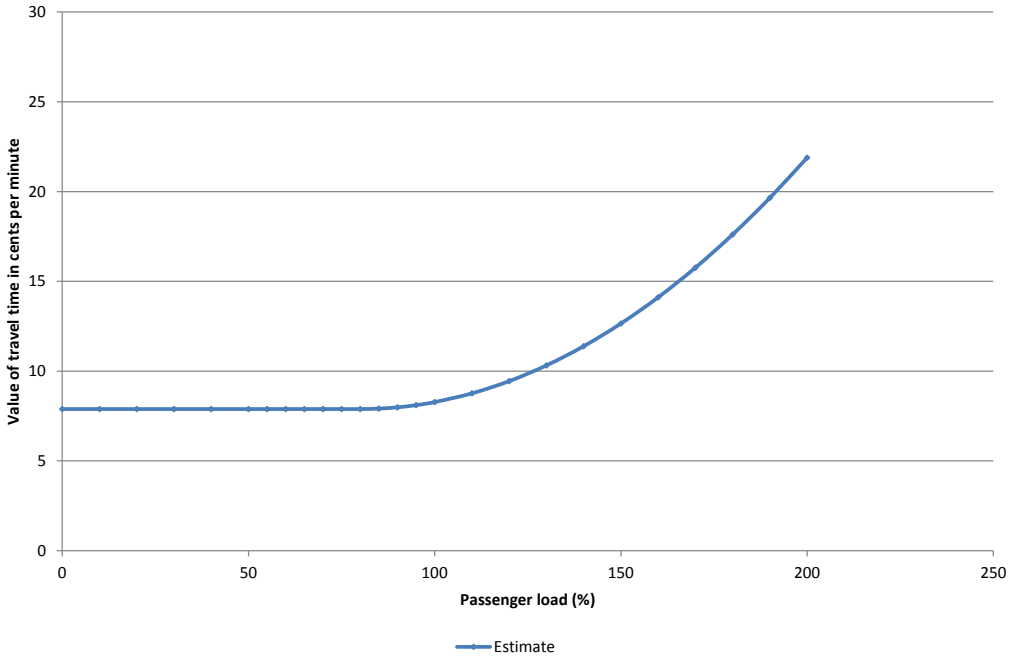


Figure 5.7 Value of travel time savings vs. passenger load in a vehicle.

l includes the waiting and boarding arcs, e.g. the waiting time at a stop as $l(i_m, j) = t(j) - t(i_m)$.

Formulation of the Day-to-Day Demand, Flow, Travel Cost Variation

In order to model the uncertainty of the demand, and corresponding link and route flows, the study, assumes that the transit demand between each O-D pair is a random variable reflecting the day-to-day variation as modeled by Shao et al. (2006), the random variable may be expressed as:

$$Q_{rs} = q_{rs} + \varepsilon_{rs} \quad \forall r \in R, s \in S \tag{5.3}$$

where q_{rs} is the mean demand such that $E[Q_{rs}] = q_{rs}$, and ε_{rs} is the random term $E[\varepsilon_{rs}] = 0.0$. The capital letter represents the random variable while the lower case letter is the deterministic variable. The standard deviation (SD) σ_{rs}^{q} of the transit O-D demand is

$$\sigma_{rs}^{q} = \sqrt{Var(Q_{rs})} \quad \forall r \in R, s \in S \tag{5.4}$$

Thus, the coefficient of variance (CV) of the transit O-D demand can be derived as:

$$cv_{rs} = \frac{\sigma_{rs}^{q}}{q_{rs}} \quad \forall r \in R, s \in S \tag{5.5}$$

The flow conservation equations using the random variables F_{rs}^{p} for route flows along a route p are as follows:

$$Q_{rs} = \sum_{p} F_{rs}^{p} \quad \forall r \in R, s \in S \tag{5.6}$$

$$q_{rs} = \sum_{p} f_{rs}^{p} \quad \forall r \in R, s \in S \tag{5.7}$$

$$\sigma_{rs}^{f,p} = \sqrt{Var(F_{rs}^{p})} = f_{rs}^{p} cv_{rs} \quad \forall p \in P_{rs}, r \in R, s \in S \tag{5.8}$$

where $\sigma_{rs}^{f,p}$, f_{rs}^{p} are the SD and the mean of the route flow of route p.

These equations are based on the assumption that that 1) the CV of the route flow is the same as that of the transit O-D demand and 2) the distribution of the route flows is the same class as the O-D demand distribution, and 3) the route flows are mutually independent.

Consequently, the route-based flow conservation equations yield the link flow V_a.

$$V_a = \sum_{rs} \sum_{p} F_{rs}^{p} \cdot \delta_{rs}^{p,a} \quad \forall a \in TA \tag{5.9}$$

$$v_a = \sum_{rs} \sum_{p} f_{rs}^{p} \cdot \delta_{rs}^{p,a} \quad \forall a \in TA \tag{5.10}$$

$$\sigma_{v}^{a} = \sqrt{Var(V_a)} = \sqrt{\sum_{rs} \sum_{p} Var(F_{rs}^{p}) \cdot \delta_{rs}^{p,a}} = \sqrt{\sum_{rs} \sum_{p} (f_{rs}^{p})^2 (cv_{rs})^2 \cdot \delta_{rs}^{p,a}} \quad \forall a \in TA \tag{5.11}$$

where v_a, σ_{v}^{a} are the mean and the SD of V_a, respectively.

If we assume that the transit O-D demand is described with the normal distribution, then the distribution of the route and link flows can be characterized by the normal distribution. Using this property, the time cost distribution and the route also can be derived.

When the time cost due to crowding and traveling the line segment is expressed by eq. (5.1) the mean and variance of the time cost are expressed as:

$$\mathrm{P}_a = l_a + \xi \cdot t_0 + \xi \cdot t_0 \cdot \frac{\alpha}{(C_a)^{\beta}} \sum_{i=0, i=even}^{\beta} \binom{\beta}{i} (\sigma_v^a)^i (v_a)^{\beta - i} (i-1)!! \tag{5.12}$$

$$(\sigma_{\mathrm{P}}^a)^2 = \left(\xi \cdot t_0 \cdot \frac{\alpha}{(C_a)^{\beta}} \right)^2 \left(\sum_{i=0, i=even}^{2\beta} \binom{2\beta}{i} (\sigma_v^a)^i (v_a)^{2\beta - i} (i-1)!! - \left(\sum_{i=0, i=even}^{\beta} \binom{\beta}{i} (\sigma_v^a)^i (v_a)^{\beta - i} (i-1)!! \right)^2 \right) \tag{5.13}$$

Because the time cost associated with each link follows the normal distribution, the addition of each link component comprising a route yields the following mean $\bar{c}_{rs}^{\,p}$ and the SD $\sigma_{rs}^{p,\mathrm{P}}$ of route cost due to crowding.

$$\bar{c}_{rs}^{\,p} = \sum_{a \in TA} \mathrm{P}_a \cdot \delta_{rs}^{p,a} + \mu \tag{5.14}$$

$$\sigma_{rs}^{p,\mathrm{P}} = \sum_{a \in TA} (\sigma_{\mathrm{P}}^a)^2 \cdot \delta_{rs}^{p,a} \tag{5.15}$$

The derived mean and the SD of the time cost due to crowding on the route level enables us to model the transit users' route choice behavior under uncertainty of the route cost incurred by crowding effect. Using the confidence level α that determines the probability of arriving at the destination within the route travel cost ω_{rs}^{p}, the effective route cost defined on route *p* is the sum of expected travel time and the additional cost due to the uncertain crowding effect.

$$\omega_{rs}^{p} = \bar{c}_{rs}^{p} + z_{\alpha} \cdot \sigma_{rs}^{p,\mathrm{P}} \tag{5.16}$$

where z_{α} is the inverse of standard normal cumulative distribution function at α confidence level. With the defined effective route cost function, the Reliability User Equilibrium (RUE) conditions for the proposed model are

$$\begin{aligned} & f_{rs}^{p}(\omega_{rs}^{p} - \pi_{rs}) = 0 \\ & \omega_{rs}^{p} - \pi_{rs} \geq 0 \end{aligned} \tag{5.17}$$

where π_{rs} denotes the reliable shortest route with the minimum additional cost for the O-D pair connecting r and s.

Solution Algorithm

One of the most important tasks in the solution algorithm for the RUE problem would be finding the reliable shortest route because the reliable shortest route cannot be calculated by the sum of the link cost comprising a route. This implies that the typical shortest path algorithm based on the Bellman's principle of optimality is not appropriate for the proposed model. For that reason, the early studies that formulated the RUE problem utilized the simple network with already enumerated route sets. However, Chen et al. (2011b) paves the way to find the Pareto-optimal (non-dominated) paths with the bi-criterion shortest path algorithm using the Mean-Variance (M-V) dominant condition. This study adopts the algorithm for finding the reliable shortest route. The detail of the M-V dominant condition can be found in Chen et al. (2011b). The flow update is conducted by the double projection algorithm due to the asymmetric feature in the crowding cost function.

The details of the solution algorithm are as follows:

Step 1. Initialization

$k = 0, \beta_k^m = \bar{\beta}, \varepsilon(0 < \varepsilon < 1), \rho(0 < \rho < 1)$

Call the bi-criterion shortest path algorithm for each origin r

Add the reliable shortest route π into the route set for each O-D pair by the user class m, $\pi \in P_{rs}^m$

Assign all trips q_{rs}^m into π.

Update link flows V_a and travel cost P_a.

Step 2. Column generation

Call the bi-criterion shortest path algorithm for each origin r

Remove the unused route in P_{rs}^m

If the newly found shortest route's cost $C_{rs}^{\pi,m}$ is less than any in P_{rs}^m, then add π in to P_{rs}^m

Step 3. Convergence test

If all route sets in each class m satisfies the convergence condition, then stop; otherwise go to Step 4.

Step 4. Flow update for a user class m

Find the appropriate step-size β_k^m

Perform the predictor projection $\bar{F}_k^m = proj_\Omega(F_k^m - \beta_k^m C(F_k^m))$

If $\beta_k^m \geq \varepsilon \left[\left\| F_k^m - \bar{F}_k^m \right\| / \left\| C(F_k^m) - C(\bar{F}_k^m) \right\| \right]$ then reduce β_k^m

$$\min \left\{ \rho \beta_k^m, \varepsilon \frac{\left\| F_k^m - \bar{F}_k^m \right\|}{\left\| C(F_k^m) - C(\bar{F}_k^m) \right\|} \right\}$$

Go to the top of the Step 4.

Step 5 Perform the correct projection

$$F_{k+1}^{m} = proj_{\Omega}(F_{k}^{m} - \beta_{k}^{m} C(\bar{F}_{k}^{m}))$$

Go to Step 2.

Bi-criterion shortest path algorithm

Step 1. Initialization

Generate a route p from an origin to itself, set $\mathrm{P}_{rr}^{p} = 0, (\sigma_{rr}^{p,\mathrm{P}})^2 = 0$ and $\bar{c}_{rs}^{p} = 0$. Add p into label-vector P_{rr} and the list of candidate label SE.

Step 2. Label selection

Take label $p \in P_{ru}$ at node u from SE in FIFO order. If SE is empty, then go to Step 4; otherwise go to Step 3.

Step 3. Route extension

For every outgoing link a of chose node u (v denotes a successor node of u)

Step 3.1 Generate a new label $p \in P_{rv}$. Set

$$\mathrm{P}_{rv}^{p} = \mathrm{P}_{ru}^{p} + \mathrm{P}_{a}, (\sigma_{rv}^{p,\mathrm{P}})^2 = (\sigma_{ru}^{p,\mathrm{P}})^2 + (\sigma_{a}^{\mathrm{P}})^2 \text{ and } \bar{c}_{rv}^{p} = \mathrm{P}_{rv}^{p} + \mathrm{z}_{\alpha}\sigma_{rv}^{p,\mathrm{P}}$$

Step 3.2 if $p \in P_{rv}$ is acyclic, then go to Step 3.3; otherwise scan the next link

Step 3.3 if p is a non-dominated route under (M-V) dominant, then add p into P_{rv}, and Remove all paths M-V dominated by p from P_{rv} and SE.

Go to Step 2.

Step 4. Output the reliable shortest route for each O-D pair. Stop.

Numerical Example

Through the numerical examples, the study investigates the changes incurred by incorporating the crowding variations into the transit assignment. The former part of the numerical example attends the details of the users' behavioral changes due to the crowding effect and its consequences while the latter part aggregates the external effect and examines the details of the cost.

As introduced, the study network for the numerical example is the UTA network covering Weber, Davis, Salt Lake, Tooele, and Utah counties in the state of Utah. Since the network is located in one of the developed countries, it is an appropriate network in terms that people using the UTA service would concern more on the comfort and the quality of the service. The UTA network is not seriously congested, but the study examines the various levels of congestions caused by three hours morning peak (6:00 – 9:00 AM) demand. In the time expanded network, the route choice is strongly associated with the starting time of the demand. In order to consider variations of the time dependent route choice behaviors, the study divides the morning peak demand by five and some aggregate measures are based on the averages of divided time periods. The transit demand obtained from the Wasatch Front Regional Council (WFRC) demand model is divided into two classes: one with 96% confidence level and the others who are indifferent to the in-vehicle crowding, and CV = 0.2 is used. The details of the WFRC model will be discussed in the following chapter that directly dealt the model for the demand maximization by the fare structure reform.

Behavioral Response to the Crowding Effect

In order to avoid crowding, transit users change their travel patterns temporally and spatially, by either travelling slightly earlier or later or by travelling with different route selection (Whelan and Crockett 2009; Sumalee et al. 2009; Ceapa et al. 2012). This section explicitly traces the users' behavioral changes responding to the crowding effect in the time expanded network and sees if the consideration causes issues relevant to the system planning.

The behavioral changes that consider the crowding effect could be classified as the temporal and spatial changes. In the model, the temporal change was observed at the transfer point and at the beginning of the trip. The upper two graphs in Figure 5.8 shows one example of changing the transfer point by the transit users originated from Clear filed to the Lakeview hospital on the time expanded network. After incorporating the crowding effect, the transit users choose to transfer to the Route 455 at State & Main because the users behave in a way that the total travel cost is minimized. Even if the passenger load factor (indicated above the time expanded transit route) is lower than 1.0 for both Route 470 and Route 455, the same crowding cost is faced by the user until the State & Main stop because both vehicles are un-crowded. However, transferring to the already occupied transit vehicle at 1000 N & Main incurs the higher cost than transferring in advance at State & Main.

The other behavioral pattern is the spatial change, taking an alternative route. Incorporating the crowding effect incurs users to take alternative routes to the destination. However, due to the limited supply of the transit services, the available options are varied by the locations.

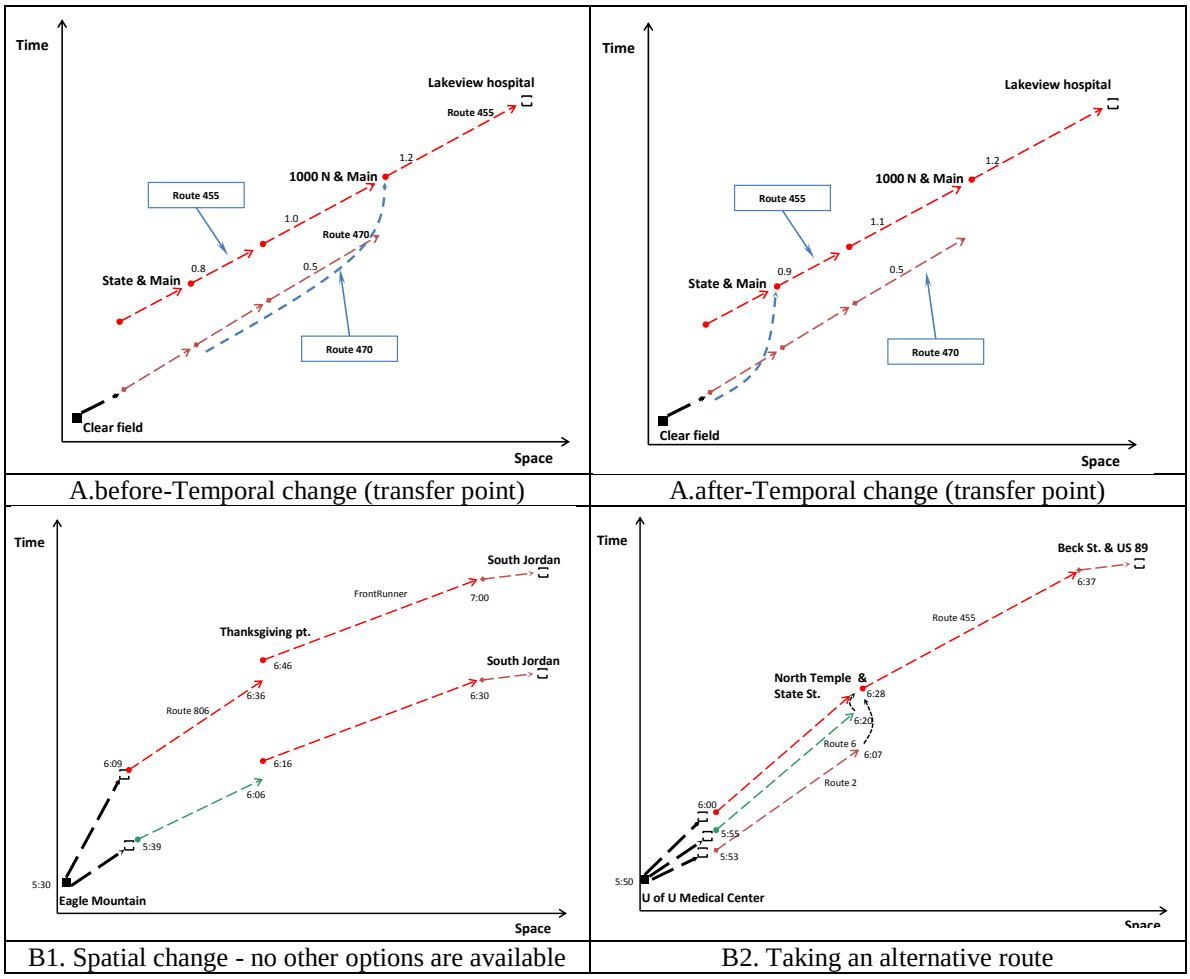


Figure 5.8 Temporal and spatial change on route choice behaviors due to crowding.

These available options are two cases; no other options are available but only one while the other one has several optional routes. As presented in the lower two graphs in Figure 5.8, the users with only one option to the destination have to wait for the next vehicle while the users with the several options have the various route choices to the destination. For the case of pattern B1, the users have to pay the crowding cost unless the cost is higher than waiting for the next vehicle while the users of pattern B2 can take immediately another vehicle.

These changes on the route choice behaviors are aggregated into the each transfer nodes in the time expanded network and are presented as the passenger load factor. The study summarizes the identified behavioral changes through the passenger load factor at the transfer point because two patterns are associated with transfer behaviors, and even if users choose a new route it is likely that the users behave as pattern A in the new route. Figure 5.9 illustrates the top 50 highly congested transfer points according to the passenger load factor before and after the crowd modeling. We can identify that both figures commonly indicate the congestion areas but with different intensities. The congestion points modeled without the crowding effect are densely populated around the core of the Salt Lake City while crowding considered model distributes the congestion along the transit lines.

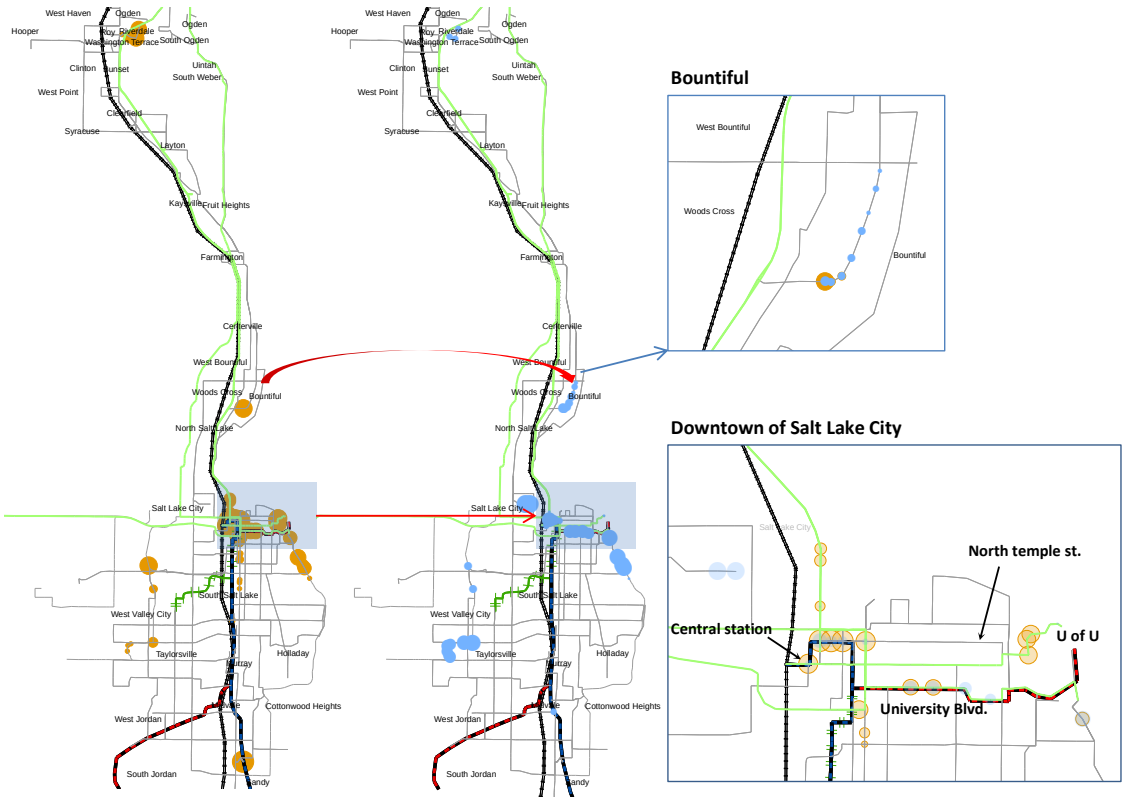


Figure 5.9 Comparison of highly congested area before and after the crowd modeling.

A closer look at the Northern Salt Lake City area where the downtown and the major business districts are located is shown in the figure. Overall, the size of the circles is smaller than that of the model without the crowding effect, and the circles are rather scattered. In the case of the corridor connecting the Salt Lake Central station and the University of Utah, the circles are crowded around the central station and the University area when the crowding effect are not considered, while users considering the crowding effect avoid those areas and incur the lower congestions than before. Spreading the crowding effect along the transit line is explicitly observed in the Bountiful area too.

The model does not exactly represent each individual's behavior; rather the study focuses on its aggregated effect in the network level. The network level analysis identified that the individual behavioral change tends to spread the concentrated crowding congestions over the wider range of the transit network. The implication of the results has the same way of Tirachini et al. (2013)'s warning on the overestimation of demand without the explicit consideration for crowding effect. While their study limited on identifying the overestimation of demand, this study's investigation on the explicit behavioral change goes one step further for identifying the overestimated transfer activities. Be noted that the overestimation is not based on the overestimated demand. The finding regarding the aggregated behavioral changes reflecting the effect of crowding potentially has significant implication for activity estimation of proposed public transport enhancements especially the facility designs (e.g. transfer passageway, bus stops). Figure 5.9 suggests that if demand for the proposed facility is estimated without explicit consideration of crowding, it is likely that the congestion level is overestimated, consequently planning for the facility design will be bigger than the actual system uses.

In the next section, the study attends the details of the cost and the variability caused by the users' attitudes in the network level.

Analysis on Crowding Cost

Before investigating the variation of the uncertainty model due to the randomness of the crowding effect, the study calculates the crowding cost without the uncertainty modeling in order to get the general overview due to the crowding effect for the UTA network.

The following Table 5.1 indicates the crowding cost for each major group of O-D pairs headed to the central business districts of Salt Lake City.

Table 5.1 Additional cost without uncertainty modeling for each major group of origins.

Group of origins	additional cost without uncertainty modeling		equivalent passenger load
	dollar/hr	cents/min	
Ogden	5.36	8.93	1.10
Layton	5.57	9.28	1.15
Bountiful	6.96	11.60	1.31
West Valley city	6.24	10.40	1.29
South Jordan	5.53	9.22	1.18
Murray	6.91	11.52	1.30
Provo	5.38	8.97	1.09

Overall, the crowding cost in the table is higher than the un-crowded situation, but it rarely touches the value equivalent to 1.5 passenger load. According to Figure 5.7 that graphed valuation for the reduced crowd, the un-crowded situation is valued as 7.9cents/min, which is equivalent to 4.74 dollar/hr. This means that when a vehicle runs the entire route without any standees the valuation of crowding for the trip is 4.74dollar/hr. Considering that the crushing passenger load is around 1.7 ~ 1.8 (Critical Transit 2014), the additional cost due to the crowd modeling indirectly indicates the

crowding condition experienced by the transit demand from the each group of origins. We can identify that each end of South (Provo) and North (Ogden) of the UTA network has the relatively low additional cost because users will have enough seats. In the meantime, as the vehicles come to close to Salt Lake City, the passenger load gradually increases, consequently, the trips that begin their journey in the middle of the corridor (red dotted line) will have the limited space from the very beginning of their trips. This pattern is represented in the growing additional cost attached to Ogden, Layton, Bountiful in order.

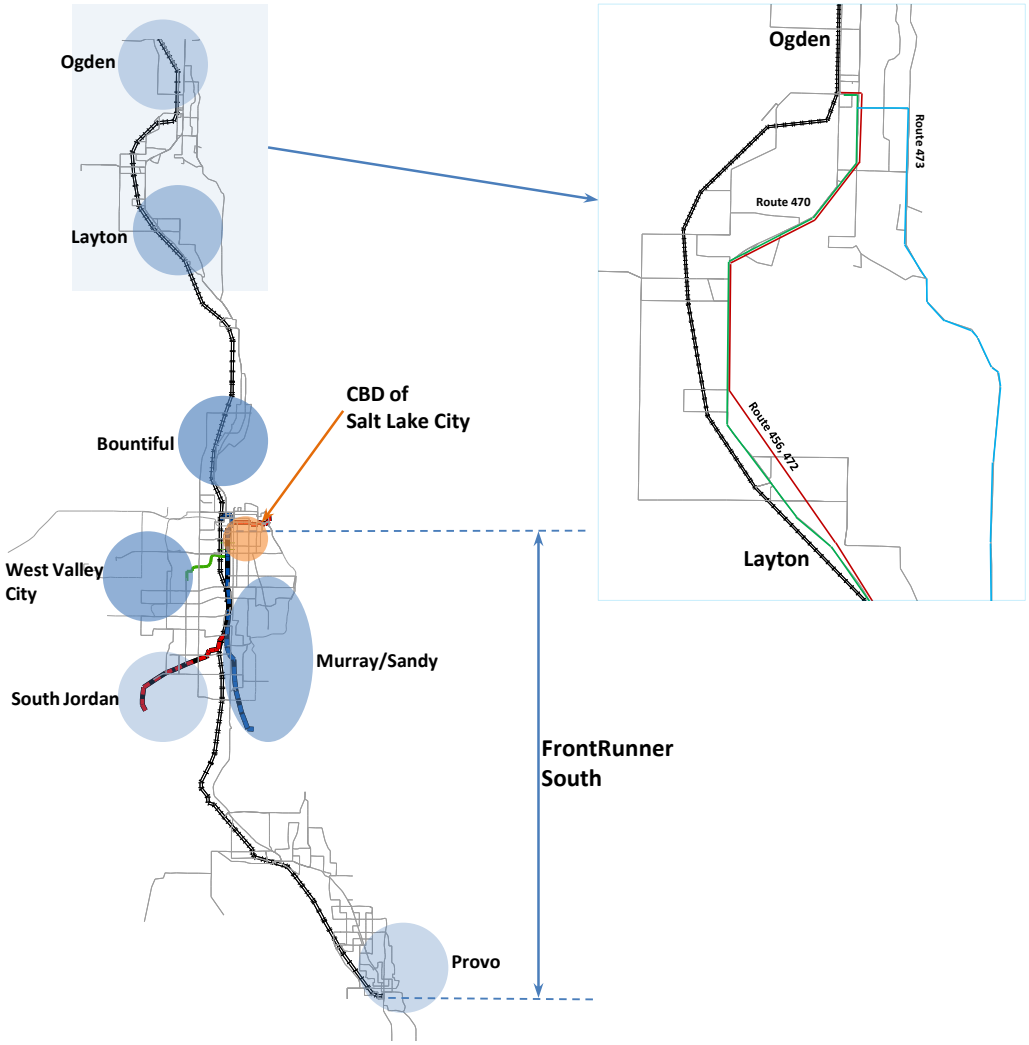


Figure 5.10 Illustration of the congestion level for each major group of origins.

The couple of operations of bus lines (Route 461, 471, 463) that begin from Bountiful and Centerville during peak hours are designed to support the demands. Similarly, the increasing cost pattern is observed in the north bound from Provo, Murray, West Valley City in order. Namely, the cities adjacent to the downtown of Salt Lake City such as West Valley City, Murray, and Sandy have the relatively high congestion level than the sub-urban areas.

The previous paragraphs provide the general idea of the congestions for each major group of origins. When the uncertainty is considered in the RUE the level of the expected crowding cost is determined according to the users' confidence levels. The following section examines the variation of the cost with different attitudes.

First, the total travel cost has been broken down for the detailed investigation on the cost composition by the different confidence level. Table 5.2 indicates each component of the total travel cost for the demand originated from Ogden and Layton. For both cases, the users indifferent to crowding choose the route that cost more than the highly concerned users. However, the difference of the total cost is not significant than that of the demand from Layton. The investigation on the route choice identifies that the most demands from Ogden split to FrontRunner and Expressbuses. Crowding experienced by the users is very low in both of the transit vehicles, but the Expressbus users experience slightly more congestions as the buses get to its final destination. On the other hand, the demand from Layton has the significantly different total travel cost. In this case, one Expressbus is excluded (Route 473). Thus, the passenger load of FrontRunner is lower than that of Expressbuses. This incurs the different route choice between user classes. The route selected by the indifferent users has the higher crowding

cost, the in-vehicle time and the SD while the highly concerned users choose the route that minimizes the total cost despite of the high fare level.

Table 5.2 Details of the total travel cost in the crowding considered model.

Component of travel cost	Ogden		Layton	
	indifferent	high confidence	indifferent	high confidence
total travel cost	159.07	156.73	120.96	95.85
Expected crowding cost	43.56	42.80	41.52	25.19
fare	5.25	5.25	2.25	3.33
in-vehicle time	80.12	78.25	64.28	47.30
SD	1.25	0.18	1.92	0.56

As the users' concerns on crowding increase, it is expected that these changes would be more distinguished. The study examines this in terms of the revenue increase. The following Table 5.3 presents the percent revenue increase to the revenue from one class model where the whole demand does not consider crowding. As one might expect, the demand with the high confidence and the revenue are in direct proportion because the users with high confidence are willing to pay more as identified above. However, the scales of the ratio are dependent on the service availability for those users. For the users who originated from Ogden and Provo, which are located at each end of FrontRunner, the revenue increase is more distinguished than other areas. Because the users in those areas can easily switch to FrontRunner, the demand increase is transferred to the revenue increase. On the other hand, the users in West Valley City, Bountiful, Murray/Sandy do not seem to have the strong incentive to switch to FrontRunner. Geographically, the FrontRunner's route is not attractive to the users from West Valley City headed to the downtown of Salt Lake City and closeness of the users in Bountiful, Murray/Sandy to the downtown contributes to the low increase of the revenue.

Table 5.3 Revenue increase ratio due to different demand composition.

Group of origins	demand compositions (high confidence: indifferent)		
	3:7	5:5	7:3
Ogden	1.8%	3.5%	4.9%
Bountiful	0.2%	0.3%	1.7%
West Valley City	0.1%	0.2%	0.2%
Murray/Sandy	0.6%	0.7%	1.4%
Provo	1.5%	3.5%	3.5%

This section intended to show that the revenue could be affected by the different classes of users. In case of UTA, the fare increase is not significant. However, this depends on the geographic distribution of demand, service availability, and the fare level of the premium service.

In addition, the study examines the impact of the extension project of FrontRunner South, which expanded the former southern terminus from Salt Lake Central to Provo Station at the end of year 2012. Because FrontRunner is the reliable commuter rail service, the service would be preferred by the users with the high confidence level. The study investigates how users are benefited from the new project in terms of the expected crowding cost.

Table 5.4 presents the comparison of the expected crowding cost before and after the extension project. Overall, the expected crowding cost is reduced after the extension. Because the new service line is relatively reliable and has the sufficient capacity, the extension provided the attractive options for the highly confidence users in the Southern part of the UTA network.

It is noteworthy that switching to FrontRunner not only provides a more comfort riding experience for highly confidence users but also reduces discomfort of indifferent users. When a new public transportation enhancement project is evaluated by the travel

time and cost only, this reduced crowding would not be assessed. The results show that considering crowding with multi-class transit users enables to identify the benefit from the service supply increase for different user classes. Reduced crowding for both classes also implies that additional consideration of crowding to the two dominant travel disutility factors (time and cost) not always cause the increase of disutility that end up with underestimation of the demand.

Table 5.4 Comparison of the expected crowding cost before and after the extension.

group of origins	before		after		% decrease	
	high confidence	indifferent	high confidence	indifferent	high confidence	indifferent
South Jordan	41.07	50.28	37.34	48.98	9.1%	2.6%
Murray/Sandy	71.99	76.37	65.45	73.84	10.0%	3.4%
Provo	49.83	56.50	43.94	52.41	13.4%	7.8%

Summary

As mentioned in the introduction, the crowding effect is easily noticeable by the users and it gains attention by academics, consultancies and transit agencies. Following the trend, the study modeled the daily variation of the in-vehicle crowding in the time expanded network representing the real large scale transit system. Crowding considered transit assignment for the real network could be modeled and implemented by constructing the crowding cost function that follows the valuation of crowding (Tirachini et al. 2013) and by using the reliable shortest path finding method introduced by Chen et al. 2011b.

In the numerical experiment, the impact caused by crowd modeling is analyzed in terms of each individual's behavioral response and the expected crowding cost for

different user classes. From the results, the insight on the consequences from the crowd modeling could be identified.

The analysis on the aggregated behavioral changes showed that each individual behavior tends to spread the concentrated crowding congestions over the wider range of the network. The results identified the possibility of the overestimated congestions without the crowding effects, which is strongly related to appropriate public transport facility designs.

The cost component analysis identified that different attitudes to crowding could result in the revenue increase for transit agencies. Because the users with the high confidence level want to use the route with the less expected crowding cost despite of the higher fare levels, the increase of revenue was observed. However, this depends on the spatial distribution of demand and the availability and the fare levels of the transit services. In addition, consideration of the multi user classes enables to identify the discomfort reduction by the transit service supply project for multiple user classes simultaneously. This showed that addition of the disutility factor, crowding, not always causes the increase of disutility from the transit uses.

In this section, the non-additivity of the reliable shortest route is created by the uncertainty modeling of the crowding effect, but in the next section, the non-additivity is directly modeled in the distance-based fare structure.

CHAPTER 6

EXAMINING THE POTENTIAL OF THE TRANSITION INTO DISTANCE-BASED FARE STRUCTURE CONSIDERING USERS' BEHAVIORS IN A TIME-EXPANDED NETWORK

In this chapter, one instance of the asymmetric interactions observed in the fare structure, especially in the distance-based fare structure, is modeled to provide insight on the fare policy approach for transit agencies. The objective is to analyze potential of the transition from a flat fare system to a distance-based fare structure. As mentioned in the literature review, the existing approaches for modeling the distance-based fare structure are mostly under profit maximization, and the route choice behaviors were not considered.

Unlike the previous studies, this study aims for examining potentials of the fare structure transition through modeling users' route choices in the large scale transportation network under demand maximization. Profit maximization is appropriate for agencies that opt for profit making and for private companies but are not desirable for most transit agencies inclined to social objectives. This study views the potential of the distance-based fare structure through demand maximization. In terms of the modeling perspective, the investigation for the fare structure reform is based on the effective modeling for the cost and the users' behaviors responding to the fare structure change with the time expanded network. Consideration for the users' behaviors was conducted by Zhou and Lam (2001), but the optimal fare is estimated based on the transit line not the distance-based fare structure.

The application of this chapter can be characterized by following features: 1) The potential of the distance-based fare structure is examined from demand maximization; 2) The distance-based fare structure is modeled on the time expanded network; 3) The interactions between the fare policy maker and the users' behaviors in the time expanded network is effectively modeled through the bi-level programming.

The first step in examining the potential of a fare structure change is to find the optimal fare levels when the fare policy is changed to the distance-based fare structure. Finding the optimized fare levels is based on the effective representation scheme for the fare structure and the riders' behaviors by using the time expanded network and the bi-level programming. The details of the methodology and the fare model are addressed in the following sections of the chapter.

The next section addresses the motivation for modeling the distance-based fare structure. In addition to the modeling perspective, the research addresses the practical perspective of incorporating the distance-based fare structure. The research continues to describe the effectiveness of the time-expanded network for modeling the distance-based fare structure. Explanation is given of how the timetable for each stop along the transit line effectively measures the distance-based fare for the time-expanded network. Then, before addressing the methodology for the optimization process, it is necessary to set up the fare model, which works as the objective function in the optimization program. This research uses the demand maximization model, namely ridership maximization. The methodology section describes the interaction between the fare policy maker and the users in the network via bi-level programming. The final section is about the proposed methodology application into the UTA system and its results.

Toward the Distance-Based Fare Structure

Public transportation is a part of the everyday lives for people living all over the world. Currently, a number of transit systems are in operation in order to provide people with basic mobility, alleviate congestion, protect the environment, and to save energy.

Efforts for maximizing those benefits through designing the system to be competitive in the transportation market have long been an interesting topic for researchers. The financial performance is a major issue for transit agencies. Raising the fare for increasing the financial independency would result in sacrificing ridership, which is another important indicator evaluating the system's performance. Currently, 98% (NTD 2013) of transit agencies in U.S. are suffering from the financial deficits for the operating expenses.

The transit agency desires to simplify fare collection makes generating revenue difficult. According to the Transit Cooperative Research Program (TCRP 2003), the percentage of agencies using the differentiated fare structure has declined since 1994. This implies that "transit agencies by and large simply do not seem willing to address the complexities associated with designing, implementing, administrating, and marketing." This long-term tendency to abandon differentiated pricing in favor of flat fares causes insufficient revenue via the fare-box. The differentiated fare structure can alleviate financial difficulties, unfortunately, the flat fare trend seems to be an industry norm (Hickey 2005; Chien and Tsai 2007; Tsai et al. 2008).

In the recent past (in 1960's), the differentiated fare system was used in most cities (Cervero and Wasch 1982). The method was gradually replaced by the flat fare in order to provide low-cost transportation to lower income people and the simplicity of the

flat fare collection. Almost all transit agencies in U.S changed to the flat fare system (Cervero and Wasch 1982). Later, this change became the primary cause for the 1980's transit fiscal crisis.

Cervero (1981) and Cervero and Wasch (1982) attributed the transit agencies financial difficulties to the insensitivity of the flat fare and the suburbanization in 1960's and 1970's. They observed that the suburbanization of U.S. at this time exacerbated an unhealthy fiscal condition due to many transit operators expanding their service coverage to outlying areas while keeping the insensitive fare structure. They supported this observation with the comparison of the average mileage between the individual bus route and the total number of buses in the fleet. From 1960 to 1974, the average mileage covered by individual bus routes more than doubled while the total bus mileage actually declined (Sale and Green 1979).

However, the difficult-to-implement differentiated fare structure of the past has become more attractive with recent development of fare collection technologies. In particular, the increasingly widespread usage of smart cards (TCRP 2003) has made the distance-based fare structure more practical. The smart card's ability to be used on multiple operators easily enables the distance-based fare structure by consolidating all transit systems into one and collecting fare based on actual passenger distance traveled. Generally, the fare structure is used by rail systems because the total in-vehicle distance determined by the strictly controlled boarding and the exiting points that travelers purchase before they ride. Advancements in global positioning technologies and fare media make possible the implementation of total in-vehicle distance at the fare level for

other transit modes. These days, the structure is being expanded to other transit vehicle types, such as rail and bus modes.

The fare structure underwent changes for years and there was a time when the differentiated fare structure was widely implemented. Seemingly, once the agency adopts the flat fare due to its simplicity, regardless of the financial deficit, they become unwilling to change to any alternatives. The general perception of the transit system as a public good attributed to the propagation of the flat fare. However, recent fare collection technological developments offer the opportunity to reconsider these practices. Successful examples of distance-based fare structure implementation have been seen and provide agencies motivation to reconsider the practicality of this system.

This research investigates the positive ridership impacts for the UTA if they undergo a fare structure change by considering the distance-based fare structure (Lee 2013). The investigation is based on the effective cost modeling and user response modeling to the fare structure change with the time-expanded network.

Effectiveness of Time Expanded Network for the Distance-Based Fare Structure

In order to model the distance-based fare structure, selection of the appropriate modeling technique that effectively traces passengers' route choices is important. Since the fare is based on the total passenger in-vehicle distance the model that closely describes the passengers' flows is needed for representing the fare structures. Due to easy access to route information, this research uses the time-expanded network for transit operation. In the time-expanded network, transit users' routes are traceable in the same manner as the conventional highway network.

In the time-expanded network, all possible ways of the passenger flows are created with virtual links and nodes. In Figure 6.1, the transit lines (upper graph) connecting the origin and the destination pair is broken into the run-based representation scheme (lower graph).

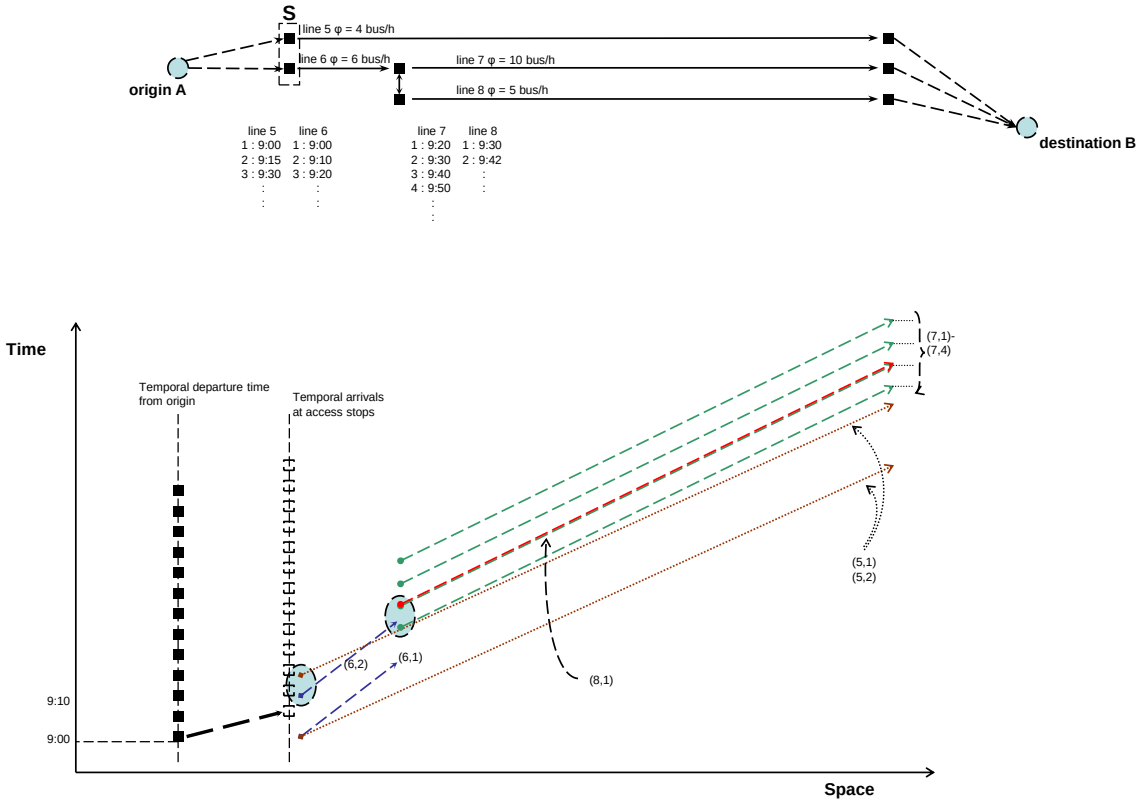


Figure 6.1 Virtual links and nodes in the time expanded network.

Each run of the transit line drawn by the dotted lines and the temporal arrivals of the runs (end of each arrow) build up all possible ways to the destinations. Under the structure, the network-handling scheme becomes very similar to the conventional highway network structure. A route comprising the sequence of links is replaced by the sequence of runs in the time-expanded network. Tracing passengers' route choices is intuitive by following the sequence of virtual links. For example, in Figure 6.1, users

departed from the origin at 9:00 AM encounter many choices of runs of vehicles as the number of virtual nodes after the temporal arrivals of the thick dotted lines. However, they would choose the first or the second leaving vehicle (indicated with the circles) in order to finish their journey with the time minimization manner. In this case, if the passenger takes the first arriving vehicles, which is the second run of line 6, then he/she misses the line 7 that departs at 9:20 AM. Taking the second run of bus 5 that connects directly to the destination B is the shortest route to the destination according to the space-time graph.

The time-expanded network is also advantageous for modeling the intermittent transit operations, which only run for specific time periods. For example, The UTA system has a few buses running a couple of times during the morning peak hours in order to support the demand heading to the downtown Salt Lake City. In the time expanded network scheme, simply adding virtual links and nodes effectively account for those operations.

Methodology

Bi-Level Programming

This study utilizes the bi-level programming for representing the relation between the fare policy maker and the passengers. The programming method effectively models the leader-follower relationship in which the policy maker is the leader and the passengers are the followers. It is assumed that the policy maker influences the passengers with the fare levels and cannot control the passengers' route choice in the network. Once the fare level has been set, passengers will choose the lowest-cost route to

their destination. The interactions between the policy maker and the passengers can be represented as follows:

$$\begin{aligned} &\max\ Z(\mu, d(\mu)) \\ &s.t. \\ &G(\mu, d(\mu)) \le 0 \end{aligned} \tag{6.1}$$

where the objective function of the upper level Z is the function of μ and the decision vector from the lower level $d(\mu)$, and is constrained by G while in the lower level, the individuals want to minimize their generalized cost to the destination.

$$\min\ c(\mu) \tag{6.2}$$

Under the distance-based fare structure, the fare μ consists of more than one variable that includes the base fare b, mileage fare m and additional variables representing the surcharge for the premier services. μ is defined by

$$\mu_{rs} = b + m \cdot s_{rs} + \{\max_{l \in L} \psi_l \cdot \delta_{rs}^{p_l}\} \tag{6.3}$$

where ψ_l is the additional fare for using premier service l.

With the given fare level μ, the lower level finds the route with the minimum generalized user costs with various time components. The generalized user cost includes various time components from the time expanded network and the fare.

$$c_{rs} = \varphi_{iv} iv_{rs} + \varphi_{iw} iw_{rs} + \varphi_{xf} xf_{rs} + \xi \mu_{rs} + \varphi_{xfw} xfw_{rs} \tag{6.4}$$

where

c_{rs} : generalized user cost

iv_{rs} : in-vehicle transit time

iw_{rs} : initial waiting time

xf_{rs} : number of transfers

xfw_{rs} : transfer waiting time

$\varphi_{iv}, \varphi_{iw}, \varphi_{xf}, \varphi_{xfw}$: weights on the time component, which are dimensionless

ξ : measure of travel time equivalent to one dollar of fare paid

Genetic Algorithm (GA) with Alpha Constrained Method

The decision variables to be optimized in the distance-based structure are constrained by budgets. In order to efficiently solve the constrained multi-variable optimization problem in the bi-level programming, the study adopted the genetic process for the upper level program. Since the Genetic Algorithm (GA) is directly applicable to only unconstrained optimization, it is necessary to use additional schemes that will keep solutions in the feasible region. Among the various approaches for the constrained genetic algorithm, this study adopts alpha constrained method (Takahama and Sakai 2004) in order to dynamically penalize the infeasible solutions, which are ordered by the alpha level comparison.

In the alpha constrained method, the satisfaction level of a constraint is introduced to indicate how much a candidate solution satisfies the constraints. A solution is prioritized by the satisfaction level of a constraint over the objective function value because ensuring the feasible solution is more important than the achievement of the objective. Moreover, as the algorithm iterates, the satisfaction level of a constraint gradually increases by the dynamically adjusted alpha. The subsequent description of the

α constrained method for the GA process begins with the general formulation of the constrained optimization problem.

Alpha constrained method

The general formulation of the constrained optimization problem may be expressed as follows:

$$\begin{aligned} &\min\ z(X) \\ &s.t. \\ &g_i(X) \le 0,\ \ i = 1, 2, ..., p \\ &h_j(X) = 0,\ \ j = 1, 2, ..., q \\ &l_i \le x_i \le u_i,\ \ i = 1, 2, ..., n \end{aligned} \tag{6.5}$$

where

$X = (x_1, x_2, ..., x_n)$: n dimensional vector

$z(X)$: objective function

$g_i(X)$: p in-equality constraints

$h_j(X)$: q equality constraints

l_i, u_i : lower and upper bounds of x_i

A satisfaction level of constraints $\sigma(X)$ is defined in order to measure how much a candidate solution X satisfies the constraints.

$$\begin{pmatrix} \sigma(X) = 1, & \text{if } g_i(X) \le 0, h_j(X) = 0\ \ \forall i, j \\ 0 \le \sigma(X) \le 1, & \text{otherwise} \end{pmatrix} \tag{6.6}$$

More specifically, the satisfaction level may be formulated as:

$$\sigma_{g_i}(X)=\begin{pmatrix}1 & if\ g_i(X)\le 0\\ 1-\dfrac{g_i(X)}{b_i} & if\ 0\le g_i(X)\le b_i\\ 0, & otherwise\end{pmatrix} \tag{6.7}$$

$$\sigma_{h_j}(X)=\begin{pmatrix}1-\dfrac{|h_j(X)|}{b_j} & if\ |h_j(X)|\le b_j\\ 0, & otherwise\end{pmatrix} \tag{6.8}$$

where

b_i, b_j are proper positive fixed values. Each satisfaction level is combined into one representative level such as the minimization operator:

$$\sigma(X)=\min_{i,j}\{\sigma_{g_i}(X),\sigma_{h_j}(X)\} \tag{6.9}$$

Once the satisfaction function has been set, an order relation is defined where the feasibility precedes the objective function value as follows:

Let $f_1(X_1)$, $f_2(X_2)$ and σ_1, σ_2 be the function values and the satisfaction levels at solutions X_1, X_2, respectively.

$$(f(X_1),\sigma_1)\le_\alpha (f(X_2),\sigma_2)\Leftrightarrow\begin{pmatrix}f(X_1)\le f(X_2), & if\ \sigma_1,\sigma_2\ge\alpha\\ f(X_1)\le f(X_2), & if\ \sigma_1=\sigma_2\\ \sigma_1>\sigma_2 & otherwise\end{pmatrix} \tag{6.10}$$

In the example case of $\alpha=0$,alpha level comparison($\le_\alpha$) is equal to the ordinal comparison for the objective function values. This order relation that ranks the infeasible solutions low penalizes the infeasible solutions.

In addition, the value of alpha also gradually increases as the algorithm iterates in order to narrow down the feasible solutions with higher satisfaction levels. However, in

the case when the feasible region is extremely narrow, such as equality constraints, it is difficult to determine the feasible region and it is highly likely that most candidate solutions gather in the neighborhood of the point with high satisfaction levels. To avoid this situation, it is necessary to properly control the alpha level so that a wide range of a feasible region is covered. The basic idea is that increasing the α level as slowly as possible in order to ensure enough individual satisfying α levels. However, the time when the alpha level time control is limited.

$$\alpha(0) = \frac{1}{2}\left(\max_{x_k} \sigma(x_k) + \frac{1}{N}\sum_{k=1}^{N} \sigma(x_k) \right)$$

$$\alpha(t) = \begin{pmatrix} 1-[1-\alpha(0)](1-\frac{2t}{T})^2, & if\ 0 < t < \frac{T}{2} \\ 1 & if\ t \geq \frac{T}{2} \end{pmatrix} \tag{6.11}$$

where

T : maximum number of iterations

t : iteration number

N : number of population

Genetic process

With the α constrained method, the genetic process for the constrained maximization problem begins with decoding the decision variables into several binary codes in order to get a suitable "string" structure akin to biological chromosomes. This genetic representation is processed through reproduction, crossover, and mutation operators iteratively until a set of optimal solutions is found. Readers are advised to refer

to Goldberg (1989), Deb (2001), and Coello et al. (2002) for more details on each step of the GA. Below provides the general description for each step of the GA method with the special features for penalizing the infeasible solution.

The first step is to randomly produce the initial population set according to the lower and the upper levels of each decision variables. Then, reproduction selects good parent solutions to form a mating pool. A simplified concept of reproduction is to select parent solutions from the existing pool with above-average fitness values and copy them into a mating pool. These above-average copied solutions act as parents for the next generation. In this study, this comparison level provides priority to the feasible solution. The selection probability ranked with level comparison uses the linear selection strategy (Baker 1984; Bäck and Hoffmeister 1991).

$$p_k = \frac{1}{N}(\eta^+ - [\eta^+ - \eta^-]\frac{r_k - 1}{N - 1}) \tag{6.12}$$

where $\eta^- = 2 - \eta^+$ and $\eta^+ = [1.0, 2.0]$ is the maximum expected value, which specifies how many times the best individual is selected more than the median individual.

In the crossover, two randomly selected parent solutions (chromosomes) exchange their gene information expecting the offspring's improved fitness created by the crossover. Not all mating pool solutions found in the search process are involved in the crossover due to the effects being either positive or negative. The following Figure 6.2 illustrates the crossover operation.

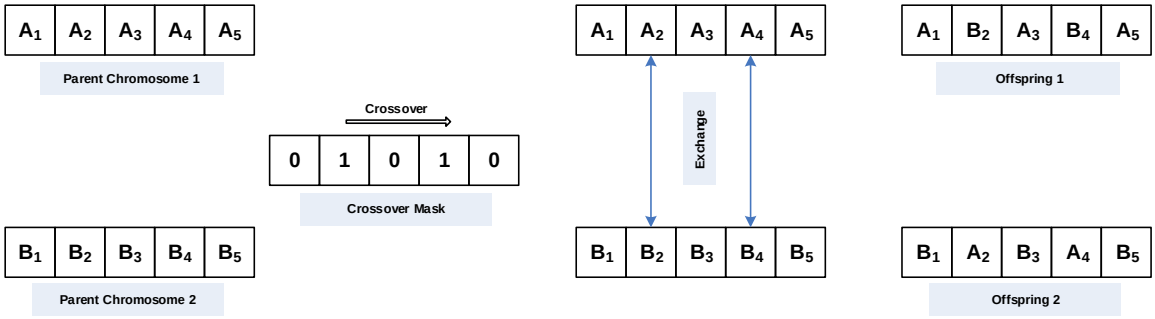


Figure 6.2 Crossover operation.

In the crossover step, it is possible that some solutions could become infeasible after going through the crossover process. However, the satisfaction level maximization is given top priority and is naturally optimized in the direction where the constraints are satisfied. Therefore it is not necessary to prevent infeasibility in the crossover step (Takahama and Sakai 2004).

Finally, the mutation is intended to create a new solution adjacent to a current solution by introducing a small change in the gene information. Despite the significant changes to the gene's information making broader solution searches possible, the mutation is intended to localize the search around the current solution. The most minuscule gene string may be altered with small probability.

Under the described methodologies, the solutions procedure can be outlined as follows (Figure 6.3). After determining the fare structure for the distance-based fare, arbitrarily created initial set of decision variables is inserted into the lower level program that finds the optimal transit route in the time expanded network. Tracing the identified route for each O-D pair enables the estimation of the initial wait time, in-vehicle time, transfer wait time and the number of transfers. The time components and fare levels are used to calculate the demand in the upper level program. The fitness of each fare level is

used for qualifying into the next step. This includes mutation and crossover unless the algorithm stops due to the predefined stopping criterion. The GA usually stops when the pre-determined number of iterations is over or the enhancement of the solution fall into the specified range.

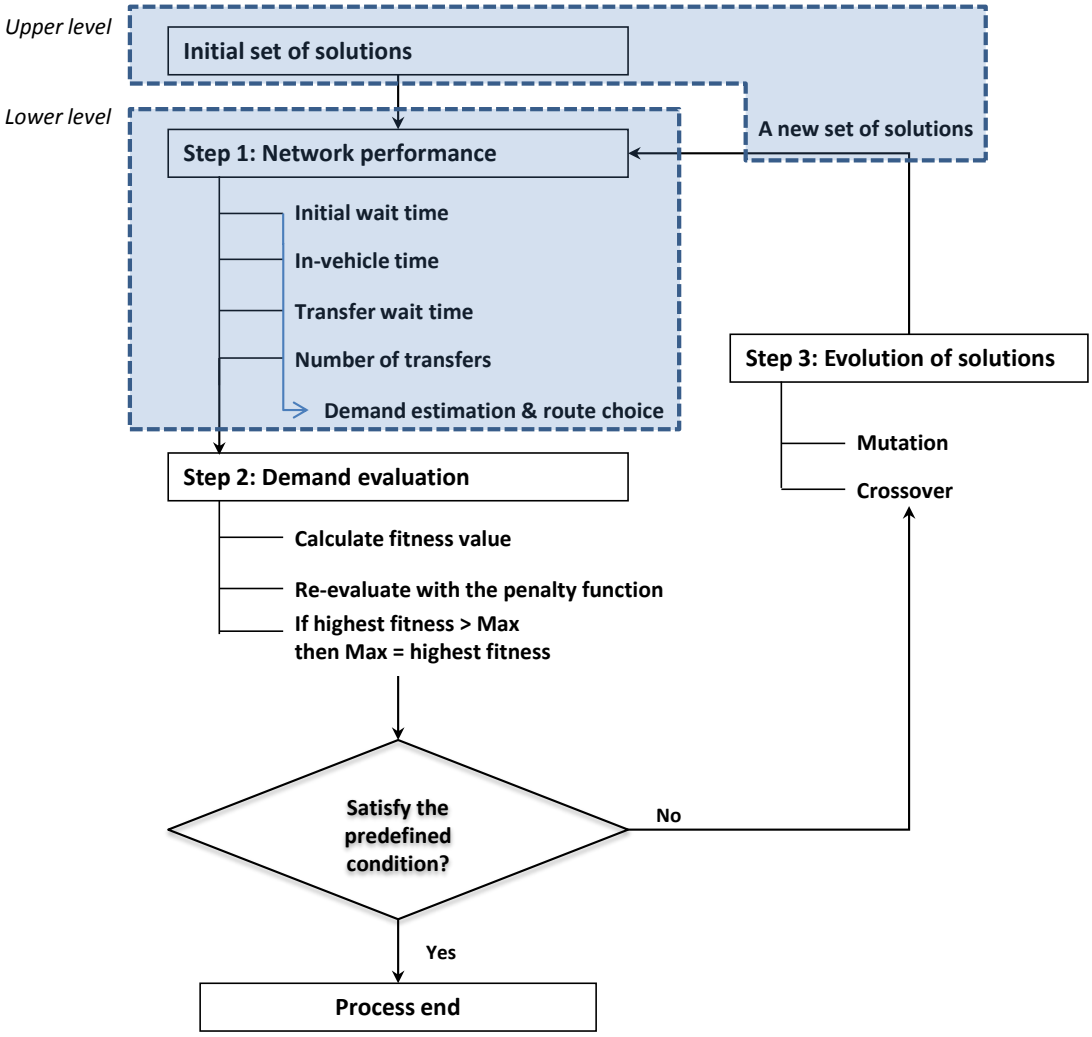


Figure 6.3 Outline of the proposed bi-level programming.

Input

This analysis is based on the peak demand and the 2012 transit network. The network and the coefficients for the utilities for transit and passenger cars are based on the planning model for the Wasatch Front Regional Council (WFRC). In the time-expanded network, the route choice is strongly associated with the starting time of the demand. In order to consider variations of the time dependent route choice behaviors, the study divides the morning peak demand by five and averages each time component for the analysis.

Demand function

The demand function is a key feature of the fare planning process. It estimates the transit demand that responds to the fare changes. The demand function employed in the WFRC model is the nested logit function that counts for the choice probability that partially avoids the IIA (Independence of Irrelevant Alternatives) properties of the traditional logit model. The model is sophisticated enough to divide the transit demand into drive-to-transit M^D and walk-to-transit M^W. The nesting structure with nesting coefficient λ of the WFRC model is illustrated in Figure 6.4.

The associated utilities for each O-D pair for choosing a transit mode accounts for various time components such as the in-vehicle time, initial wait time, number of transfers, transfer wait time, and fares.

$$U_{rs}^{m} = V_{rs}^{m} + e_{rs}^{m} \tag{6.13}$$

where

$$V_{rs}^{m} = \theta_{iv} iv_{rs}^{m} + \theta_{iw} iw_{rs}^{m} + \theta_{xf} Xf_{rs}^{m} + \theta_{\pi} \mu_{rs}^{m} + \theta_{xfw} Xfw_{rs}^{m} + \theta_{prk} prk_{rs}^{m} + \theta_{opr} opr_{rs}^{m} + \theta_{toll} toll_{rs}^{m}$$

V_{rs}^{m} : utilities for choosing a mode in the nesting structure between O-D pair r and s

θ_{iv} : coefficient of in-vehicle time

θ_{iw} : coefficient of initial waiting time

θ_{xf} : coefficient of number of transfers

θ_{μ} : coefficient of fare

θ_{xfw} : coefficient of transfer waiting time

θ_{prk} : coefficient of parking cost

θ_{opr} : coefficient of operating cost

θ_{toll} : coefficient of toll fare

e_{rs}^{m} : unobservable error term

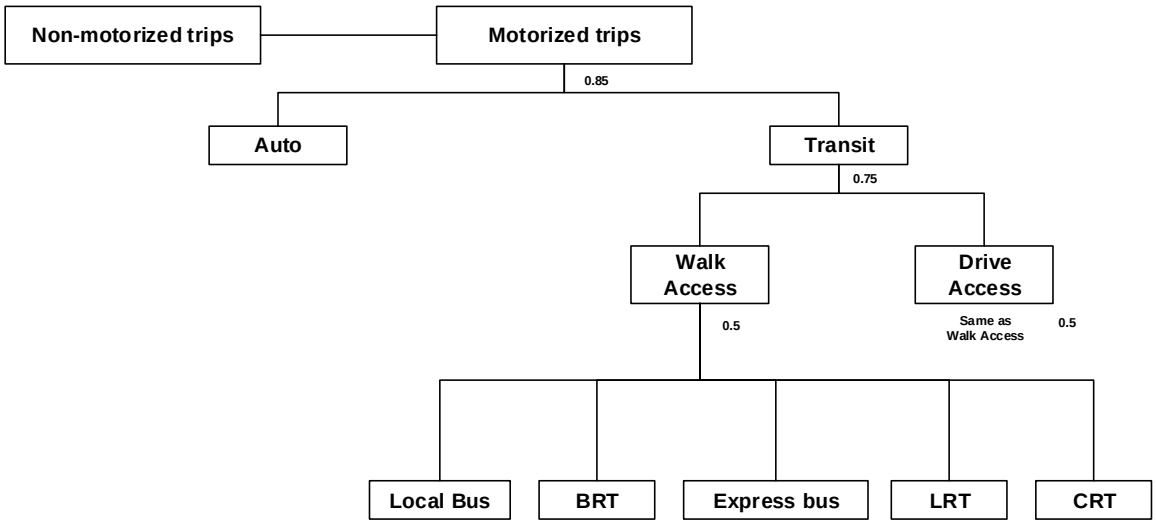


Figure 6.4 Nesting structure of the WFRC demand model.

Under the specification, the expected demand d_{rs}^{m} for choosing mode $m \subset M_{rs}^{W}$ is the product of the total demand q_{rs} and the probability p_{rs}^{m}.

$$d_{rs}^{m} = q_{rs} \cdot p_{rs}^{m} = q_{rs} \cdot \frac{e^{V_m/\lambda}\left(\sum_{m=1}^{M^W} e^{V_m/\lambda}\right)^{\lambda-1}}{\left(\sum_{m=1}^{M^W} e^{V_m/\lambda}\right)^{\lambda} + \left(\sum_{m=1}^{M^D} e^{V_m/\lambda}\right)^{\lambda}} \tag{6.14}$$

The values of coefficients used in the model are listed in Table 6.1.

Table 6.1 Input coefficients.

coefficient	value	coefficient	value
φ_{iv}	1.000	θ_{iv}	-0.0221
φ_{iw}	1.932	θ_{iw}	-0.0427
φ_{xf}	2.262	θ_{xf}	-0.2210
φ_{xfw}	10.00	θ_{xfw}	-0.0500
ξ	6.740	$\theta_{\mu}, \theta_{prk}, \theta_{opr}$	-0.0099
		θ_{toll}	-0.0050
ASC	value	ASC	value
walk to local/express bus	-0.3757	drive to local/express bus	-1.5892
walk to BRT	-0.0500	drive to BRT	-1.0209
walk to rail	0.2757	drive to rail	-0.4525

The WFRC model uses the same coefficients for money related components including transit fare, parking, and vehicle operating cost with the exception of the toll. The Alternative Specific Constant (ASC) for modeling mode-bias from unobserved factors among transit vehicles by access modes is also represented.

Operating cost

There are three vehicle types for the UTA. The NTD (2013) gathers the annual transit performance from most agencies in U.S. and provides the relevant statistical data. For the UTA, the operating cost per passenger-mile is simply calculated by dividing the

total operating expenses by passenger-miles. The cost per passenger-mile for the bus is relatively higher due to the price of gas and driver wages. The two rail systems incur almost the same cost; however, FrontRunner carries passengers a much greater distance.

Table 6.2 Operating cost per passenger trip by each vehicle type in UTA.

Mode	Total operating cost ($)	Total passenger-miles	Cost per passenger-mile ($)	Average mile per passenger (miles)
Bus	107,815,451	124,309,210	0.867	5.8
TRAX	34,821,026	71,081,431	0.490	4.6
FrontRunner	20,517,540	41,565,944	0.494	25.8

Fare model specification

The primary objective of this research is to determine the appropriate fare levels of the distance-based fare structure after a fare structure reform. By doing so, the financial performance of the new fare structure is kept the same as the current fare structure. Thus, the constraints that limit the revenue achievement to the operating cost are designed to have the same fare box recovery ratio, which is 20% of the operating costs (OULAG 2012). However, for the FrontRunner, the higher fare box recovery ratio has been applied due to its relatively high total cost. The specified expression for the demand maximized fare model can be presented as:

$$\max \sum_r \sum_s q_{rs} \tag{6.15}$$

s.t.

$$\sum_r \sum_s r_{rs}^l d_{rs} = 0.2 \cdot \left(\sum_{l \in TA} \sum_{p \in P_{rs}} d_{rs} \cdot c_l \cdot s_l \cdot \delta_{rs}^{pl}\right) \quad l \in \{buses, TRAX, BRT\} \tag{6.16}$$

$$\sum_r \sum_s r_{rs}^F d_{rs} = 0.4 \cdot \left(\sum_{l \in TA} \sum_{p \in P_{rs}} d_{rs} \cdot c_l \cdot s_l \cdot \delta_{rs}^{pl}\right) \quad l \in \{FrontRunner\}$$

where

$r_{rs}^F = \psi_F - b + s_{rs}^F \cdot m$

$r_{rs}^{l} = b + \sum_{l \in TA} \sum_{p \in P_{rs}} s_l \cdot m \cdot \delta_{rs}^{pl} - r_{rs}^{F} \cdot \delta_{rs}^{pF}$

r_{rs}^{F} : revenue allocated to FrontRunner

r_{rs}^{l} : revenue allocated to the other modes

$\delta_{rs}^{p,F}$: one if route *p* passes through FrontRunner, zero otherwise

ψ_F : premium for using FrontRunner

s_{rs}^{F} : FrontRunner in-vehicle distance

GA settings

Table 6.3 summarizes the initial settings used in the alpha constrained genetic algorithm.

Table 6.3 Initial settings for alpha constrained genetic algorithm.

Variables		Setting
N	number of populations	20
T	maximum iterations	30
	binary substrings of length	6
	cross over probability	0.8
	mutation probability	0.01
	interval for the base fare ($)	[0,5.0]
	interval for the mileage fare (cents)	[0.00, 30.00]
	interval for the premium fare ($)	[0,5.0]
b_1	constant in eq. (6.7)	0.02
b_2	constant in eq. (6.7)	0.4
η^+		1.8
η^-		0.2

Convergence Characteristics by the Alpha Constrained Method

Before discussing the optimal fare levels from the proposed methodology, the solution evolutions, especially the satisfaction of the constraints, are described in this section. Because the alpha constrained GA method places the priority on the feasibility

of the solution which is measured by the satisfaction levels of the solutions, the solution's fare recovery ratio for each constraint has been graphed on Figure 6.5.

Each point of the graph is the average of the solutions of each iteration. The constraints for the fare recovery ratio for FrontRunner and the other transit vehicles are 40% and 20%, respectively. As we can see, the satisfaction level for the fare recovery constraint improves as the algorithm continues because the alpha constrained GA method strictly prioritize only the feasible solutions. Generally, the solutions have been identified before the allowed maximum iterations.

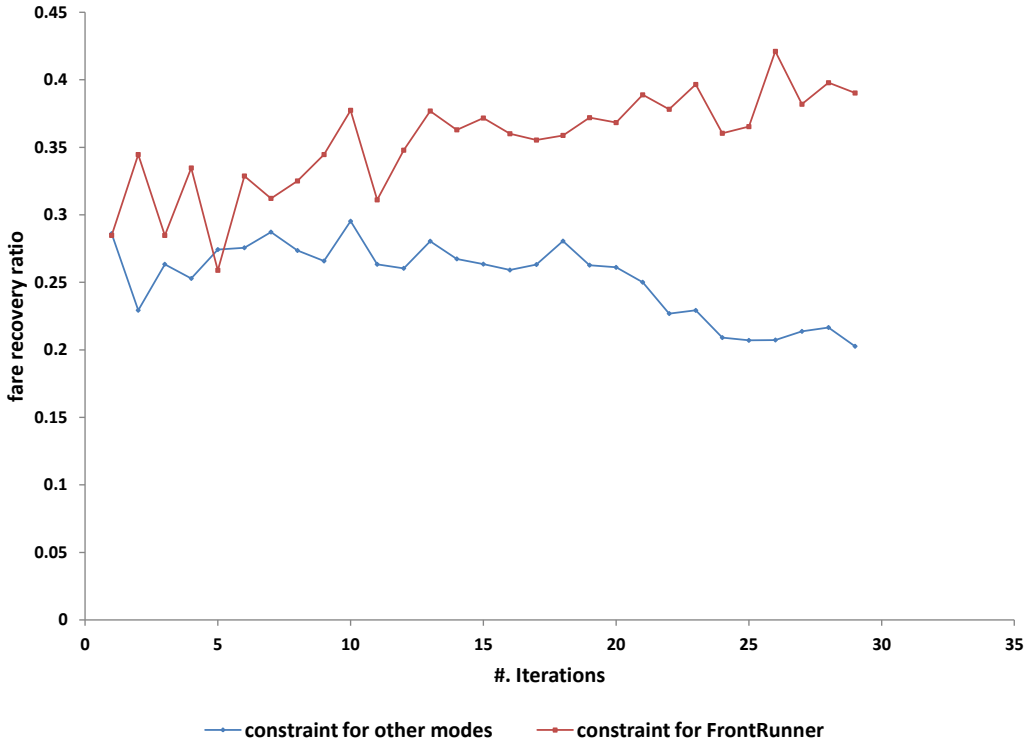


Figure 6.5 Improvement of the fare recovery ratio of solutions as the algorithm iterates.

Analysis on Various Fare Levels of Distance-Based Fare

Using the proposed methodology, this section finds the distance-based fare structure's optimal fare levels and describes the expected changes resulting from the transition.

The evolution of the optimal fare levels is demonstrated on Figure 6.6. Increasing order of demand does not necessarily coincide with the sequence of the solution improvement in the algorithm. Considering that the estimated demand from the model with the current fare structure is about 69,061 trips, the presented fare levels show the higher performance in terms of demand maximization. The following pattern is easily identified: when the base fare decreases, the demand increases. The demand touches its peak when the base fare becomes zero. The implication is that potential transit demands can be found where the users' travel distance is short enough to be less than the base fare. It has been regarded that the distance-based fare is an alternative for achieving fairness of fare charging. However, the increasing demand pattern coupled with the low base fare implies the possibility of overpayment by some based on their actual distance traveled despite the more fair distance-based fare structure. Furthermore, this is opposite to the case of the previous studies that sought the revenue potential from long distance trips.

Table 6.4 Optimal fare evolution.

Base fare ($)	Mileage fare (cents/mile)	FrontRunner	Demand
0.00	15.71	0.83	88,130
0.25	14.33	0.90	85,364
0.50	13.10	1.27	82,543
1.02	10.60	2.29	76,920
1.52	8.00	3.55	72,351

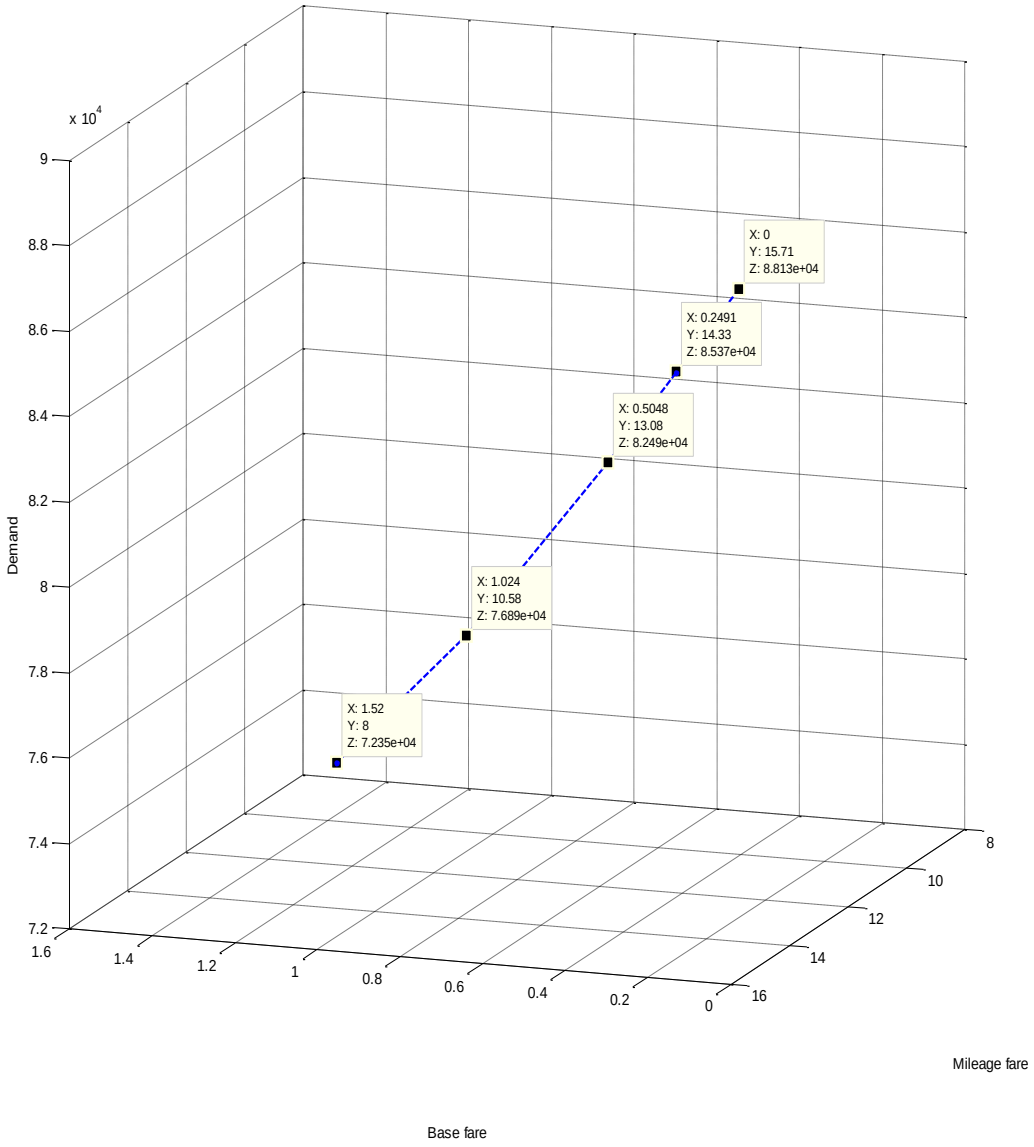


Figure 6.6 Evolutions of optimal solution for distance-based fare structure.

The time-expanded network also effectively represents the transfer behaviors caused by the fare structure change. The increased number of transfers from surrounding local buses is demonstrated when the current fare structure is changed to the distance-based fare which lowered the transfer fare to express buses (Figure 6.7). The time expanded network explicitly produces the vehicle information from which the transfer

flows come and the number of transfers and the vehicle information from which the transfer flows come.

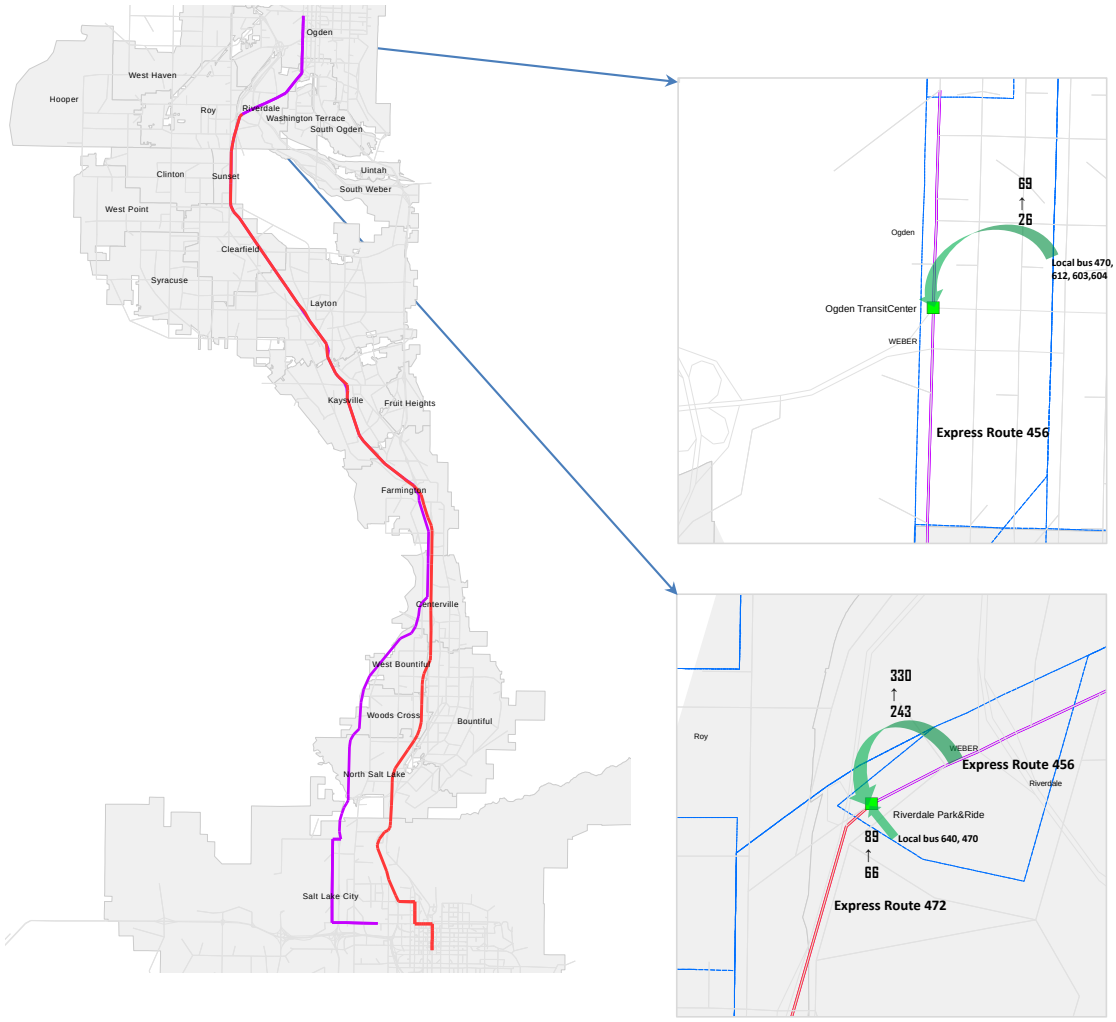


Figure 6.7 Transfer from adjacent local transit vehicles.

Appropriate Distance-Based Fare Structure for UTA

This section analyzes the appropriate fare levels for the distance-based fare structure that allows for the potential to increase ridership. This analytical approach to the ridership potential is based on investigating the transit demand characteristics rather than simply comparing the demand variations of different fare structures.

According to the optimal fare levels estimated from the proposed methodology, the mileage-only fare has the highest potential for maximizing the transit demand. However, most agencies implementing the distance-based fare structure collect the base fare in addition to the mileage fare. They do this because the base fare ensures minimum revenue for the agencies. This section of the study identifies the appropriate base fare level for the UTA that maintains the potential ridership increase.

Assume a virtual flat fare that divides the transit market with distance-based fare structure into two: one that pays more than the flat fare (higher priced market) and one that pays less than the flat fare (lower priced market). With this specification, the conditions for maximizing ridership are dependent on the market segmentation. Specifically, in that positive maximum ridership change can be expected when the lower priced market is greater than the higher priced market (Ling 1998).

$$\frac{\partial \Delta T}{\partial q_1} \geq 0 \quad if \quad \mu_d \leq \mu_f \tag{6.17}$$

$$\frac{\partial \Delta T}{\partial q_2} \leq 0 \quad if \quad \mu_d \geq \mu_f \tag{6.18}$$

where

$$q_i = k\mu_f^{-e_i}$$

$$q_i^{'} = q_i\left(\frac{\mu_d}{\mu_f}\right)^{-e_i}$$

$$\Delta T = \sum_{i=1}^{n} q_i^{'} - \sum_{i=1}^{n} q_i = \sum_{i=1}^{n} q_i\left[\left(\frac{\mu_d}{\mu_f}\right)^{-e_i} - 1\right]$$

μ_f : flat fare

μ_d : distance-based fare

q_i : initial demand in market i

q_i' : ridership under distance-based fare in market i

e_i : market i 's fare elasticity of demand

ΔT : change of total ridership

k : constant

$i = \{1, 2\}$: 1= lower priced market, 2 = higher priced market

The next section applies this market segmentation analysis to the UTA system with the fixed fare that divides the UTA market into two in order to examine the potential of the ridership increase.

UTA Actual Fare Payment

According to the analysis, the eligibility of the distance-based fare structure bearing the increased ridership potential depends on the level of fixed fare in the market. In this section, the appropriate base fare levels are determined through the consideration of the actual payment of the UTA users.

The current UTA's fare structure is divided into two fare categories: users paying $2.35, riding any local and BRT buses and TRAX, and users paying additional fare to use the premium services. The scope for the fixed fare estimation is confined to the first user group. This group represents the demand majority (>85%) of rides. The second user group is regarded as insensitive to the fare changes due to long distance trips (Daskin et al. 1988).

The $2.35 fixed fare for the majority user group does not indicate the actual amount of payment. There are various fare types that provide discounted rates including Eco pass, Education Pass, normal pass, and tokens. These discount fare arrangements, primarily accepted on buses and TRAX, diminish the actual fixed fare level for these user group. According to the Office of the Utah Legislative Auditor General (OULAG 2012), the average fare was $0.30 for the Education Pass, $0.64 for the Eco Pass and $1.28 for cash/tokens and normal passes per boarding using the 2008 database. Taking this into consideration, the average total fare per boarding was $0.81. The study roughly estimated the total fare per trip by multiplying the total fare per boarding ($0.81) to the average number of transfers. American Public Transportation Association (2007) estimated the average number of transfers as 1.53 through the large-scale survey that analyzed 496,000 public transit riders from 2000 through 2005. Thus, the converted fixed fare per trip in the UTA system is approximately $1.32 considering the adjustment of the price using the consumer price index between 2008 and 2012.

A graph has been plotted that shows the proportional change of the lower priced market as the fixed fare increases. Take a look at the point that crosses 50% of the lower priced market in Figure 6.8. Drawing a vertical line on $1.32 easily let us make a conclusion that the feasible base fare that bears the potential of demand increase is less than or equal to $0.5. Assuming that the ridership change is proportional to the size of the lower priced market the best fare structure that has the highest potential for maximizing ridership is the same as the previous analysis. However, the figure shows more options for the base fare up to $0.5 with the corresponding mileage charge while keeping the potential of increasing ridership with the distance-based fare structure.

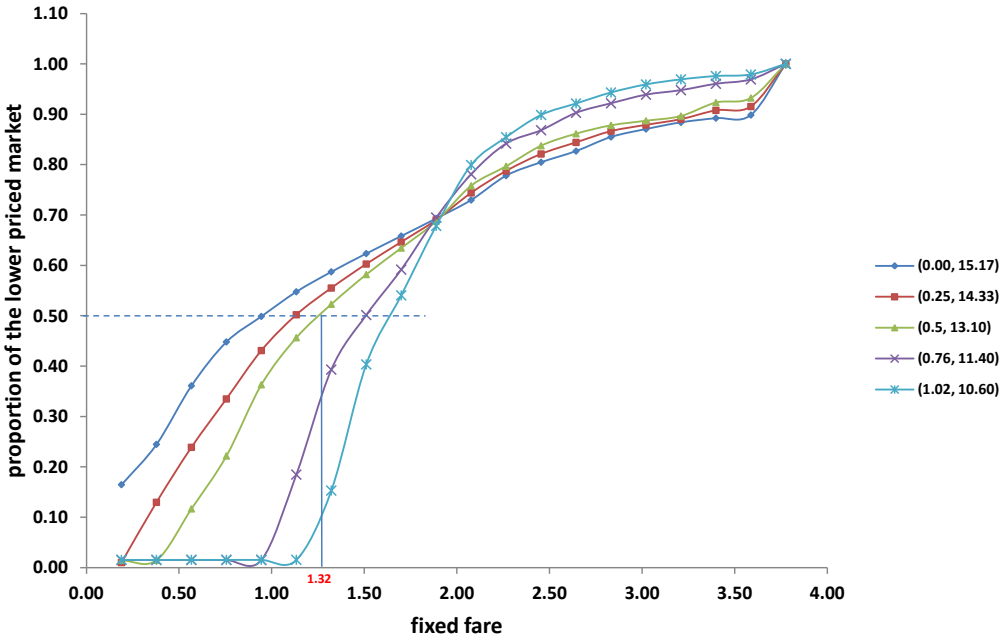


Figure 6.8 Cumulative lower priced market share when the fixed fare is $1.32.

Effects of Buffer Distance

When the base fare is $0.5 demand maximization is not significant because the portion of each market segment is almost even. The research investigates the effect of introducing a buffer distance for increasing the market share of the lower priced market. Placing the buffer distance that waives the fare to a certain miles is one of practically implemented distance-based fare structure types. Also, the properly determined buffer distance can be used to replace the fare free zone in the downtown area in Salt Lake City. The mathematical expression of eq. (6.3) is slightly modified as:

$$\mu_{rs} = b + m \cdot \max[(\omega_{rs} - \lambda), 0] + \{\max_{l \in L} \psi_l \cdot \delta_{rs}^{pl}\} \tag{6.19}$$

where λ is the buffer distance

Using the same methodology, the optimized fare levels for different buffer distances have been identified for the base fares less than or equal to $0.50.

Table 6.5 Optimized fare levels of the distance-based fare with the buffer distances.

Base fare	Mileage fare (cents)	FrontRunner ($)	Buffer distance
0.00	16.30	1.06	1
	16.92	0.75	2
	17.55	0.45	3
0.25	15.09	1.68	1
	15.90	1.48	2
	16.75	1.29	3
0.50	13.75	2.50	1
	14.43	2.25	2
	15.25	2.00	3

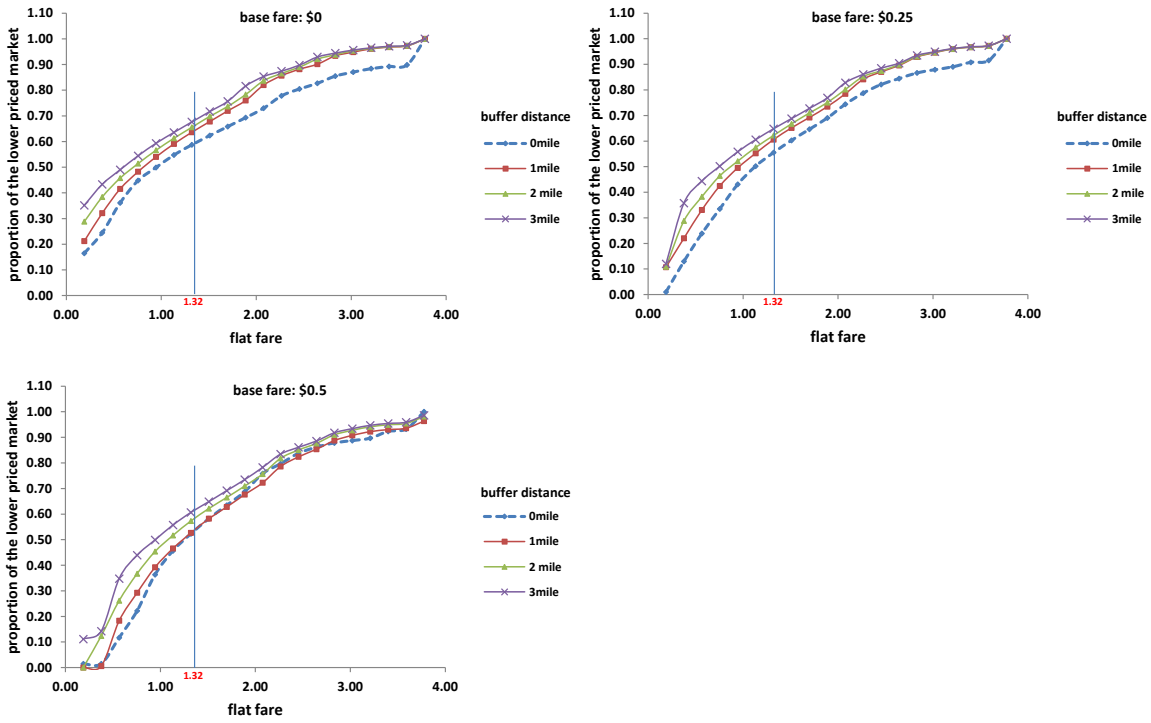


Figure 6.9 Changing the lower market shares with different buffer distances.

By providing fare free services for a certain miles, the mileage fare has been increased slightly from the zero buffer distance examples. However, the fare free effect is greater than the increased mileage fare because the lower priced market shares step up as the buffer distance increases (Figure 6.9). For all three cases of different base fares, introducing the buffer distance causes the increase of the market share of the lower priced market. The dotted line indicating no-buffer-distance case is located at the bottom for three cases. This implies that increasing the base fare can be offset by the buffer distance for keeping and increasing the potential of the demand maximization.

In addition to the investigation on the market share, the expected ridership increase is estimated assuming a constant elasticity proposed by APTA (1991), which is -0.403. Ling's (1998) condition for the ridership increase is based on the assumption that the fixed fare is the average of the distance-based fare, which links the positive ridership change with the higher proportions of the lower priced market. However, this study examines the ridership increase based on the actual fare that does not correspond to the average of estimated distance-based fare paid by the UTA customers. Thus, the total ridership change could be negative even if the lower priced market is higher than 50%. The estimated total ridership change is based on the aggregated level, but it reflects the effects due to the ratio between distance-based fare and the fixed fare. Be noted that this approach is somewhat conservative since it applies the same elasticity value for both higher and lower priced markets.

The market shares of the lower priced market when the fixed fare is $1.32 is presented in Table 6.6. In the previous section, the zero base fare has the highest potential for demand maximization. However, introducing the buffer distance makes

other fare levels are more attractive because the market share and the expected ridership increase as the free ride distance increases. The market shares of the fare levels of $0.25 with 2, 3 mile buffer distance and $0.5 with 3 mile buffer distance have the similar or higher potential than the zero base fare case. In general, providing the free service for a certain miles could be unfair because absolutely free service is provided for the trips less than the buffer distance, these fare levels would be considered as attractive fare structure for the UTA.

Table 6.6 Market shares of the lower priced market with different fare levels.

base fare and buffer distance	market shares of the lower priced market	expected ridership increase (trips)
0.00, 0 mile	58.7%	4,263
0.00, 1 mile	63.6%	7,239
0.00, 2 mile	65.6%	8,131
0.00, 3 mile	67.6%	9,140
0.25, 0 mile	55.5%	-863
0.25, 1 mile	60.6%	3,496
0.25, 2 mile	62.2%	6,043
0.25, 3 mile	64.8%	8,296
0.50, 0 mile	52.3%	-4,632
0.50, 1 mile	52.8%	-2,971
0.50, 2 mile	57.1%	86
0.50, 3 mile	60.6%	4,003

Because the fixed fare is approximated with the estimated fare per boarding and the number of transfers this section conducts a sensitivity analysis with the different fixed fare levels. Table 6.7 shows the lower priced market shares with two different fixed fare levels; one is lower than the estimated UTA fare and the other one is higher than the UTA fare. As one might expect, as the fixed fare increases the lower priced market easily achieves the 50% market shares associated with the positive ridership changes because the higher fixed fare incurs the greater portion of the lower priced market. When the fixed fare is $1.51 the ridership increase is expected for all buffer distances with

$0.25 base fare. For more accurate estimation on the lower priced market share, it is desirable to estimate the fixed fare per trip in the UTA as closely as possible.

Table 6.7 Market shares of the lower priced market when fixed fare changes.

base fare and buffer distance	fixed fare $1.13		fixed fare $1.51	
	market share of lower priced market	expected ridership increase (trips)	market share of lower priced market	expected ridership increase (trips)
0.00, 0 mile	54.8%	1,262	62.3%	7,380
0.00, 1 mile	59.1%	3,433	67.8%	10,034
0.00, 2 mile	61.3%	5,016	69.8%	10,926
0.00, 3 mile	63.5%	6,737	71.7%	11,827
0.25, 0 mile	50.2%	-4,874	60.3%	2,705
0.25, 1 mile	55.2%	-446	65.1%	6,839
0.25, 2 mile	57.5%	2,198	66.7%	9,007
0.25, 3 mile	60.5%	4,614	68.8%	11,043
0.50, 0 mile	45.6%	-9,001	58.2%	-908
0.50, 1 mile	46.6%	-7,549	58.4%	598
0.50, 2 mile	51.4%	-4,020	61.9%	3,716
0.50, 3 mile	55.7%	190	64.8%	7,229

Readiness of the UTA for Distance-Based Fare Structure

The successful implementation of the fare structure reform should be accompanied by properly designed infrastructure and a detailed fare policy. When implementing the distance-based fare structure, the policy regarding technology should incorporate users' ability to indicate the beginning and conclusion of their trip. Technology that measures passenger travel distance by boarding/exiting monitoring (for example via card swipe) is the most critical factor when trying to create a fare enforcement policy. By requiring passengers to tap their electronic payment media to readers in a vehicle or on a platform, the agencies are able to charge the fare based on passenger usage. With said policy, the travel distance of a trip will be tracked when the passenger boards and exits using a device that identifies the location.

The global positioning system (GPS) would be used to do this tracking. Typically, readers are installed at a controlled entrance such as a rail line platform or in buses. Thus, it is necessary that the transit system is ready for the infrastructure deployment such as the GPS devices and the electronic payment system.

The current UTA fare collection system is ready for a fare structure change to the distance-based because every UTA bus and train stop has been equipped with tap on-and-off fare boxes allowing customers to pay their fare electronically.

UTA's The Electronic Fare Collection (EFC) system launched on January 1st, 2009 enabling passengers to use electronic media for paying the fare. The EFC facilitates payment processing and the transfers for users. Fare payment as tapping the card to the reader. Furthermore, with the EFC, there is no need to ask for a transfer from the driver. Once passengers tap off, the system automatically applies transfer credits for all transfers the passengers make within a 2-hour period. With the EFC system, the total distance of a passenger's trip can be estimated because all UTA buses are equipped with the GPS devices and readers are installed at the entrance of TRAX and FrontRunner.

The currently implemented tap on-and-off policy, originally designed to discount the transfer credits, can be smoothly adopted for the distance-based fare structure with minor public resistance. Also, the tapping on-and-off data can be used by the UTA to better evaluate and improve service.

Summary

This chapter examined the potential fare structure change for the UTA with an effective representation proposal to their fare structure. The time expanded network and the bi-level programming that modeled interaction between the fare policy maker and the

passengers was constructive in describing the distance-based fare structure. Using the alpha constrained GA method, the appropriate fare levels for the distance-based fare structure have been identified under demand maximization. Unlike the previous studies that investigated the revenue increase from the distance-based fare structure, this chapter examined the potential of the maximization of the transit demand from the fare structure reform. The analysis identified potential transit demand where users' travel distance is short enough to be less than the base fare based on the analysis that showed the highest demand increase when the distance-based fare is associated with zero base fare. However, the chapter provided a few more options for the base fare levels other than zero using the market segmentation analysis. Also, the study identified that providing the buffer distance for certain miles with base fare equal to or less than $0.50 is acceptable for attracting the users to the UTA's transit market.

In addition to the system analysis, the last section addressed the readiness of the UTA for this distance-based fare structure reform. The current enforcement policy and technology requires users to tap-on-and-off. The infrastructure deployment of the UTA is appropriate for the change. The methodology and the results serve as guidance for the agency while considering the fare structure reform.

CHAPTER 7

CONCLUSIONS AND EXPECTED CONTRIBUTIONS

Transportation is related to most aspects of our daily lives. It affects how we commute to work, to school or to the public places. People make the decision on the route and mode of transportation for traveling to their destinations through the complicated process that simultaneously accounts for the supply and demand sides of the transportation system,

Modeling the transportation system is important because it provides a "common ground" for discussing policy and examining the future transportation plan required in practices (Ortuzar and Willumsen 1994). Generally, modeling is a simplified representation of the real world, however, this research added value to the modeling practice by investigating the asymmetric interactions observed in the real world in order to explore potential improvements of the transportation modeling. Thus, the objective of this research was to address the potentials of the ATEP in terms of the solution algorithm and its application. In order to achieve the objective, the dissertation systematically organized the procedure that addressed the literature review, solution algorithm, and model applications.

The research began with investigating the asymmetric situations in both traffic and public transportation systems. Beginning with the simple example of the two-way street link interactions, various instances of the asymmetric interactions have been reviewed. The review classified those interactions according to the source of flows that asymmetrically interact: link, mode, and route. The basic concepts and the general cost

function's mathematical formulation that accounts for interactions were also presented for each example.

Chapter 3 investigated the contributing factors that exist behind the established cost function on the convergence of the ATEP through the direct application of the DP method to each case of identified asymmetric interactions. The experimental design is featured with different asymmetric complexities on the small network and its extension to the real networks where the experiments are designed to examine and to confirm the new findings. As the results of the experiments, contributing factors on the convergence for different types of interactions have been identified. In the case of the link interactions, the study found the possibility of the reduced asymmetric complexities due to both of the supply and demand sides of the problem. The convergence of the mode interactions was dependent on the route composition for each vehicle type. It has been identified that the resemblance of the route composition among modes creates the favorable conditions for the realization of the mode interactions, consequently the convergence is also affected by the situation. In the case of the route interactions, it has been identified that the route costs' sensitivity due to the complexities of the route cost function in the route interactions incurred a patterned computational efficiencies between link and route interactions.

In Chapter 4, an enhancement strategy for the DP method has been introduced. The efficiency of the new algorithm was achieved by fitting the solution algorithm to the ATEP structure. Formulating the ATEP in the space of route flows, both customized step-sizes for each O-D pair and for the entire route flow space were produced. The availability of these step-sizes enables to devise a selection strategy for allowing the

corrector projection to choose the larger step-size. The strategy was successfully implemented because the DP method provides an opportunity to select an alternative step-size between the two projections. The numerical tests demonstrated that the new algorithm significantly relieved the computational burdens of the basic algorithm.

The application of the modeling practice for the asymmetric interactions is conducted on the public transportation system. In Chapter 5, the asymmetric interaction is modeled through the crowding cost function. Especially, the chapter modeled the daily variation of the in-vehicle crowding in the time expanded network representing the real large scale transit system. The crowding considered transit assignment for the real network could be implemented by using the reliable shortest path finding method (Chen et al. 2011b) and the DP method. Through, the direct application of the model into the real network, several issues related to crowd modeling could be identified. First, the aggregated behavioral response to crowding revealed the possibility of overestimation of concentrated crowding congestion of the no-crowding model. Because the estimation on the transit users' activities directly affects the public transport facility design, the finding is suggestive for the system planning. Second, the different attitudes to crowding could result in the revenue increase for transit agencies. The user's behavior analysis showed that the users with the high confidence level want to use the route with the less expected crowding cost despite of the higher fares, which incurs the increase of revenue. Lastly, the comparison of the expected crowding cost before and after the extension project of FrontRunner South showed that consideration of the multi user classes enables to identify the discomfort reduction by the transit service supply project for multiple user classes simultaneously.

Chapter 6 examined the advantages of the properly structured transit fare system by simulating the passenger behaviors in the public transportation system. Through the effective modeling of the fare structure, the study investigated potential transition to the distance-based fare structure of the UTA network. The time expanded network and the bi-level programming that modeled interaction between the fare policy maker and the passengers was constructive in describing the distance-based fare structure. The major findings of the chapter was identification of potential transit demand where users' travel distance is short enough to be less than the base fare. However, the study provided a few more options for the UTA on the base fare levels other than zero using the market segmentation analysis. The study identified that providing the buffer distance for certain miles with base fare less than $0.50 is acceptable for attracting the users to the UTA's transit market.

In summary, the dissertation makes following conclusions based on the experiments that attended the DP based algorithm for the ATEP and its applications.

First, there are various contributing factors that affect the convergence of different types of the ATEP. This could be resulted from the reduced asymmetric complexities affected by demand intensities and the network configuration, and route composition between different modes, and the sensitivity of the cost function. In addition, the proposed acceleration strategy for enhancing the DP method was successful for achieving the computational efficiencies for the different types of the ATEP.

In the application perspective, the research identified that modeling ATEP is useful to provide insight on crowd modeling and fare policy approaches. Based on the results of the crowd modeling application, a conclusion has been made that crowd

modeling with multi user classes could influence the public transportation system planning and the revenue achievement of transit agencies. Moreover, addition of the disutility factor, crowding, not always causes the increase of disutility from the transit uses. Investigation on the ATEP application modeling the distance-based fare structure for the UTA found that the zero base fare has the highest potential for increasing the transit demand. However, the model provides more options on the base fare levels for attracting the users to UTA's transit market upon the fare structure change.

The contributions of the research can be summarized as follows:

1. The contributing factors influencing the convergence of the different types of the ATEP have been identified
2. A strategy for enhancing the computational efficiencies for the DP method has been developed
3. The in-vehicle crowd modeling is conducted in the real large scale transit network under the reliability based transit assignment
4. Identified the users' behavioral change and its consequences on the system planning and the revenue change due to crowd modeling with multi user classes.
5. An advanced transit assignment model using the time expanded network that considers the non-additive transit fare structure has been developed
6. Using the proposed asymmetric modeling for the fare structure, it has been identified that reform to the distance-based fare structure will have the potential to increase the demand in the Utah Transit Authority.

REFERENCES

Amijo, L. (1966). "Minimization of functions having continuous partial derivatives." *Pacific J. Math.*, 16, 1-3.

American Public Transportation Association (APTA). (1991). "Fare elasticity and its application to forecasting transit demand." Washington D.C.

APTA. (2007). "A profile of public transportation passenger demographics and travel characteristics reported in on-board surveys." Washington D.C.

APTA. (2013), "Ridership report archives." http://www.apta.com/resources/statistics/Pages/RidershipArchives.aspx (Apr. 2 2013).

Australian Transport Council. (2006). "National guidelines for transport system management in Australia." Canberra, Australia, 69-70.

Bäck, T., and Hoffmeister, F. (1991). "Extended selection mechanisms in genetic algorithms." *Proc. 4th Int. Conf. Genetic Algorithms*, Morgan Kaufmann, Burlington, MA, 92-99.

Baker, J.E. (1984). "Adaptive selection methods for genetic algorithms." *Proc. 1st Int. Conf. Genetic Algorithms and Their Applications*. Lawrence Erlbaum Associates, Carnegie-Mellon University, Pittsburg, PA, 101-111.

Bekhor, S., Toledo, T., and Prashker, J.N. (2006). "Implementation issues of route choice models in path-based algorithms." *Proc. 11th Int. Conf. Travel Behav. Res.*, Kyoto, Japan, 16-20.

Beckmann, M.J., Mcquire, C.B., and Winsten, C.B. (1956). *Studies in the economics of transportation*. Yale University Press, New Haven, Conn.

Bernstein, D., and Gabriel, G. (1997). "Solving the nonadditive traffic equilibrium problem." *Netw. Optimizat.*, Springer, 72-102

Bertsekas, D.P. (1976). "On the Goldstein-Levitin-Poylak gradient projection method." *IEEE Transp. Autom. Cntr.*, 21, 74-184.

Bertsekas, D.P., and Gafni, E.M. (1982). "Projection methods for variational inequalities with application to the traffic assignment problem." *Math. Program. Stud.*, 17, 139–159.

Borndörfer, R., Neumann, M., and Pfetsch, M. (2012). "Models for fare planning in public transport." *Discrete Appl. Math.,* 160(18), 2591-2605.

Boyce, D., Ralevic-Dekic, B., and Bar-Gera H. (2004). "Convergence of traffic assignments How Much is Enough?" *J. Transp. Eng.*, 130(1), 49–55.

Ceapa, I., Smith, C., and Capra, L. (2012). "Avoiding the crowds: Understanding Tube station congestion patterns from trip data." *Proc. of the ACM SIGKDD Intl. Workshop on Urban Computing*, Beijing, China, 134-141.

Cervero, R. (1981). "Flat versus differentiated transit pricing: what's a fair fare?" *Transportation* (10), 211-232.

Cervero, R., and Wachs, M. (1982). "An answer to the transit crisis: the case for distance-based fares." *J. Contemp. Stud.* 5(2), 59-70.

Chen, A. (2001). "Effects of flow update strategies on the implementation of the Frank-Wolfe algorithm for the traffic assignment problem." *Transp. Res. Rec.*,1771, 132-139.

Chen, A., and Ji, Z. (2005). "Path finding under uncertainty." *J.Adv. Transp.*, 39 (1), 19–37.

Chen, A., and Zhou, Z. (2010). "The alpha-reliable mean-excess traffic equilibrium model with stochastic travel times." *Transp. Res. B*, 44, 493-513.

Chen, A., Zhou, Z., and Ryu, S. (2011a). "Modeling physical and environmental side constraints in traffic equilibrium problem." *Intl. J. Sustainable Transp.*, 5(3), 72-197.

Chen, B.Y., Lam, W.H.K., Sumalee, A., and Shao, H. (2011b). "An efficient solution algorithm for solving multi-class reliability-based traffic assignment problem." *Math. Comp. Mod.*, 54, 1428-1439.

Chien, S.I., and Tsai, C.F.M. (2007). "Optimization of fare structure and service frequency for maximum profitability of transit systems." *Transp. Planning and Tech.*, 30(5), 477-500.

Chien, S.I., Tsai, C.F.M., and Spasovic, L.N. (2001). "Optimization of grid bus transit systems with elastic demand." *J. Adv. Transp.*, 36(1), 63-91.

Coelloc, C., Van Velhhuizen, D.A., and Lamont, G.B. (2002). *Evolutionary algorithm for solving multi-objective problems.* Springer Science+Business, New York, NY.

Cohen, G., and Chaplais, F. (1988). "Nested monotonicity for variational inequalities over product of spaces and convergence of iterative algorithms." *J. Optimizat. Theor. and Appl.,* 59, 369-390.

Critical Transit. (2014). "Operating the right transit mode." <http://www.criticaltransit.com/2012/08/13/operating-the-right-transit-mode/> (Mar. 5 2014).

Dafermos, S.C. (1971). "An extended assignment model with applications to two-way traffic." *Transp. Sci.,* 5(1), 366-389.

Dafermos, S.C. (1972). "The traffic assignment problem for multiclass-user transportation networks." *Transp. Sci.,* 6(1), 73-87.

Dafermos, S.C. (1980). "Traffic equilibrium and variational inequalities." *Transp. Sci.,* 14(1), 42-54.

Dafermos, S.C. (1982). "Relaxation algorithms for the general asymmetric traffic equilibrium problem," *Transp. Sci.,* 16(2), 231-240.

Dafermos, S.C. (1983). "An iterative scheme for variational inequalities," *Math. Program.* 29, 40-47.

Daskin, M.S., Schofer, J.L., and Haghani, A.E. (1988). "A quadratic programming model for designing and evaluating distance-based and zone fares for urban transit." *Transp. Res. B,* 22(1), 25-44

Deb, K. (2001). *Multi-objective optimization using evolutionary algorithms*. John Wiley and Sons, New York.

De Cea, J., and Fernandez, E. (1993). "Transit assignment for congested public transport systems: an equilibrium model." *Transp. Sci.,* 27 (2), 133–147.

Dijkstra, E. W. (1959). "A note on two problems in connexion with graphs." *Numerische Mathematik,* 1, 269–271.

Dupuis, C., and Darveau, J.M. (1986). "The convergence conditions of diagonalization and projection methods for fixed demand asymmetric network equilibrium problems." *Oper. Res. Lett.,* 5, 149-155.

Dupuis, P., and Nagurney, A. (1993). "Dynamical systems and variational inequalities." *Ann. Oper. Res.,* 44, 9-42.

Fisk, C., and Nguyen, S. (1982). "Solution algorithms for network equilibrium with asymmetric user costs." *Transp. Sci.,* 16, 361-381.

Florian, M., and Spiess, H. (1982). "The convergence of diagonalization algorithms for asymmetric network equilibrium problems." *Transp. Res. B*, 16, 447-483.

Federal Highway Administration. (2004). "Traffic congestion and reliability: linking solutions to problems." U.S. Department of Transportation, Washington, D.C.

Fu, X., Lam, W.H.K., and Chen, B.Y. (2014). "A reliability-based traffic assignment model for multi-modal transport network under demand uncertainty." *J. Adv. Transp.*, 48, 66-85.

Gabriel, L.S., and Bernstein, D. (1997). "The traffic equilibrium problem with nonadditive path costs." *Transp. Sci.*, 31, 337-348.

Glaister, S., and Collings, J.J. (1978). "Maximization of passenger miles in theory and practice." *J. Transp. and Econ. and Pol.* 12(3), 304-321.

Goldberg, D.E. (1989). *Genetic algorithm in search, optimization and machine learning*. Addison-Wesley, Reading, Mass.

Grange, L.D., and Munoz, J.C. (2009). "An equivalent optimization formulation for the traffic assignment problem with asymmetric linear costs." *Transp. Plan. and Tech.*, 32, 1-25.

Green, G. (1828). *An essay on the application of mathematical analysis to the theories of electricity and magnetism*, Wheelhouse, Nottingham, England.

Hamdouch, Y., Marcotte, P., and Nguyen, S. (2004). "Capacitated transit assignment with loading priorities." *Math. Program. B*, 101, 205–230.

Hamdouch, Y., and Lawphongpanich, S. (2008). "Schedule-based transit assignment model with travel strategies and capacity constraints." *Transp. Res. B*, 42, 663-684.

Han, D.R., and He, B.S. (2001). "An implementable stepsize rule for Goldstein-Levitin-Polyak projection method." *J. Chinese Univ.*, 23, 56-62.

Han, D., and Sun, W. (2004). "A new modified Goldstein-Levitin-Polyak projection method for variational inequality problems." *Compu. and Math. with Appl.*, 47(12), 1817-1825.

He, B.S., Yang, H., Meng, Q., and Han, D.R. (2002). "Modified Goldstein-Levitin-Polyak projection method for asymmetric strongly monotone variational inequalities." *J. Optimizat., Theor. and Appl.*, 112, 129-143.

Hensher, D.A., and McLeod, P.B. (1977). "Towards an integrated approach to the identification and evaluation of the transport determinations of travel choices." *Transp Res.,* 11(2), 77-93.

Heydecker, H.G. (1983). "Some consequences of detailed junction modeling in road traffic assignment." *Transp. Sci.,* 17(3), 263-281.

Hickey, R. (2005). "Impact of transit fare increase on ridership and revenue: metropolitan transportation authority, New York City." *Transp. Res. Rec.*,1927, 239-248.

Kepaptsoglou, K., and Karlaftis, M. (2009). "Transit route network design problem: Review." *J. Transp. Eng.* 135(8), 491-505.

Khobotov, E.N., (1987). "Modification of the extra-gradient method for solving variational inequalities and certain optimization problems." *USSR Comp. Math. and Mathematical Phys.*, 27(5), 120–127.

Kinderlehrer, D., and Stampacchia, G. (1980). *An introduction to variational inequalities and their applications*. Academic Press, New York.

Korpelevich, G.M. (1977). "The extra-gradient method for finding saddle points and other problems." *Matecon,* 12, 747-756.

Koutsopoulos, H.N., and Habbal, M. (1994). "Effect of intersection delay modeling on the performance of traffic equilibrium models." *Transp. Res. A,* 28, 113-149.

Lam, W.H.K., Gao, Z.Y., Chan, K.S., and Yang, H. (1999). "A stochastic user equilibrium assignment model for congested transit networks." *Transp. Res. B,* 33, 351-368.

Lawphongpanich, S., and Hearn, D.W. (1894). "Simplicial decomposition of the asymmetric traffic assignment problem." *Transp. Res. B*, 18, 123-133.

Lee, C., and Machemehl, R.B. (2005). "Combined traffic signal control and traffic assignment: algorithms, implementation and numerical results." Center for Transportation Research, University of Texas at Austin.

Lee, J. (2013). "UTA considers pay-by-distance plan, changes to free fare zone." <http://www.ksl.com/index.php?sid=18615672&nid=481> (Jan.14 2013).

Li, Z., and Hensher, D.A. (2011). "Crowding and public transport: a review of willingness to pay evidence and its relevance in project appraisal." *Transp. Pol.*, 18, 880-887.

Ling, J. (1998). "Transit fare differentials: a theoretical analysis." *J. Adv. Transp.*, 32(3), 297-314.

Lo, H.K., and Chen, A. (2000). "Traffic equilibrium problem with route-specific costs: formulation and algorithms." *Transp. Res. B,* 34(6), 493-513.

Lo, H.K., Luo, X.W., and Siu, B.W.Y. (2006). "Degradable transport network: travel time budget of travelers with heterogeneous risk aversion." *Transp. Res. B,* 40(9), 792-806.

Mahmassani, H. S., and Mouskos, K.C. (1988). "Some numerical results on the diagonalization algorithm for network assignment with asymmetric interactions between cars and trucks." *Transp. Res. B,* 22, 275-290.

Marcotte, P. (1991). "Application of Khobotov's algorithm to variational inequalities and network equilibrium problems." *Inform. Syst. and Oper. Res.,* 29, 258-270.

Marcotte, P., and Guelat, J. (1988). "Adaptation of a modified Newton method for solving the asymmetric traffic equilibrium problem." *Transp. Sci.,* 22(2), 112-124.

Marcotte, P., and Wynter, R. (2004). "A new look at the multiclass network equilibrium problem." *Transp. Sci.,* 38(3), 282-292.

Meneguzzer, C. (1995). "An equilibrium route choice model with explicit treatment of the effect of intersections." *Transp. Res. B,* 29(5), 329-356.

National Transit Database. (2013). <http://www.ntdprogram.gov/ntdprogram/> (Apr. 14 2013).

Moller-Pedersen, J. (1999). "Assignment model of timetable based systems (TPSCHEDULE)." *Proc. 27th European Transportation Forum, Seminar F*, Cambridge, England, 159–168.

Nagurney, A. (1984). "Comparative tests of multimodal traffic equilibrium methods." *Transp. Res. B,* 13, 183-187.

Nagurney, A. (1986). Computational comparisons of algorithms for general asymmetric traffic equilibrium problems with fixed and elastic demands. *Transp. Res. B,* 20, 78-84.

Nagurney, A. (2002). "Variational Inequalities." <http://supernet.isenberg.umass.edu/austria_lectures/fvisli.pdf> (March 23 2013).

Nagurney, A., and Zhang, D. (1996). *Projected dynamical systems and variational inequalities with applications,* Kluwer Academic Publishers, Boston, Mass.

Nash, C. A. (1978). "Management objectives, fares and service levels in bus transport." *J. Transp. Econ. and Pol.*, 12(1), 70-85.

Netter, M., and Sender, J.G. (1970). "Equilibre offre-demande et tarification sur un reseau de transport." Institute de Recherche des Transports, Arcueil, France.

Nguyen, S., and Dupuis, C. (1984). "An efficient method for computing traffic equilibria in networks with asymmetric transportation costs." *Transp. Sci.*, 18(2), 185-202.

Nguyen, S., Pallottino, S., and Malucelli, F. (2001). "A modeling framework for passenger assignment on a transport network with timetables." *Transp. Sci.*, 35(3), 238-249.

Nielsen, O.A., and Jovicic, G. (1999). "A large-scale stochastic timetable-based transit assignment model for route and submode choices." *Proc.27th European Transportation Forum, Seminar F*, Cambridge, England, 169-184.

Office of the Utah Legislative Auditor General. (2012). "A performance audit of the Utah Transit Authority." <http://le.utah.gov/audit/12_01rpt.pdf> (Apr. 24 2013).

Ortega, J.M., and Rheinboldt, W.C. (1970). *Iterative solution of nonlinear equations in several variables.* Academic Press, New York.

Ortuza, J.D., and Willumsen, L.G. (1994), *Modeling transport*, John Wiley & Sons, New York.

Panicucci, B., Pappalardo, M., and Passacantando, M.(2007). "A path-based double projection method for solving the asymmetric traffic network equilibrium problem." *Optimizat Let.*, 1, 171-185.

Pang, J.S., and Chan, D. (1982). "Iterative methods for variational and complementarity problems," *Math. Program.*, 24, 284-313.

Patriksson, M. (1993). "A unified description of iterative algorithms for traffic equilibria." *European J. Oper. Res.*, 71(2), 154-176.

Patriksson. M. (1994). *The traffic assignment problem: models and methods*, V.S.P. Intl Science, Netherlands.

Poon, M.H., Wong, S.C., and Tong, C.O. (2004). "A dynamic schedule-based model for congested transit networks." *Transp. Res. B,* 38, 343-368

Prager, W. (1954). "Problems of traffic and transportation.", *Proc. the Symp. Oper. Res. in Bus. and Ind.,* Kansas City, Kansas, 105-113.

Pratelli, A. (2004). "The combined zone and fare planning problem." *Urban Transport X. urban transport and the environment in the 21st century*. WIT Press, Southampton, UK, 311-320.

Qian, Z.S., Shen, W., and Zhang, H.M. (2012). "System-optimal dynamic traffic assignment with and without queue spillback: its path-based formulation and solution via approximate path marginal cost." *Transp. Res. B,* 46(7) 874–893.

RideUTA. (2013). <http://www.rideuta.com/ridinguta/routes/routefinder.aspx> (Apr. 14 2013).

Sale, J.E. and Green, B. (1979). "Operating costs and performance of American public transit systems." *J. Am. Planning Assoc.*, 45, 22-27.

Schmocker, J.D. (2006). "Dynamic capacity constrained transit assignment." PhD thesis, Imperial College, London, England.

Shao, H., Lam, W.H.K., Meng, Q., and Tam, M.L. (2006). "Demand-driven traffic assignment problem based on travel time reliability." *Transp. Res. Rec.*, 1985, 220-230.

Shao, H., Lam, W.H.K., and Tam, M.L. (2006). "Reliability-based stochastic traffic assignment model for network with multiple user classes under uncertainty in demand." *Netw. and Spatial Econ.*, 6(3-4), 173-204.

Sheffi, Y. (1985). *Urban transportation networks: equilibrium analysis with mathematical programming methods*. Prentice Hall, Englewood Cliffs, N.J.

Siu, B., and Lo, H.K. (2006). "Doubly uncertain transport network: degradable link capacity and perception variations in traffic conditions." *Transp. Res. Rec.*, 1964, 59-69.

Small, K., and Verhoef, E. (2007). *The economics of urban transportation* – 2nd edition, Routledge, New York.

Smith, M.J. (1979). "The existence, uniqueness and stability of traffic equilibria." *Transp. Res. B,* 13(4), 295–304.

Smith, M.J. (1981). "Properties of a traffic control Policy which ensure the existence of a traffic equilibrium consistent with the policy." *Transp. Res. B*, 15, 453-462.

Spiess, H., and Florian, M. (1989). "Optimal strategies: a new assignment model for transit networks." *Transp. Res. B,* 23(2), 83-102.

Sumalee, A., Tan, Z., and Lam, W.H.K. (2009). "Dynamic stochastic transit assignment with explicit seat allocation model." *Transp. Res. B,* 43, 895-912.

Szeto, W.Y., Solayappan, M., and Jiang, Y. (2011). "Reliability-based transit assignment for congested stochastic transit networks." *Comp.-Aided Civil and Infstr. Eng.* 26, 311-326.

Szeto, W.Y., Jiang, Y., Wong, K.I., and Solayappan, M. (2013). "Reliability-based stochastic transit assignment with capacity constraints: formulation and solution method." *Transp. Res. C,* 35, 286-304.

Takahama, T., and Sakai, S. (2004). "Constrained optimization by alpha constrained genetic algorithm." *Sys. and Comp. in Japan,* 35, 11-21.

Transit Cooperative Research Program (TCRP). (1996). "Fare policies, structure and technologies." Transportation Research Board, Washington D.C., report 10 Chapter 3.

TCRP. (2003). "Fare policies, structure and technologies: update." Transportation Research Board, Transportation Research Board, Washington D.C., report 94 Chapter 2.

Tirachini, A., Hensher, D.A., and Rose, J.M. (2013). "Crowding in public transport systems: effects on users, operation and implications for the estimation of demand." *Transp. Res. A,* 53, 36-52.

Tong, C.O., and Richardson, A.J. (1984). "A computer model for finding the time-dependent minimum path in a transit system with fixed schedules." *J. Adv. Transp.,* 18(2), 145-161.

Tsai, F.M., Chien, S. I., and Spasovic, L.N. (2008). "Optimizing distance-based fares and headway of an intercity transportation system with elastic demand and trip length differentiation." *Transp. Res. Rec.*, 2089, 101-109.

Tsai, F.M., Chien, S. I., and Wei, C.H. (2013). "Joint optimization of temporal headway and differential fare for transit systems considering heterogeneous demand elasticity." *J. Transp. Eng.*, 139(1), 30-39.

Van Vuren, T., and Van Vliet, D. (1992). *Route choice and signal control: the potential for integrated route guidance.* Athenaeum Press Ltd., Newcastle, Great Britain.

Wallace, C.E., Courage, K.G., Hadi, M.A., and Gan, A.G. (1998). *"TRANSYT-7F user's guide."* University of Florida, Gainesville, Fla.

Wardrop, J.G. (1952). "Some theoretical aspects of road traffic research." *Proc. Inst. Civil Engr.*, 2(1), 325-378

Webster, F.V. (1958). "Traffic signal setting." Road Research Laboratory Report 39. Crowthorne, Berkshire, England.

Whelan, A., and Crockett, J. (2009). "An Investigation of the willingness to pay to reduce rail overcrowding." *Proc. 1st Int. Conf. Choice Mdlng.*, Harrogate, England.

Wong, S.C. (1995). "Derivatives of the performance index for the traffic model from TRANSYT." *Transp. Res. B,* 29(5), 303-327.

Wong, S.C., Yang, C., and Lo, H. (2001). "A path-based traffic assignment algorithm based on the TRANSYT traffic model." *Transp. Res. B,* 35(2), 163-181.

Wu, J.H., Florian, M., and Marcotte, P. (1994). "Transit equilibrium assignment: a model and solution algorithms." *Transp. Sci.,* 28(3), 193-203

Yang, H., Zhang, X., and Meng, Q. (2004). "Modeling private highways in networks with entry-exit based toll charges." *Transp. Res. B,* 38(3), 191-213.

Zhang, Y., Lam, W.H.K., Sumalee, A., Lo, H.K., and Tong, C. (2010). "The multi-class schedule-based transit assignment model under network uncertainties." *Plc. Transp.* 2, 69-86.

Zhou, J., and Lam, H.K.W. (2001). "A bi-level programming approach – optimal transit fare under line capacity constraints." *J. Adv. Transp.,* 34(2), 105-124.

APPENDICES

A. Proof of the convergence of the double projection method

Khobotov (1987) presented the convergence theorem for the extra gradient method (double projection method) which follows the same line as the proof of Korperlevich (1977).

Theorem : Let the set Ω of solutions of the VI, eq. (2.19), let Q be a closed convex set, and $C(x)$ a continuous monotonic operator in Q. Then, the extra gradient method eq. (3.1, 3.2) is convergent to a solution of eq. (2.19) from any initial point $u^0 \in Q$.

$$\lim_{k\to\infty} \min_{u^*} \rho(u^*, u^k) = 0, \quad u^* \in \Omega$$

where $\rho(u^*, u^k)$ is the Euclidian distance between points u^* and u^k.

Proof: we estimate the difference $\left\|u^{k+1} - u^*\right\|^2$ for any point $u^* \in \Omega$. By the properties of the projection onto the convex set Q, we see that, for all $v \in Q$ and any u,

$$(u - proj_Q(u), v - proj_Q(u)) \le 0 \tag{7.1}$$

Hence it follows that

$$\|u - v\|^2 = \|u - proj_Q(u) + proj_Q(u) - v\|^2 = \|u - proj_Q(u)\|^2 -$$
$$2(u - proj_Q(u), v - proj_Q(u)) + \|v - proj_Q(u)\|^2 \ge \|u - proj_Q(u)\|^2 + \|v - proj_Q(u)\|^2$$

with $v = u^*(u^* \in \Omega, \Omega \in Q), u = u^k - \alpha_k C(\bar{u}^k)$, we obtain from this inequality, using eq. (6.1)

$$\left\|u^k - \alpha_k C(\bar{u}^k) - u^*\right\|^2 \ge \left\|u^k - \alpha_k C(\bar{u}^k) - u^{k+1}\right\|^2 + \left\|u^* - u^{k+1}\right\|^2$$

which leads to the inequality

$$\begin{aligned} \left\|u^{k+1} - u^*\right\|^2 &\le \left\|u^k - \alpha_k C(\bar{u}^k) - u^*\right\|^2 - \left\|u^k - \alpha_k C(\bar{u}^k) - u^{k+1}\right\|^2 \\ &= \left\|u^k - u^*\right\|^2 - \left\|u^k - u^{k+1}\right\|^2 + 2\alpha_k (C(\bar{u}^k), u^* - u^{k+1}) \end{aligned} \tag{7.2}$$

Since the operator $C(u)$ is monotonic

$$0 \le (C(u)-C(u^*),u-u^*) = (C(u),u-u^*)-(C(u^*),u-u^*) \le (C(u),u-u^*),\ u \in Q$$

Hence

$$(C(\bar{u}^k),u^*-u^{k+1}) = (C(\bar{u}^k),u^*-\bar{u}^k)+(C(\bar{u}^k),\bar{u}^k-u^{k+1}) \le (C(\bar{u}^k),\bar{u}^k-u^{k+1})$$

We have from eq. (6.2)

$$\begin{aligned}
&\left\|u^{k+1}-u^*\right\|^2 \le \left\|u^k-u^*\right\|^2 - \left\|u^k-u^{k+1}\right\|^2 + 2\alpha_k (C(\bar{u}^k),\bar{u}^k-u^{k+1}) \\
&= \left\|u^k-u^*\right\|^2 - \left\|u^k-\bar{u}^k\right\|^2 - \left\|\bar{u}^k-u^{k+1}\right\|^2 - 2(u^k-\bar{u}^k,\bar{u}^k-u^{k+1}) \\
&+2\alpha_k (C(\bar{u}^k),\bar{u}^k-u^{k+1}) \\
&= \left\|u^k-u^*\right\|^2 - \left\|u^k-\bar{u}^k\right\|^2 + 2(u^k-\alpha_k C(\bar{u}^k)-\bar{u}^k,u^{k+1}-\bar{u}^k) - \left\|\bar{u}^k-u^{k+1}\right\|^2 \qquad (7.3) \\
&= \left\|u^k-u^*\right\|^2 - \left\|u^k-\bar{u}^k\right\|^2 - \left\|\bar{u}^k-u^{k+1}\right\|^2 + 2(u^k-\alpha_k C(u^k)-\bar{u}^k,u^{k+1}-\bar{u}^k) \\
&+2(\alpha_k C(u^k)-\alpha_k C(\bar{u}^k),u^{k+1}-\bar{u}^k) \\
&\le \left\|u^k-u^*\right\|^2 - \left\|u^k-\bar{u}^k\right\|^2 - \left\|\bar{u}^k-u^{k+1}\right\|^2 + 2\alpha_k \left\|C(u^k)-C(\bar{u}^k)\right\|\left\|u^{k+1}-\bar{u}^k\right\|
\end{aligned}$$

The last inequality holds since $(u^k-\alpha_k C(u^k)-\bar{u}^k,u^{k+1}-\bar{u}^k) \le 0$. This follows from eq. (6.1) with $v = u^{k+1}, u = u^k-\alpha_k C(u^k)$

For any $u^{k+1},u^k,\bar{u}^k,\alpha_k$ we have

$$\left\|u^{k+1}-\bar{u}^k\right\|^2 + \alpha_k^2\left\|C(u^k)-C(\bar{u}^k)\right\|^2 \ge 2\alpha_k \left\|C(u^k)-C(\bar{u}^k)\right\|\left\|u^{k+1}-\bar{u}^k\right\|$$

In view of this, we obtain from eq. (6.3)

$$\begin{aligned}
&\left\|u^{k+1}-u^*\right\|^2 \le \left\|u^k-u^*\right\|^2 - \left\|u^k-\bar{u}^k\right\|^2 - \left\|\bar{u}^k-u^{k+1}\right\|^2 + \left\|u^{k+1}-\bar{u}^k\right\|^2 \\
&+\alpha_k^2\left\|C(u^k)-C(\bar{u}^k)\right\|^2
\end{aligned} \qquad (7.4)$$

Consider the set $R_k(u^*)$ given by

$$R_k(u^*) = \{u \mid u \in Q \cap \bar{R}_k(u^*)\}$$

where $\bar{R}_k(u^*) = \{u \mid \|u - u^*\|^2 \le \|u^k - u^*\|^2\}$

with $k = 0$ for any point $u^0 \in Q$ and the point $u^* \in \Omega$ for which the difference $\|u^{k+1} - u^*\|^2$ is estimated in eq. (6.2), the set $R_0(u^*)$ is bounded. This follows from the fact that $\bar{R}_0(u^*)$ is bounded.

We now define the set $\hat{R}_0(u^*)$:

$$\hat{R}_0(u^*) = \{\bar{u} \mid \bar{u} = proj_Q(u - \tilde{\alpha}\tilde{b}_0), u \in R_0(u^*), \tilde{b}_0 \in M_0\}$$

where $0 \le \tilde{\alpha}_0 \le \bar{\alpha}$, while

$$M_0 = \{z \mid z \in E^n, \|z\| \le \max_u \|C(u)\|, u \in R_0(u^*)\}$$

Then, for all $\tilde{b}_0 \in M_0, u \in R_0(u^*)$, the set $\hat{R}_0(u^*)$ is bounded, and $R_0(u^*) \subset \hat{R}_0(u^*)$.

The continuous operator $C(\bar{u})$ is also bounded set $\hat{R}_0(u^*)$. Hence there is a constant $L_0 < \infty$ such that

$$\|C(u^0) - C(\bar{u}^0)\| L_0 \|u^0 - \bar{u}^0\| \tag{7.5}$$

for all $u^0 \in \hat{R}_0(u^*), \bar{u}^0 \in \hat{R}_0(u^*)$. Then, with $k = 0$, we obtain from eq. (6.4):

$$\|u^1 - u^*\|^2 \le \|u^0 - u^*\|^2 - \|u^0 - \bar{u}^0\|^2 (1 - \alpha_0^2 L_0^2)$$

It is clear from this that α_0 can always be chosen in such a way that $(1 - \alpha_0^2 L_0^2) > 0$.

In our case the point $u^1 \in R_0(u^*)$, so that $R_1(u^*) \subset R_0(u^*)$, $u^0 \ne \bar{u}^0$. For this, α_0 can be chosen (e.g., from the condition $0 < \alpha_0 \le \min\{\bar{\alpha}, \varepsilon / L_0\}, 0 < \varepsilon < 1$). Moreover, from eq. (6.4) with $k = 0, u^0 \ne \bar{u}^0$, we have

$$\left\|u^1 - u^*\right\|^2 \le \left\|u^0 - u^*\right\|^2 - \left\|u^0 - \bar{u}^0\right\|^2 (1 - \alpha_0^2 \frac{\left\|C(u^0) - C(\bar{u}^0)\right\|^2}{\left\|u^0 - \bar{u}^0\right\|^2})$$

whence it is clear that $u^1 \in R_0(u^*)$ also if

$$0 < \alpha_0 \le \min\{\bar{\alpha}, \varepsilon \frac{\left\|u^0 - \bar{u}^0\right\|}{\left\|C(u^0) - C(\bar{u}^0)\right\|}\}, \ \varepsilon = const, \ 0 < \varepsilon < 1$$

Here, by eq. (6.5),

$$1/L_0 \le \left\|u^0 - \bar{u}^0\right\| / \left\|C(u^0) - C(\bar{u}^0)\right\|$$

If $u^k = \bar{u}^k$ at some iteration of the method with $\alpha_k \ne 0$, then $u^k \in \Omega$. For, by the property of the projection onto a convex closed set Q, given any $v \in Q$ and $u = u^k - C(u^k)$, we have

$$(u^k - \alpha_k C(u^k) - \bar{u}^k, v - \bar{u}^k) \le 0$$

Hence it follows that, with $u^k = \bar{u}^k$, we have $(C(\bar{u}^k), v - \bar{u}^k) \ge 0$ for any $v \in Q$. This means that $\bar{u}^k \in \Omega$.

We now take the set

$$\hat{R}_1(u^*) = \{\bar{u} \mid \bar{u} = proj_Q(u - \tilde{\alpha}\tilde{b}_1), u \in R_1(u^*), \tilde{b}_1 \in M_1\}$$

where

$$M_1 = \{z \mid z \in E^n, \|z\| \le \max_u \|C(u)\|, u \in R_1(u^*)\}, \ 0 \le \tilde{\alpha} \le \bar{\alpha}$$

The set $\hat{R}_1(u^*) \subset \hat{R}_0(u^*)$, since $R_1(u^*) \subset R_0(u^*)$, while

$$\max_{\tilde{u}} \|C(\tilde{u})\| \le \max_u \|C(u)\| \quad for \ \tilde{u} \in R_1(u^*) \ and \ u \in R_0(u^*)$$

On repeating the above procedure for $k = 2, 3, ...,$ we find that, at each iteration, α_k can be chosen in accordance with condition

$$0 < \alpha_k \le \min\{\bar{\alpha}, \varepsilon \frac{\left\|u^k - \bar{u}^k\right\|}{\left\|C(u^k) - C(\bar{u}^k)\right\|}\} \tag{7.6}$$

in such a way that

$$R_k(u^*) \subset ... \subset R_0(u^*) \text{ and } \hat{R}_k(u^*) \subset ... \subset \hat{R}_0(u^*) \ \forall k$$

where

$$\hat{R}_k(u^*) = \{\bar{u} \mid \bar{u} = proj_Q(u - \tilde{\alpha}\tilde{b}_k), u \in R_k(u^*), \tilde{b}_k \in M_k\}$$

while

$$M_k = \{z \mid z \in E^n, \|z\| \le \max\|C(u)\|, u \in R_k(u^*)\}$$

Since the $\hat{R}_k(u^*)$, $k = 0,1,...,$ are bounded, there exist $L_k < \infty$ such that

$\left\|C(u^k) - C(\bar{u}^k)\right\| \le L_k \left\|u^k - \bar{u}^k\right\|$ for all u^k and $\bar{u}^k \in \hat{R}_k(u^*)$ and since

$\hat{R}_k(u^*) \subset ... \subset \hat{R}_0(u^*)$, then $L_0 \ge ... \ge L_k \ge ...$

There is thus a constant $\hat{\alpha} = \min\{\bar{\alpha}, \varepsilon / L_0\}$, $0 < \varepsilon < 1$, such that, at each iteration, we can choose α_k in accordance with eq. (6.6) in such a way that

$$0 < \hat{\alpha} \le \alpha_k \le \min\{\bar{\alpha}, \varepsilon \frac{\left\|u^k - \bar{u}^k\right\|}{\left\|C(u^k) - C(\bar{u}^k)\right\|}\} \tag{7.7}$$

Then from eq. (6.4) with $u^k \ne \bar{u}^k$, we have

$$\left\|u^{k+1} - u^*\right\|^2 \le \left\|u^k - u^*\right\|^2 - \left\|u^k - \bar{u}^k\right\|^2 (1 - \alpha_k^2 \frac{\left\|C(u^k) - C(\bar{u}^k)\right\|^2}{\left\|u^k - \bar{u}^k\right\|^2}) \tag{7.8}$$

which along with eq. (6.7) and the inclusions $R_k(u^*) \subset ... \subset R_0(u^*)$, shows that, as $k \to \infty$, either $u^{k+1} \to u^*$ or else $\left\|u^k - \bar{u}^k\right\| \to 0$. If the first alternative is true, then theorem is proved. If the second is true and $\alpha_k \ge \alpha$, we find, in accordance with the property of

the projection eq. (6.1), that $(u^k - \alpha_k C(u^k) - \bar{u}^k, v - \bar{u}^k) \le 0$ for any $v \in Q$. Hence, as $k \to \infty$, for any $v \in Q$, we have $(C(u^k), v - u^k) \ge 0$. This means that $\hat{u} \in \Omega$, where

$$\hat{u} = \lim_{k \to \infty} u^k$$

The theorem is proved.

B. Asymmetric matrix A for Nguyen-Dupuis Network

$$A = \begin{pmatrix}
18 & 10 & 5 & 0 & 1 & 1.2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 19 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.2 & 0.1 & 0 \\
2 & 0 & 15 & 9 & 2 & 1.1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 5 & 11 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0.3 & 0.1 & 0 & 0 & 0 & 0 & 0 & 0 \\
3 & 0 & 2 & 0 & 13.5 & 6 & 0.4 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
2 & 0 & 1 & 0.9 & 2 & 17.5 & 0 & 0 & 0 & 0 & 0 & 0.4 & 0.2 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0.6 & 0 & 12.5 & 0.1 & 4 & 4.4 & 0 & 0 & 0 & 0 & 0 & 0 & 0.3 & 0 & 0 \\
0 & 0 & 0 & 0 & 0.5 & 0 & 2 & 5.5 & 0 & 0 & 0 & 1.1 & 0 & 0.1 & 0 & 0 & 0.4 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0.5 & 0 & 13.5 & 1 & 0.6 & 0 & 0 & 0 & 0 & 0 & 0 & 5 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0.2 & 0 & 5 & 33.3 & 0 & 0 & 0 & 2 & 13 & 12 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.2 & 0 & 12.5 & 0 & 0 & 0 & 3 & 0 & 0 & 2 & 0 \\
0 & 0 & 0 & 0.5 & 0 & 4 & 0 & 2 & 0 & 0 & 0 & 9.8 & 2 & 0.3 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0.1 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 1 & 5 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.3 & 0 & 10 & 0 & 3 & 0 & 37 & 0.1 & 14 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.1 & 11 & 0 & 0 & 0.2 & 20 & 0.5 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.2 & 0 & 0 & 0 & 0.1 & 2 & 50 & 0 & 0 & 6 \\
0 & 4 & 0 & 0 & 3 & 0 & 0.3 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 20 & 1 & 0 \\
0 & 4 & 0 & 0 & 0 & 0 & 0 & 0 & 4 & 0 & 0.3 & 0 & 0 & 0 & 0 & 0 & 2 & 11 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 5 & 0 & 0 & 10
\end{pmatrix}$$

$$h = \begin{pmatrix} 7 & 9 & 9 & 12 & 3 & 9 & 5 & 13 & 5 & 9 & 9 & 10 & 9 & 6 & 9 & 8 & 7 & 14 & 11 \end{pmatrix}$$

C. Equilibrium solutions with different degrees of asymmetric mode interactions

direction	Link sequence	Path costs and path flows							
		$\rho_{md} = 0.2$				$\rho_{md} = 0.3$			
		Cars		Trucks		Cars		Trucks	
Origin 1 to destination 2	2 18 11	70.02	300.00	55.68	100.00	70.03	300.00	57.49	100.00
	2 17 7 9 11	86.90	0.00	67.32	0.00	86.94	0.00	69.80	0.00
	2 17 8 14 15	91.80	0.00	72.50	0.00	91.83	0.00	74.94	0.00
	1 5 7 9 11	87.37	0.00	67.32	0.00	87.35	0.00	69.80	0.00
	1 6 12 14 15	109.75	0.00	83.91	0.00	109.74	0.00	87.12	0.00
	1 5 8 14 15	92.27	0.00	72.50	0.00	92.24	0.00	74.94	0.00
	2 17 7 10 15	88.95	0.00	69.65	0.00	88.97	0.00	72.07	0.00
	1 5 7 10 15	89.42	0.00	69.65	0.00	89.38	0.00	72.07	0.00
Origin 1 to destination 3	2 17 8 14 16	95.96	0.00	75.19	0.00	95.97	0.00	77.80	0.00
	2 17 7 10 16	93.11	130.34	72.34	24.77	93.11	130.34	74.93	25.00
	1 5 7 10 16	93.58	0.00	72.34	21.58	93.52	0.00	74.93	21.25
	1 5 8 14 16	96.43	0.00	75.19	0.00	96.38	0.00	77.80	0.00
	1 6 12 14 16	113.91	0.00	86.60	0.00	113.88	0.00	89.98	0.00
	1 6 13 19	93.11	469.66	72.34	153.65	93.11	469.66	74.93	153.76
Origin 4 to destination 2	3 5 7 9 11	77.40	215.92	61.53	81.59	77.39	215.89	63.50	81.64
	3 5 7 10 15	79.45	0.00	63.85	0.00	79.42	0.00	65.77	0.00
	3 5 8 14 15	82.30	0.00	66.70	0.00	82.28	0.00	68.63	0.00
	3 6 12 14 15	99.78	0.00	78.11	0.00	99.78	0.00	80.82	0.00
	4 12 14 15	77.41	234.08	61.52	68.41	77.40	234.11	63.50	68.36
Origin 4 to destination 3	3 5 7 10 16	83.61	0.00	66.54	0.00	83.56	0.00	68.63	0.00
	3 5 8 14 16	86.46	0.00	69.39	0.00	86.42	0.00	71.49	0.00
	3 6 12 14 16	103.94	0.00	80.80	0.00	103.92	0.00	83.67	0.00
	3 6 13 19	83.14	0.00	66.54	0.00	83.15	0.00	68.63	0.00
	4 12 14 16	81.57	0.00	64.21	0.00	81.55	0.00	66.36	0.00
	4 13 19	60.77	150.00	49.95	50.00	60.77	150.00	51.31	50.00

Appendix C continued

direction	Link sequence	Path costs and path flows							
		$\rho_{md} = 0.4$				$\rho_{md} = 0.5$			
		Cars		Trucks		Cars		Trucks	
Origin 1 to destination 2	2 18 11	70.05	300.00	59.30	100.00	70.07	300.00	61.11	100.00
	2 17 7 9 11	86.97	0.00	72.28	0.00	87.00	0.00	74.76	0.00
	2 17 8 14 15	91.85	0.00	77.37	0.00	91.87	0.00	79.81	0.00
	1 5 7 9 11	87.32	0.00	72.28	0.00	87.29	0.00	74.76	0.00
	1 6 12 14 15	109.72	0.00	90.33	0.00	109.70	0.00	93.55	0.00
	1 5 8 14 15	92.20	0.00	77.37	0.00	92.17	0.00	79.81	0.00
	2 17 7 10 15	88.98	0.00	74.50	0.00	88.99	0.00	76.92	0.00
	1 5 7 10 15	89.33	0.00	74.50	0.00	89.28	0.00	76.92	0.00
Origin 1 to destination 3	2 17 8 14 16	95.98	0.00	80.40	0.00	95.99	0.00	83.00	0.00
	2 17 7 10 16	93.11	130.34	77.52	25.22	93.10	130.35	80.12	25.45
	1 5 7 10 16	93.46	0.00	77.52	20.92	93.40	0.00	80.12	20.59
	1 5 8 14 16	96.33	0.00	80.40	0.00	96.28	0.00	83.00	0.00
	1 6 12 14 16	113.85	0.00	93.36	0.00	113.81	0.00	96.74	0.00
	1 6 13 19	93.11	469.66	77.52	153.86	93.10	469.65	80.12	153.97
Origin 4 to destination 2	3 5 7 9 11	77.38	215.87	65.47	81.69	77.37	215.82	67.45	81.74
	3 5 7 10 15	79.39	0.00	67.69	0.00	79.36	0.00	69.61	0.00
	3 5 8 14 15	82.26	0.00	70.56	0.00	82.24	0.00	72.49	0.00
	3 6 12 14 15	99.78	0.00	83.52	0.00	99.78	0.00	86.23	0.00
	4 12 14 15	77.39	234.13	65.47	68.31	77.38	234.18	67.45	68.26
Origin 4 to destination 3	3 5 7 10 16	83.52	0.00	70.72	0.00	83.47	0.00	72.81	0.00
	3 5 8 14 16	86.39	0.00	73.59	0.00	86.35	0.00	75.69	0.00
	3 6 12 14 16	103.91	0.00	86.55	0.00	103.89	0.00	89.43	0.00
	3 6 13 19	83.17	0.00	70.72	0.00	83.18	0.00	72.81	0.00
	4 12 14 16	81.52	0.00	68.50	0.00	81.49	0.00	70.64	0.00
	4 13 19	60.78	150.00	52.67	50.00	60.78	150.00	54.02	50.00

Appendix C continued

direction	Link sequence	Path costs and path flows							
		$\rho_{md} = 0.6$				$\rho_{md} = 0.7$			
		Cars		Trucks		Cars		Trucks	
Origin 1 to destination 2	2 18 11	70.08	300.00	62.91	100.00	70.10	300.00	64.72	100.00
	2 17 7 9 11	87.03	0.00	77.24	0.00	87.06	0.00	79.72	0.00
	2 17 8 14 15	91.90	0.00	82.24	0.00	91.92	0.00	84.68	0.00
	1 5 7 9 11	87.26	0.00	77.24	0.00	87.24	0.00	79.72	0.00
	1 6 12 14 15	109.69	0.00	96.76	0.00	109.67	0.00	99.97	0.00
	1 5 8 14 15	92.13	0.00	82.24	0.00	92.09	0.00	84.68	0.00
	2 17 7 10 15	89.01	0.00	79.35	0.00	89.02	0.00	81.78	0.00
	1 5 7 10 15	89.24	0.00	79.35	0.00	89.19	0.00	81.78	0.00
Origin 1 to destination 3	2 17 8 14 16	95.99	0.00	85.61	0.00	96.00	0.00	88.21	0.00
	2 17 7 10 16	93.10	130.35	82.71	25.67	93.10	130.35	85.31	25.90
	1 5 7 10 16	93.33	0.00	82.71	20.26	93.27	0.00	85.31	19.93
	1 5 8 14 16	96.23	0.00	85.61	0.00	96.17	0.00	88.21	0.00
	1 6 12 14 16	113.78	0.00	100.12	0.00	113.75	0.00	103.50	0.00
	1 6 13 19	93.10	469.65	82.71	154.07	93.10	469.65	85.31	154.18
Origin 4 to destination 2	3 5 7 9 11	77.36	215.76	69.42	81.79	77.35	215.67	71.40	81.86
	3 5 7 10 15	79.33	0.00	71.53	0.00	79.30	0.00	73.45	0.00
	3 5 8 14 15	82.22	0.00	74.42	0.00	82.20	0.00	76.35	0.00
	3 6 12 14 15	99.78	0.00	88.94	0.00	99.78	0.00	91.65	0.00
	4 12 14 15	77.37	234.24	69.42	68.21	77.36	234.33	71.40	68.14
Origin 4 to destination 3	3 5 7 10 16	83.43	0.00	74.89	0.00	83.38	0.00	76.98	0.00
	3 5 8 14 16	86.32	0.00	77.79	0.00	86.28	0.00	79.88	0.00
	3 6 12 14 16	103.88	0.00	92.30	0.00	103.86	0.00	95.18	0.00
	3 6 13 19	83.19	0.00	74.89	0.00	83.21	0.00	76.98	0.00
	4 12 14 16	81.47	0.00	72.78	0.00	81.44	0.00	74.93	0.00
	4 13 19	60.79	150.00	55.38	50.00	60.79	150.00	56.73	50.00

Appendix C continued

direction	Link sequence	Path costs and path flows							
		$\rho_{md} = 0.8$							
		Cars		Trucks		Cars		Trucks	
Origin 1 to destination 2	2 18 11	70.12	300.00	66.53	100.00				
	2 17 7 9 11	87.09	0.00	82.20	0.00				
	2 17 8 14 15	91.94	0.00	87.11	0.00				
	1 5 7 9 11	87.21	0.00	82.20	0.00				
	1 6 12 14 15	109.65	0.00	103.18	0.00				
	1 5 8 14 15	92.06	0.00	87.11	0.00				
	2 17 7 10 15	89.03	0.00	84.20	0.00				
	1 5 7 10 15	89.15	0.00	84.20	0.00				
Origin 1 to destination 3	2 17 8 14 16	96.01	0.00	90.81	0.00				
	2 17 7 10 16	93.10	130.36	87.90	26.12				
	1 5 7 10 16	93.21	0.00	87.90	19.60				
	1 5 8 14 16	96.12	0.00	90.81	0.00				
	1 6 12 14 16	113.72	0.00	106.88	0.00				
	1 6 13 19	93.10	469.64	87.90	154.28				
Origin 4 to destination 2	3 5 7 9 11	77.34	215.48	73.37	81.95				
	3 5 7 10 15	79.27	0.00	75.37	0.00				
	3 5 8 14 15	82.18	0.00	78.28	0.00				
	3 6 12 14 15	99.78	0.00	94.36	0.00				
	4 12 14 15	77.35	234.52	73.37	68.05				
Origin 4 to destination 3	3 5 7 10 16	83.34	0.00	79.07	0.00				
	3 5 8 14 16	86.25	0.00	81.98	0.00				
	3 6 12 14 16	103.84	0.00	98.05	0.00				
	3 6 13 19	83.22	0.00	79.07	0.00				
	4 12 14 16	81.41	0.00	77.07	0.00				
	4 13 19	60.79	150.00	58.09	50.00				