Logit-based transit assignment: Approach-based formulation and paradox revisit

S. Sun and W.Y. Szeto[1]
Department of Civil Engineering, The University of Hong Kong, Hong Kong, P. R. China
[1] Phone: (852) 28578552, Fax: (852) 25595337, Email: ceszeto@hku.hk

ABSTRACT

This paper proposes an approach-based transit assignment model under the assumption of logit-based stochastic user equilibrium (SUE) with fixed demand. This model is proven to have a unique solution. A cost-averaging version of the self-regulated averaging method (SRAM) is developed to solve the proposed approach-based SUE transit assignment problem. It is proven that the algorithm converges to the model solution. Numerical examples with discussions are presented to investigate the model properties, a paradoxical phenomenon due to the stochastic nature of the model, capacity paradox, and the performance of the proposed algorithm. The sensitivity analysis of different model and algorithm parameters are performed. A performance comparison between the cost-averaging SRAM, the flow-averaging SRAM, and the method of successive averages is made. The proposed methodology is demonstrated to be able to solve the Winnipeg transit network.

Keywords: transit assignment; logit-based stochastic user equilibrium; paradox; approach-based formulation.

1. INTRODUCTION

The transit assignment problem has received considerable attention, as finding solutions to this problem is essential for planning, designing, controlling, and managing transit networks and evaluating transit system performance. This problem requires determining passenger flows in transit networks, which depends on the underlying assumptions of the route choice behavior of passengers. Earlier transit assignment models assume that passengers select routes based on Wardrop's user equilibrium (UE) principle. Most of the transit assignment models in the literature are also user equilibrium-based models (e.g., Wu et al., 1994; Poon et al., 2004; Hamdouch et al., 2011; Schmöcker et al., 2011; Sun et al., 2013; Trozzi et al., 2013; Verbas et al., 2016; Binder et al., 2017).

Daganzo and Sheffi (1977) and Fisk (1980) extended Wardrop's UE principle to the stochastic user equilibrium (SUE) principle to capture the random effect in travelers' route choice behavior. In contrast with the perfect information assumption in the UE principle, the SUE principle assumes that passengers may not know the precise travel time of available routes and make route choice decisions based on their perceived travel time. This extension is more realistic and leads to the development of SUE transit assignment models (e.g., Nielsen, 2000; Nielsen and Frederiksen, 2006; Liu and Meng, 2014).

Traditionally, transit assignment problems are either formulated as link-based models (e.g., Wu et al., 1994; Kurauchi et al., 2003; Cepeda et al., 2006; Hamdouch and Lawphongpanich, 2008; Hamdouch et al., 2014; Codina and Rosell, 2017) or path-based models (e.g., Lam et al., 1999; Wu and Lam, 2003; Teklu, 2008; Li et al., 2010; Szeto et al., 2011, 2013; Cats et al., 2016; Nuzzolo et al., 2016) depending on the presented form of passenger flows (i.e., link flow or path flow variables). Link-based models can be solved without knowing the path set and hence path set generation heuristic and time-consuming path enumeration procedures can be avoided during solution processes. Solving the link-based formulations directly can also be quicker than solving path-based formulations with path enumeration and guarantee convergence to obtain solutions to transit assignment models for practical size networks. However, path flow information cannot be obtained using link-based formulations. This information is useful to determine the impact of path-specific cost (or cost-saving) for a group of passengers. For example, it is common in Hong Kong that there is a fare discount when a passenger transfers from one specific transit line to another. This fare discount can be modeled to be a path-specific cost saving.

In contrast to link-based models, path-based models can provide path flow information for passengers' route choice behavior modeling which allows the modeler to evaluate the impact of path-specific cost to a specific group of passengers. Moreover, solution methods for logit-based stochastic assignment problems with a considerably faster convergence rate than those for deterministic counterparts can be easily applied to solve the path-based formulations. However, the use of path-based methods requires an explicit path set, which can be obtained by path enumeration. The process of path set enumeration can be very time-consuming for practical size transit networks. As a result, the path set generation approach is commonly used instead of path set enumeration to solve most of the path-based models. This approach only generates paths when needed and unused paths are deleted. However, this approach is heuristic and cannot guarantee convergence. Recently, efficient methods, such as event dominance (Florian, 1998, 2004) or equilibrated choice sets (Watling et al., 2015; Rasmussen et al., 2015), have been developed to overcome these issues and have been applied to commercial software packages (e.g., Emme).

In order to retain the advantages of link-based models while retaining the path choice information used in *traffic assignment* models, Long et al. (2013) proposed an alternative methodology. They proposed the approach-based formulation for their deterministic dynamic traffic assignment problem. In their formulation, approach proportions are used to describe traffic movements in the network and are used as decision variables. (An approach proportion associated with a link emanated from a node is defined as the probability of the link chosen by traffic flows at that node.) In this approach-based formulation, the path flow information is implicitly included in the formulation and can be obtained through a forward pass method. This formulation approach has only been applied to very limited transit assignment studies. For example, Szeto and Jiang (2014) proposed the approach-based formulation of the UE transit assignment problem; Jiang and Szeto (2016) also formulated their reliability-based stochastic transit assignment problem as an approach-based transit assignment problem. However, *extensions to SUE transit assignment have not been found*.

In this paper, we propose an approach-based logit SUE transit assignment model, which can be formulated as a fixed-point (FP) problem in terms of approach proportions (or namely approach probabilities). In our proposed model, the approach proportions are destination specific. Compared with an origin-destination based model, the number of decision variables in our model is reduced significantly. We also prove that our model has a unique solution.

SUE transit assignment models are usually solved by the techniques for FP problems including the method of successive averages (MSA) (e.g., Wu and Lam, 2003; Nielsen and Frederiksen, 2006; Sumalee et al., 2009). The MSA adopts a fixed and predetermined step size during the solution process and is known to have a slow convergence rate. Regarding this issue, Liu et al. (2009) proposed a self-regulated averaging method (SRAM) for traffic assignment problems, which in contrast adopted varying step sizes during the solution process to improve the convergence rate. Long et al. (2014) further reformed the traditional flow-averaging SRAM into a cost-averaging version to solve their traffic assignment model. In contrast with the original flow-averaging version of the SRAM, the cost-averaging version solves the FP problem formulated in terms of link costs instead of *passenger flows*. However, the application of this cost-averaging method for solving transit assignment problems, including our approach-based problems, has not been reported in the literature. *It is unclear whether the cost-averaging SRAM is more efficient than the traditional flow-averaging SRAM and the traditional MSA to solve transit assignment problems and whether the cost-averaging SRAM is convergent to the solution of the proposed approach-based model.*

In this paper, in order to improve computational efficiency, the convergent cost-averaging version of the SRAM is proposed to solve the approach-based SUE transit assignment model. Moreover, in this paper, the effect of the algorithmic parameters on the speed of convergence is examined. In addition, a performance comparison among the cost-averaging and flow-averaging SRAM and the MSA is made based on numerical examples. The Winnipeg transit network is used to demonstrate the convergence of the cost-averaging SRAM.

The proposed approach-based SUE transit assignment model can be used to evaluate network design strategies and identify possible paradox occurrences. In the literature, limited effort has been spent on the identification and analysis of paradoxical phenomena of transit assignment problems. Cominetti and Correa (2001) presented a paradox on the demand side of transit assignment, showing that a certain range of demand increments may not affect the transit time of the system. Szeto and Jiang (2014) and Jiang and Szeto (2016), on the other hand, presented the Braess-like and capacity paradoxes on the supply side of transit assignment, showing that providing a new transit line or increasing service frequency may not necessarily enhance the system performance in terms of expected total system cost or network capacity/throughput. However, in the aforementioned models, they assume that passengers follow the UE principle; *little attention has been paid to the paradoxical phenomenon caused by the stochastic nature of SUE transit assignment. It is also unclear whether capacity paradox can still be observed under the SUE condition.*

In this paper, we illustrate the paradoxes associated with the stochastic nature of the model as well as passengers' non-cooperative behavior based on numerical examples: adding a new transit line to the network or improving the frequency of an existing transit line in a transit network can cause an increase in expected total system cost and a reduction in network throughput. The occurrences of the two paradoxes are also investigated. We also examine the effect of passengers' perception of travel cost (measured by the value of θ) on the paradoxical phenomena.

Overall, this paper presents an alternative methodology to solve large-scale SUE transit assignment based on the approach-based formulation and the cost-averaging SRAM. This paper also revisits two types of paradoxes that have been found in the UE traffic or transit assignment literature but have not been discussed in the SUE transit assignment literature. This paper enriches the literature by extending the approach-based theory and concept as well as the discussion of paradoxes to SUE transit assignment. Specifically, this paper makes the following contributions.

- It proposes an approach-based transit assignment model under logit-based stochastic user equilibrium; this model is not a simple or straightforward extension of the UE counterpart but is more general than the counterpart; the model is proven to have a unique solution.
- It proposes to use the cost-averaging SRAM to solve the FP problem; the cost-averaging SRAM is proven to be convergent and can solve large transit networks.
- It illustrates the existence of the paradoxes caused by the stochastic nature of the model and passengers' non-cooperative behavior in transit networks in the context of stochastic user equilibrium. It shows that improving a transit route in the network or adding a new transit route to the network may not necessarily improve the performance of the transit system in terms of expected total system cost and network capacity.
- It investigates the occurrences of the two paradoxes and provides insights and suggestions on transit network design to avoid both the occurrence of these paradoxes.
- It demonstrates the effects of different parameters in the proposed model and algorithm.

The remainder of this paper is organized as follows: Section 2 introduces the notations and network presentation used in this paper, followed by the assumptions. Then, the approach-based logit SUE transit assignment model is presented. The solution algorithm proposed to solve the model is detailed in Section 3. In Section 4, various numerical studies are carried out to show the model properties, the occurrence of the two paradoxes as well as the performance of the proposed algorithm. Finally, Section 5 gives the conclusion.

2. MODEL FORMULATION

2.1. Notations

The following notations are used throughout this paper.

L	the set of lines in the transit network;
l	the line index;
S	the set of links;
s	the link index;
b	the approach index;
A_s	the set of attractive lines associated with link s;
A_i^+	the set of links emanating from node i or the set of approaches associated with node i;
A_i^-	the set of links going into node i;
N	the set of nodes (i.e., stops) in the transit network;
i	the node index;
$t(s)$, $h(s)$	the tail and head nodes of link s;
$u(b)$	the underlying link of approach b;
R	the set of origins;
r	the origin index;
D	the set of destinations;
d	the destination index;
P^{rd}	the set of paths associated with origin-destination (O-D) pair rd;
p	the path index;
S^p	the set of links on path p;
S^{id}	the set of links on efficient paths connecting node i and destination d;
t_s, ω_s, ϕ_s	the in-vehicle travel time, waiting time, and perceived congestion time of link s;
t_s^l, w_s^l	the in-vehicle travel time and relative frequency of line l associated with link s;
f^l, κ^l	the frequency and capacity of a single vehicle of line l;
L_s^{id}	the likelihood of passengers who use link s and travel from node i to destination d;
L_b^d	the likelihood of passengers who use approach b and travel to

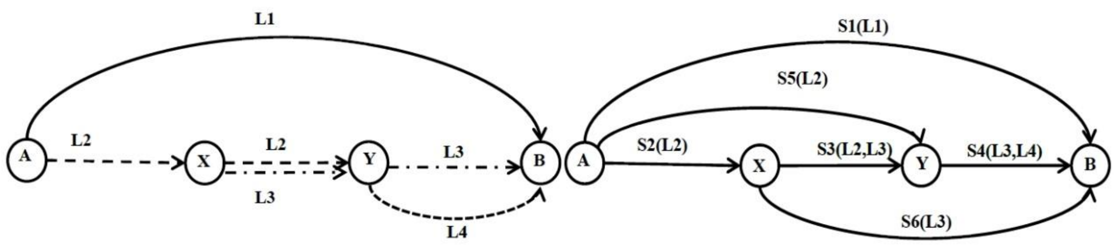
	destination d;
π^{id}	the minimum travel cost for passengers from node i to destination d;
π_s^{id}	the minimum travel cost for passengers from node i to destination d using link s;
W_s^{id}	the weight for passengers who enter link s and travel from node i to destination d;
W_b^d	the weight for passengers who choose approach b and travel to destination d;
q^{rd}	passenger demand from origin r to destination d;
v_s	passenger flow on link s;
y_p^{rd}	passenger flow on path p associated with O-D pair rd;
v_{sl}^d	passenger flow on link s to destination d through line l;
$\overline{v}_s$	passenger flow on the competing links of link s;
v_s^{id}	passenger flow on link s from node i to destination d;
v_s^d	passenger flow on link s towards destination d;
$\mathbf{v}$	the vector $\left(v_s^d\right)$ with a dimension of $\lvert S\rvert\times\lvert D\rvert$;
a_s^{id}	the choice probability of link s chosen by passengers traveling from node i to destination d;
a_s^d	the choice probability of link s chosen by passengers traveling to destination d;
$\mathbf{a}$	the vector $\left(a_s^d\right)$ with a dimension of $\lvert S\rvert\times\lvert D\rvert$;
α_b^d	the approach probability of passengers using approach b and traveling to destination d;

α_p^{rd}	the probability of passengers using path p traveling from origin r to destination d;
$\boldsymbol{\alpha}$	the vector $\left(\alpha_b^d\right)$ with a dimension of $\lvert S\rvert\times\lvert D\rvert$;
c_s	the total expected travel cost of link s;
c_p	the total expected travel cost of path p;
$\mathbf{c}$	the vector $\left(c_s\right)$ with a dimension of $\lvert S\rvert$;
μ_{T} (μ_{W})	the value of in-vehicle travel time (waiting time);
σ	a unit conversion parameter;
δ', ε, n	the parameters of the congestion cost function;
ϖ_s	the parameter of the congestion cost function associated with link s;
θ	the positive dispersion parameter that reflects an aggregate measure of passengers' perception of travel cost, the parameter of logit-based SUE model;
k	the iteration number in the SRAM;
η, γ	the parameters of step sizes in the SRAM;
β^k	the parameter of the step size at the k-th iteration in the SRAM;
λ^k	the step size at the k-th iteration in the SRAM;
$\mathbf{h}^k$	the vector of descent directions at the k-th iteration in the SRAM.

2.2. Network representations

A transit network consists of a set of nodes (i.e., transit stops) and a set of transit routes (i.e., transit lines) serving the network. Passengers can board or alight at transit stops. Following the definition of attractive lines specified by De Cea and Fernández (1993), passengers at each stop are classified into groups according to their alighting stops. Then, a link (also called route section) is created for passengers boarding and alighting at the same pair of transit stops. Transit routes providing direct services between that pair of stops (i.e., common lines) are associated with the link connecting the pair of stops. As a result, a link-route representation of the transit network is developed.

For illustration purposes, a small sample transit network is presented in Figure 1(a). The link-route representation of this network is shown in Figure 1(b). There are four transit stops (i.e., nodes A, B, X, and Y) and four transit lines (i.e., L1 to L4) in the network. There are 6 different pairs of stops in total (i.e., A-B, A-X, X-Y, Y-B, A-Y, and X-B) which are connected by links S1-S6, respectively. Take S1 and S3 as examples, only L1 provides a direct service between nodes A and B, and therefore L1 is associated with S1 as marked in the brackets next to S1 in Figure 1(b). For S3 connecting nodes X and Y, both L2 and L3 provide direct services, and therefore they are both associated with S3 as marked in the brackets next to S3 in Figure 1(b) and they are common lines of S3. This link-route representation considers only attractive lines (i.e., excluding routes with exceptionally high cost). The set of attractive lines associated with each link can be determined by a method proposed by Chriqui and Robillard (1975).



(a) A small transit network (b) Link-route representation

Figure 1. A small transit network and its link-route representation (De Cea and Fernández (1993))

2.3. Assumptions

As in the literature (e.g., Spiess and Florian, 1989; De Cea and Fernández, 1993; Lam et al., 1999), the following classical assumptions are made throughout this paper. A1) Passengers are assumed to arrive at transit stops randomly. A2) A passenger waiting at a transfer node considers a set of attractive lines before boarding, and he/she boards the first arriving bus if possible. A3) The waiting time for a transit line on a link is independent of the waiting times for other lines on the same link. A4) Vehicle headways are assumed to follow an exponential distribution. A5) Passengers' route choice behaviors are in a stochastic manner, which means that the stochastic user equilibrium condition is assumed. Passengers make their route choice according to their perception of the total expected travel cost of each feasible route (path). A6) The travel demand between each origin-destination (OD) pair in the system is assumed to be known and fixed. This assumption is reasonable for strategic planning when the day-to-day variation, especially during the peak hour period, is small or negligible. A7) For simplicity, the capacity of each transit vehicle is assumed to be the same. However, there is no conceptual difficulty in extending the formulation to a scenario in which vehicles of different capacities traverse different routes. A8) When passengers making route choice decisions, only "efficient paths" as defined by Ran and Boyce (1996) are considered. The definition of "efficient paths" is given by the following:

Definition D1. *A path between an OD pair is efficient if it includes only links that take travelers closer to the destination.*

It has been proved by Long et al. (2015) that, under this assumption, the sub-network between each OD pair is acyclic.

2.4. Cost components

Based on the preceding notations, link-route representation, and assumptions, the cost components, and the formulations are presented below.

Three cost components are included in the expected total cost: mean in-vehicle travel time cost, mean waiting time cost, and perceived congestion cost. For link s, the expected total travel cost is given by

$$c_s = \mu_{\mathrm{T}} t_s + \mu_{\mathrm{W}} \omega_s + \mu_{\mathrm{W}} \phi_s, \quad \forall s \in S, \tag{1}$$

where μ_{T} and μ_{W} are the corresponding values of time.

The formulations of the three cost components are described individually in the following subsections.

2.4.1. *Mean in-vehicle travel time cost*

The mean in-vehicle travel time of link s is defined to be the weighted sum of the in-vehicle travel time of all the attractive lines associated with link s as

$$t_s = \sum_{l \in A_s} w_s^l t_s^l, \quad \forall s \in S, \tag{2}$$

where w_s^l is defined by

$$w_s^l = \frac{f^l}{\sum_{j \in A_s} f^j}, \quad \forall s \in S, l \in A_s. \tag{3}$$

The value of in-vehicle travel time μ_{T} is multiplied by t_s to obtain the mean in-vehicle travel time cost.

2.4.2. Mean waiting time cost

The mean waiting time of link s is defined to be the waiting time for the first arriving vehicle from the set of attractive lines associated with link s. Under assumptions A1) – A4), the mean waiting time of link s is given by

$$\omega_s = \frac{\sigma}{\sum_{l \in A_s} f_s^l}, \quad \forall s \in S. \tag{4}$$

In this case, the unit of frequency is veh/hr; and that of waiting time is in minutes. σ = 60 min/hr. Based on Eq. (4), we can obtain the mean waiting time cost $\mu_{\mathrm{W}}\omega_s$.

2.4.3. Perceived congestion cost

The congestion cost function approach is one of the main approaches in the literature used to model the additional waiting time caused by insufficient vehicle capacity on a link. The perceived congestion delay of link s is given by a function of its link flow as well as link flows from a set of its competing links.

The set of competing links of link s consists of links associated with two groups of passengers. One group of passengers is the group of passengers who board before the tail node of link s, $t(s)$, and alight after the head node of link s, $h(s)$, using at least one of the attractive lines associated with link s. These passengers occupy vehicle spaces of the attractive lines of link s and therefore contribute to congestion delays experienced by passengers on link s. As a result, links associated with this group of passengers are included in the set of competing links of link s.

The other group of passengers is the group of passengers who board at $t(s)$ and alight after $h(s)$, using at least one of the attractive lines associated with link s. These passengers directly compete with passengers on link *s* for boarding vehicles from the attractive lines. Therefore, links associated with this group of passengers are included in the set of competing links of link *s*.

Let v_s be the passenger flow on link *s* in pass/hr, and $\bar{v}_s$ be the passenger flow contributed by the competing links of link *s* in pass/hr. Then, v_s and $\bar{v}_s$ can be respectively calculated by

$$v_s = \sum_{d \in D} \sum_{l \in A_s} v_{sl}^d, \quad \forall s \in S \quad \text{and} \tag{5}$$

$$\bar{v}_s = \sum_{d \in D} \sum_{m \in S, m \neq s} \delta_s^m \sum_{l \in A_s \cap A_m} v_{ml}^d, \quad \forall s \in S, \tag{6}$$

where $\delta_s^m = 1$ if *m* is a competing link of link *s*, = 0 otherwise.

Passenger flows on link *s* and competing links of link *s* are contributed by flows on attractive lines associated with link *s* and competing links of link *s* using attractive lines of link *s* toward all destinations as described in Eqs. (5) and (6), respectively.

The line flow v_{sl}^d is calculated by

$$v_{sl}^d = v_s^d w_s^l, \quad \forall s \in S, l \in A_s, d \in D. \tag{7}$$

Eq. (7) states that the flows on link *s* are distributed to the transit lines on that link based on the relative frequencies determined by Eq. (3).

The additional waiting time of link *s* due to congestion can be calculated by the congestion cost function proposed by Szeto and Jiang (2014) as

$$\phi_s = \varpi_s \left(\frac{\delta' v_s + \varepsilon \bar{v}_s}{\sum_{l \in A_s} f^l \kappa^l} \right)^n, \quad \forall s \in S, \tag{8}$$

where the denominator is interpreted as the capacity of link *s*; the parameters δ', ε, ϖ_s, and *n* are calibration parameters used to model different effects contributed by various passenger flows. These parameters are related to the passengers' perceptions of the level of congestion. A larger value means that the congestion level has a higher effect on the travel cost of passengers, leading to a higher congestion cost for a given ratio of flow to capacity. Based on Eq. (8), the congestion cost function can be expressed as $\mu_{\mathrm{W}}\phi_s$.

2.5. The approach-based SUE transit assignment model

Traditionally, transit assignment problems are either formulated as link-based models (e.g., Cepeda et al., 2006; Hamdouch and Lawphongpanich, 2008; Hamdouch et al., 2014) or path-based models (e.g., Lam and Zhou, 1999; Poon et al., 2004; Szeto et al., 2013), in which the link- and path-based models adopt link and path flows as decision variables, respectively. This section develops a link-based model for SUE transit assignment based on the theory of Dial's (1971) STOCH algorithm and reformulates the link-based model into an approach-based model.

2.5.1. The link-based model

Based on the theory of Dial's (1971) STOCH algorithm, the likelihood of a link *s* to be chosen by a passenger traveling from node *i* to destination *d* is expressed as

$$L_s^{id} = \begin{cases} \exp\left(\theta(\pi^{t(s)d} - \pi_s^{t(s)d})\right), & \text{if } s \in S^{id} \\ 0, \quad \text{if } s \notin S^{id} \end{cases}, \quad \forall s \in S, i \in N, d \in D, \tag{9}$$

where

$$\pi_s^{t(s)d} = c_s + \pi^{h(s)d}, \quad \forall s \in S, d \in D. \tag{10}$$

A backward pass method is used to calculate the weight of each link:

$$W_s^{id} = L_s^{id}\left(\delta_{h(s)}^d + \sum_{m \in A_{h(s)}^+} W_m^{id}\right), \quad \forall s \in S, i \in N, d \in D, \tag{11}$$

where $\delta_{h(s)}^d = 1$ if $h(s) = d$; $\delta_{h(s)}^d = 0$ otherwise.

The choice probability of a link *s* chosen by passengers traveling from node *i* to destination *d* is given as

$$a_s^{id} = \frac{W_s^{id}}{\sum_{m \in A_{t(s)}^+} W_m^{id}}, \quad \forall s \in S, i \in N, d \in D. \tag{12}$$

Note that the likelihoods, weights, and choice probabilities of all unused links are equal to zero. According to Eq. (12), we have $0 \le a_s^{id} \le 1$ and $\sum_{s \in A_{t(s)}^+} a_s^{id} = 1$ for $\forall s \in S, i \in N, d \in D$.

The passenger flow of link *s* from node *i* towards destination *d* equals the choice probability of this link multiplied by the total inflow rate of the tail node, given by

$$v_s^{id} = a_s^{id} \left(q^{t(s)d} + \sum_{m \in A_{t(s)}^-} v_m^{id} \right), \quad \forall s \in S, i \in N, d \in D, \tag{13}$$

and the passenger flow of link *s*, v_s, can be calculated by $v_s = \sum_{i \in N} \sum_{d \in D} v_s^{id}$. Note that $q^{t(s)d} = 0$ if $t(s) \notin R$.

Based on (9), (11), and (12), we can prove the following.

Proposition 1. Under assumption A8, $a_s^{id} = a_s^{jd}$ for $\forall s \in S^{id} \cap S^{jd}, i, j \in N, d \in D$.

The proof can be found in the appendix.

According to Proposition 1, link choice probability is independent of the start point. Then, Eq. (9), (11) to (13) can be rewritten as

$$L_s^d = \begin{cases} \exp\left(\theta(\pi^{t(s)d} - \pi_s^{t(s)d})\right), & \text{if } s \in S^d \\ 0, & \text{if } s \notin S^d \end{cases}, \quad \forall s \in S, d \in D, \tag{14}$$

$$W_s^d = L_s^d \left(\delta_{h(s)}^d + \sum_{m \in A_{h(s)}^+} W_m^d \right), \quad \forall s \in S, d \in D, \tag{15}$$

$$a_s^d = \frac{W_s^d}{\sum_{m \in A_{t(s)}^+} W_m^d}, \quad \forall s \in S, d \in D, \text{ and} \tag{16}$$

$$v_s^d = a_s^d \left(q^{t(s)d} + \sum_{m \in A_{t(s)}^-} v_m^d \right), \quad \forall s \in S, d \in D, \tag{17}$$

where v_s^d is given by $v_s^d = \sum_{i \in N} v_s^{id}$. Note that $q^{t(s)d} = 0$ if $t(s) \notin R$.

The link choice probability a_s^d in (17) is a function of link travel costs, which in turns are functions of link flows. Therefore, the link-based SUE assignment problem can be formulated as an FP problem: to find $\mathbf{v} = \left(v_s^d\right)$ such that

$$\mathbf{v} = \mathbf{f}(\mathbf{v}), \tag{18}$$

where $\mathbf{f}$ is the mapping function defined by Eqs. (1) to (8), (10), (14) to (17).

The FP problem (18) satisfies the solution existence conditions pointed out by Theorem 1 in the study of Cantarella (1997): at least one route is available to each passenger; passenger demand is non-negative; the link cost-flow functions described by Eqs. (1) to (8) are defined over the non-empty set of link flows and take values in the non-empty, compact, and convex set of link costs; the network loading map defined by Eqs. (14) to (17) is defined over the non-empty set of link costs and take values in the non-empty, compact, and convex set of link flows; the link cost-flow functions described by Eqs. (1) to (8) are continuous; the network loading map defined by Eqs. (14) to (17) is an upper semicontinuous point-to-set map. Therefore, the FP problem (18) has at least one solution.

The FP problem (18) also satisfies the solution uniqueness conditions mentioned in Theorem 2 in the study of Cantarella (1997): the link cost-flow functions described by Eqs. (1) to (8) are strictly monotone increasing; the network loading map defined by Eqs. (14) to (17) is monotone non-increasing. There is exactly one solution to the problem.

2.5.2. The approach-based model

The outputs of the link-based model are in terms of passenger link flows, which do not allow determining path flows easily. The path flow information is useful to determine the impact of path-specific cost (or cost-saving) for a group of passengers such as fare discounts resulting from transfers between specific transit lines. In order to overcome this problem while retaining the advantages of the link-based model, we propose to use approach probabilities to formulate the problem. Following Szeto and Jiang (2014), an approach of a node is defined by a link *coming out* from that node. An approach is a link but must be associated with one node. (A link is defined by two nodes but an approach only concerns the tail node.) The approach probability is defined as the probability of an approach to be chosen by passengers leaving the node via that approach. To illustrate the concept of approaches associated with a node, Figure 2 shows the three approaches of node A in the example network shown in Figure 1.

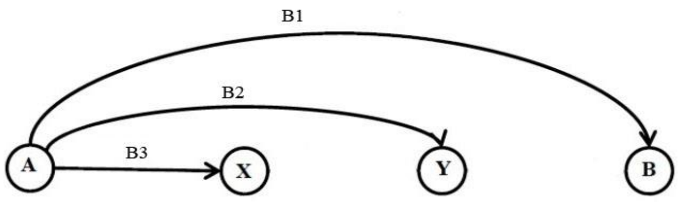


Figure 2. Three approaches of node A in the example network shown in Figure 1

The likelihood of an approach *b* to be chosen by a passenger traveling to destination *d* is expressed as

$$L_b^d = \begin{cases} \exp\left(\theta(\pi^{t(u(b))d} - \pi_{u(b)}^{t(u(b))d})\right), & \text{if } u(b) \in S^d \\ 0, \quad \text{if } u(b) \notin S^d \end{cases}, \quad \forall b \in A_i^+, i \in N, d \in D. \tag{19}$$

A backward pass method is used to calculate the weight of each approach as follows:

$$W_b^d = L_b^d \left(\delta_{h(u(b))}^d + \sum_{m \in A_{h(u(b))}^+} W_m^d \right), \quad \forall b \in A_i^+, i \in N, d \in D, \tag{20}$$

where $\delta_{h(u(b))}^d = 1$ if $h(u(b)) = d$; $\delta_{h(u(b))}^d = 0$ otherwise.

The approach probability of passengers using an approach *b* and heading towards destination *d* is given as

$$\alpha_b^d = \frac{W_b^d}{\sum_{m \in A_{t(u(b))}^+} W_m^d}, \quad \forall b \in A_i^+, i \in N, d \in D. \tag{21}$$

Note that the likelihoods, the weights, and the choice probabilities of all unused approaches are equal to zero. According to Eq. (21), we have $0 \le \alpha_b^d \le 1$ and $\sum_{b \in A_i^+} \alpha_b^d = 1$ for $\forall i \in N, d \in D$.

The passenger flow of link *s* towards destination *d* equals the approach probability of the approach using this link towards destination *d* multiplied by the total inflow rate of the tail node, given by

$$v_s^d = \sum_{b \in A_{t(s)}^+} \left(\delta_s^b \cdot \alpha_b^d\right)\left(q^{t(s)d} + \sum_{m \in A_{t(s)}^-} v_m^d\right), \ \ \forall s \in S, d \in D. \tag{22}$$

where $\delta_s^b = 1$ if $u(b) = s$; $\delta_s^b = 0$ otherwise. Note that $q^{t(s)d} = 0$ if $t(s) \notin R$.

The link flow function $\mathbf{v} = \left(v_s^d\right)$ defined by Eq. (22) is bijective, and $\mathbf{v}$ can be *unilaterally* determined by the approach probability vector $\boldsymbol{\alpha} = \left(\alpha_b^d\right)$. Therefore, passenger flows are functions of approach probabilities; link travel costs are functions of link flows as described by Eqs. (1) to (8); approach probabilities are functions of minimum travel costs towards destinations according to Eqs. (10), (19) to (21), which in turns are functions of link travel costs. Therefore, approach probabilities are functions of themselves, and the approach-based SUE problem can be formulated as an FP problem: to find $\boldsymbol{\alpha}$ such that

$$\boldsymbol{\alpha} = \mathbf{g}(\boldsymbol{\alpha}), \tag{23}$$

where $\mathbf{g}$ is the mapping function defined by (1)-(8), (10), (19)-(21).

As shown in (23), the approach-based SUE transit assignment formulation uses approach proportions as the decision variables that take values between zero and one. This formulation is different from the link-based SUE transit assignment formulation (18), which uses link flows as the nonnegative decision variables. The approach-based problem has a unique solution as stated below.

Proposition 2. The FP problem (23) has exactly one solution.

The proof can be found in the appendix.

A path flow probability can be obtained by multiplying the approach probabilities of their associated links on the path. Let S^p be the set of links on path *p*. According to this definition, the path cost c_p and the probability of path *p* between O-D pair *rd* being used α_p^{rd} can be respectively expressed as

$$c_p = \sum_{s \in S^p} c_s, \ \ \forall p \in P^{rd}, r \in R, d \in D \ \text{ and} \tag{24}$$

$$\alpha_p^{rd} = \prod_{s \in S^p} \left(\sum_{b \in A_{t(s)}^+} \delta_s^b \cdot \alpha_b^d \right), \ \ \forall p \in P^{rd}, r \in R, d \in D. \tag{25}$$

where $\delta_s^b = 1$ if $u(b) = s$; $\delta_s^b = 0$ otherwise. With the preceding notations, we can state the following proposition:

Proposition 3. The solution to the approach-based SUE problem (23) satisfies the logit-based SUE condition: $\alpha_p^{rd} = \dfrac{\exp\left(-\theta \cdot c_p\right)}{\sum_{p' \in P^{rd}} \left(\exp\left(-\theta \cdot c_{p'}\right)\right)}, \quad \forall p \in P^{rd}, r \in R, d \in D$.

The proof of Proposition 3 can be found in the appendix. With the result of Proposition 3, it is logical to have the following finding:

Proposition 4. The approach-based SUE problem (23) is equivalent to the link-based SUE problem (18).

The proof of Proposition 4 can be found in the appendix.

3. SOLUTION ALGORITHM: COST-AVERAGING SRAM

The MSA is a classic solution method for FP problems. However, the most commonly used flow-averaging version of the MSA, which updates *passenger flows* with a predetermined step size at each iteration, is known to have a slow convergence rate. Cantarella (1997) presented the cost-averaging version of the MSA that updates *link costs* at each iteration and discussed its convergence. Liu et al. (2009) proposed the flow-averaging version of the SRAM, which adopted varying step sizes during solution process to improve the convergence rate of the MSA. Long et al. (2014) further reformed the flow-averaging SRAM into a cost-averaging version. In this paper, we propose to use the cost-averaging SRAM (c-SRAM) for solving the proposed approach-based SUE transit assignment problem. The convergence rate of the proposed c-SRAM will be compared with the other three methods for FP problems mentioned above.

As shown in Eq. (19), the likelihoods of approaches to be chosen by passengers traveling to a specific destination are functions of link travel costs. Following Eqs. (20) and (21), the corresponding weights and approach probabilities are also functions of link travel costs. Therefore, we can express the approach probability vector $\boldsymbol{\alpha}$ as a function of the link travel cost vector $\mathbf{c}$, written as

$$\boldsymbol{\alpha} = \mathbf{X}(\mathbf{c}), \tag{26}$$

where $\mathbf{X}$ is the mapping function.

According to Eq. (22), passenger flows are functions of approach probabilities; according to Eq. (8), the link travel costs are functions of passenger flows and thus are also functions of link approach probabilities. The vector form can be written as

$$\mathbf{c} = \mathbf{Y}(\boldsymbol{\alpha}), \tag{27}$$

where $\mathbf{Y}$ is the mapping function.

Substituting Eq. (26) into Eq. (27), we can obtain an FP problem in terms of link travel costs for the SUE transit assignment problem, expressed as

$$\mathbf{c} = \mathbf{Y}(\mathbf{X}(\mathbf{c})). \tag{28}$$

The solution existence and uniqueness of the FP problems (18) and (23) also ensure the solution existence and uniqueness of the FP problem (28) as discussed by Cantarella (1997).

The descent direction denoted by $\mathbf{h}$ can be applied to solve the problem (28), given by

$$\mathbf{h} = \mathbf{Y}(\mathbf{X}(\mathbf{c})) - \mathbf{c}. \tag{29}$$

The cost-averaging version of the SRAM is outlined as follows:

1) Set the initial expected link travel cost equal to its free flow travel cost. Set $k = 1$, $\eta > 1$, $0 < \gamma < 1$, $\beta^0 = 1$, and the convergence tolerance $\varepsilon > 0$.

2) Calculate the interim link approach probability vector $\hat{\boldsymbol{\alpha}}^k = \mathbf{X}(\mathbf{c}^k)$ and the interim expected link travel cost vector $\tilde{\mathbf{c}}^k = \mathbf{Y}(\hat{\boldsymbol{\alpha}}^k)$, and obtain the descent direction $\mathbf{h}^k = \tilde{\mathbf{c}}^k - \mathbf{c}^k$.

3) Obtain the step size $\lambda^k = 1/\beta^k$, where

$$\beta^k = \begin{cases} \beta^{k-1} + \eta, & \text{if } \left\|\mathbf{h}^k\right\| \geq \left\|\mathbf{h}^{k-1}\right\|, \\ \beta^{k-1} + \gamma, & \text{otherwise.} \end{cases}$$

4) Update the expected link travel cost by $\mathbf{c}^{k+1} = \mathbf{c}^k + \lambda^k \mathbf{h}^k$.

5) If $\left\|\mathbf{h}^k\right\| \leq \varepsilon$, stop; otherwise, let $k = k + 1$, and go to step 2).

The proof of convergence of the proposed c-SRAM for solving the approach-based SUE transit assignment problem is similar to Cantarella's (1997) cost-averaging algorithm with the following differences: our problem is approach-based and therefore the convergence is derived based on link cost-approach probability functions instead of link cost-flow functions; the step size of the SRAM is adaptive. Following Theorem 4 in Cantarella (1997), the convergence conditions for the proposed c-SRAM are listed as follows: 1) the link cost-approach probability functions are defined over a non-empty, compact, and convex set; 2) the approach-based FP problem (23) satisfies the solution existence and uniqueness conditions (see Proposition 2); 3) all the choice maps of the problem are additive, probabilistic, continuous with continuous first derivatives over an approach probability space. These three conditions are satisfied by the proposed approach-based SUE transit assignment problem. Moreover, the sequence of step sizes of the SRAM satisfies $\sum_k \lambda^k = \infty$ and $\lim_{k \to \infty} \lambda^k = 0$. Therefore, the cost-averaging SRAM guarantees convergence for solving the proposed SUE transit assignment model.

4. NUMERICAL EXAMPLES

4.1. Model properties

In this section, the properties of the proposed approach-based SUE transit assignment model are illustrated based on the example network as shown in Figure 3. The link characteristics of the example network are listed in Table 1. The following values of parameters are used in this section: the congestion function parameters $\delta' = \varepsilon = n = 1$, $\varpi_s = 10$, $\mu_T = \mu_W = 0.5$ HK$/min.

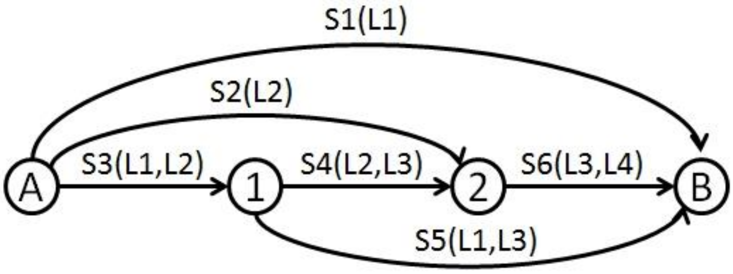


Figure 3. Small example network I

Table 1. The link characteristics of the example network I

	In-vehicle time (min)	Waiting time (min)	Capacity (pass/hr)
S1	67.00	6.00	100
S2	30.00	7.50	80
S3	15.00	3.33	180
S4	10.00	4.29	140
S5	32.50	3.75	160
S6	12.50	5.00	120

4.1.1. Equivalency between the approach-based and path-based solutions

In this example, we show that the solution obtained from the approach-based SUE model is equivalent to that of the path-based SUE model under the same network setting. The values of parameters used in this example are given as follows: demand from A to B = 300 pass/hr; the dispersion parameter $\theta = 0.5$. The solution of the approach-based SUE model is shown in Table 2:

Table 2. The solution of the approach-based SUE model

Link No.	1	2	3	4	5	6
Approach probability	0.35	0.31	0.34	0.10	0.90	1.00
Flow (pass/hr)	105.21	93.03	101.76	10.59	91.17	103.62
Congestion delay (min)	10.52	17.28	16.67	9.84	12.56	11.48
Total cost (HK$)	41.76	27.39	17.50	12.07	24.41	14.49

The path cost c_p and the path use probability α_p^{rd} can be calculated using Eqs. (24) and (25), and the path flow y_p^{rd} between O-D pair *rd* can, therefore, be calculated by

$$y_p^{rd} = \alpha_p^{rd} \cdot q^{rd}, \quad \forall p \in P^{rd}, r \in R, d \in D, \tag{30}$$

where q^{rd} is the passenger demand between O-D pair *rd*. The path flow solution is shown in Table 3. The solution shown in Table 3 satisfies the path-based logit SUE condition given by

$$y_p^{rd} = \frac{\exp(-\theta c_p)}{\sum_{p' \in P^{rd}} \exp(-\theta c_{p'})} \cdot q^{rd}, \quad \forall p \in P^{rd}, r \in R, d \in D. \tag{31}$$

Therefore, the approach-based SUE model is equivalent to the path-based SUE model.

Table 3. The path flow solution

Path No.	1	2	3	4
Link sequence	1	2-6	3-4-6	3-5
Flow (pass/hr)	105.21	93.03	10.59	91.17
Total cost (HK$)	41.76	41.885	44.055	41.9

4.1.2. The effect of passengers' perception of travel cost

In this example, we illustrate the effect of the parameter of passengers' perception of travel cost θ on passengers' route choice behavior. In this example, the demand from A to B q^{AB} = 300 pass/hr.

The results of expected total system cost obtained by varying θ are shown in Figure 4. As shown in this figure, as θ increases, the expected total system cost decreases and approaches a certain limit. It is because when θ is large, the stochastic effect of passengers' perception of path travel time decreases and the stochastic user equilibrium condition approaches the user equilibrium condition. When passengers have perfect information about the network condition, they can choose the most efficient way to their destination, and therefore the expected total system cost drops as θ increases. The implication is that the transit network performance may be underestimated by assuming user equilibrium to be held when passengers have imperfect information about the component of travel cost, such as in-vehicle travel time and congestion costs, especially when they have a very inaccurate perception of travel times or low information quality on the in-vehicle congestion condition.

The results of path flows obtained by varying θ are shown in Figure 5. As shown in this figure, the most attractive path switches from path 1 to path 2 as the value of θ increases. This example illustrates the importance of the choice of the θ value. Even under the same network settings, different θ values can result in different passenger flow patterns. It implies that it is important to accurately calibrate the θ value to determine the flow on each transit line and the utilization of each transit line in order to redesign transit services, e.g., adjusting the frequency of attractive lines, for improving system performance.

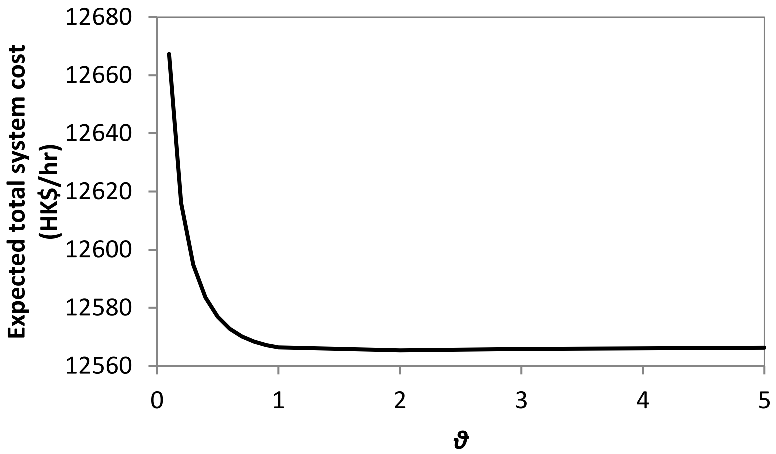


Figure 4. Expected total system cost against θ

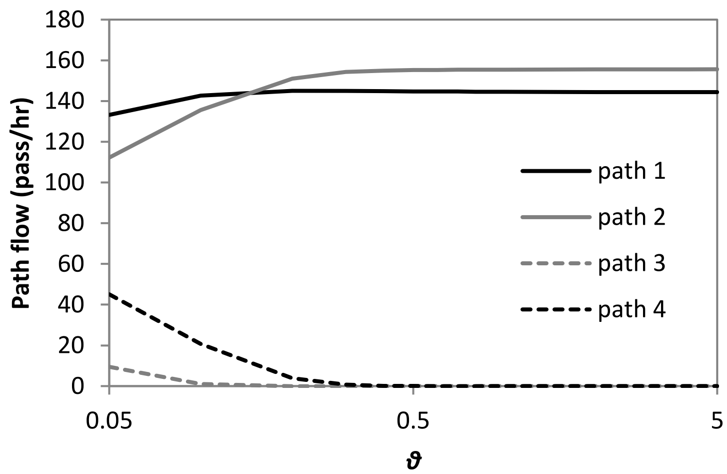


Figure 5. Path flows against θ

4.2. Paradoxical phenomena

Sheffi and Daganzo (1978) discussed the paradoxical phenomenon arising from stochastic traffic assignment problems. They found that, under the assumption of constant link cost, an improvement of a route in a road network sometimes results in increased total system travel cost. Moreover, the total system cost may increase when a new link is added to the road network due to the stochastic nature of the problem.

Such kind of paradoxical phenomenon can also be observed in the proposed stochastic transit assignment model, in which the link cost is not a constant.

Jiang and Szeto (2016) also discussed the capacity paradox arising from with their reliability-based user equilibrium transit assignment problem, in which the network maximum throughput may be reduced after new transit lines are added to a transit network or after the frequency of an existing line increases. A similar paradoxical phenomenon can also be observed using the proposed SUE transit assignment model.

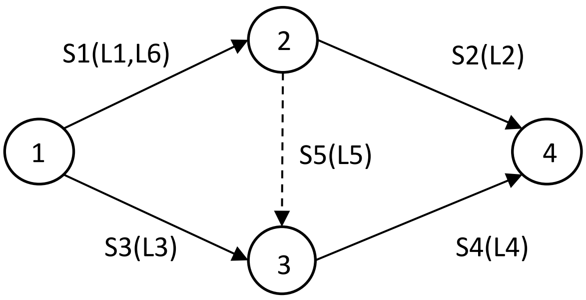


Figure 6. Small example network II

In order to illustrate the aforementioned paradoxical phenomena, a small sample transit network is developed, as presented in Figure 6. The following values of parameters are used in this section: the congestion function parameters $\delta' = \varepsilon = n = 1$, $\varpi_s = 10$; the capacity of a single vehicle is 30 pass/veh; the demand from node 1 to node 4 is 300 pass/hr; $\mu_{\mathrm{T}} = \mu_{\mathrm{W}} = 0.5$ HK$/min.

The following scenarios are discussed to illustrate the paradoxical phenomena:

1) Adding a new line to the network;
2) Improving an existing line in the network.

For the first scenario, two cases are developed: In case 1, L5 in the network (associated with S5) is not provided, while in case 2, L5 is provided with varying its frequency. The link characteristics of the example network II are shown in Table 4. For the second scenario, only case 2 is considered.

Table 4. The link characteristics of the example network II

	In-vehicle time (min)	Frequency (veh/hr)	Capacity (pass/hr)
S1 (L1, L6)	10	2+2	120
S2 (L2)	60	3	90
S3 (L3)	60	3	90
S4 (L4)	10	4	120
S5 (L5)	10	Varying	Varying

4.2.1. The paradox due to the stochastic nature of the problem

The expected total system cost against the frequency of S5 (L5) is shown in Figure 7. In case 1, since the two available paths (i.e., S1-S2 and S3-S4) are symmetric, the expected total system cost remains unchanged (and equals 17500 HK$/hr) irrespective of the value of θ. In case 2, the expected total system cost first increases and then decreases as the frequency of L5 increases. The expected total system cost in case 2 also depends on the value of θ.

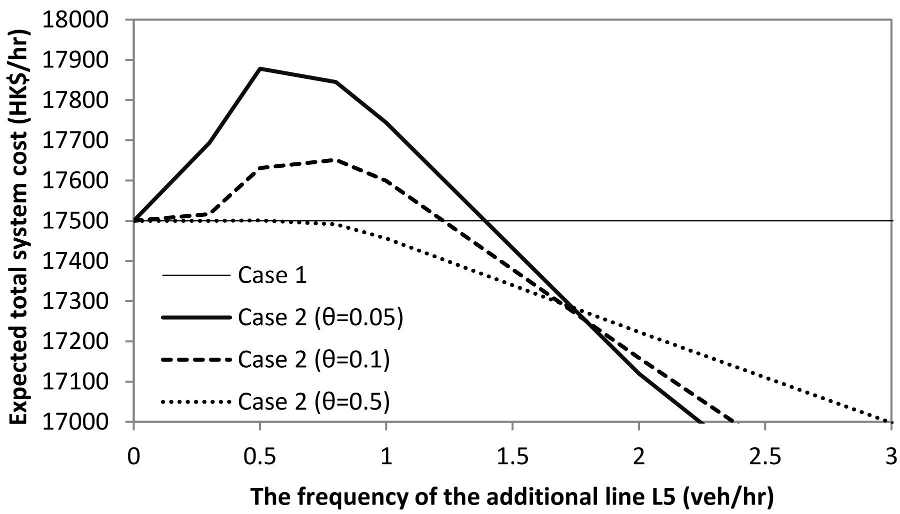


Figure 7. Expected total system cost against the frequency of L5

For the first scenario, comparing the curves of case 1 and case 2 in Figure 8, we can see that providing the additional line can lead to an increase in expected total system cost. A paradox occurs when the expected total system cost obtained in case 2 is higher than the base value in case 1 (i.e., the shaded area in Figure 8). This result shows that when a new line is *added* to the network, the expected total system cost can increase.

The occurrence of the paradox is also affected by the frequency of the additional link S5 or line L5. When the additional link S5 is "good enough" (i.e., the cost of the path using S5 is lower than or close to the costs of competitive paths using S1-S2 and S3-S4), the obtained expected total system cost in case 2 is smaller than the value obtained in case 1 (i.e., 17500 HK$/hr) and the network performance is improved; when the additional link S5 is not "good enough", the paradox occurs.

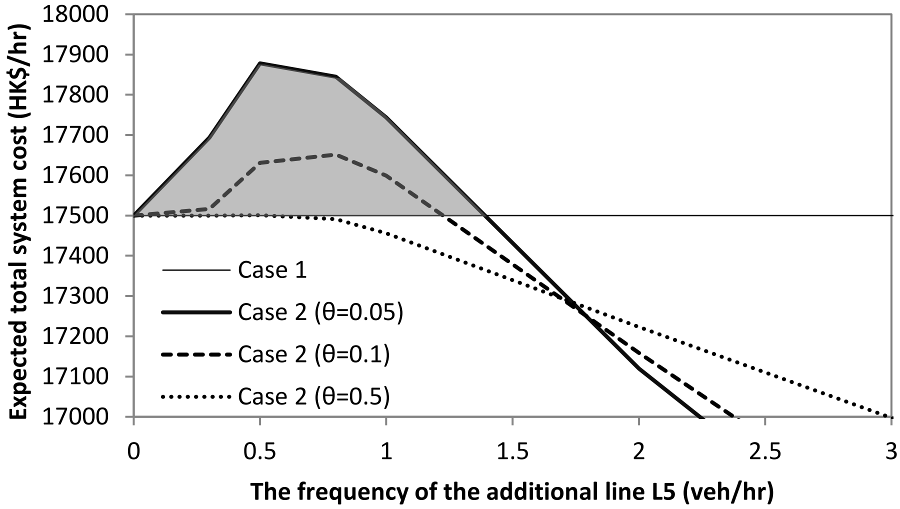


Figure 8. The paradox region of the first scenario

For the second scenario, considering the curves of case 2 only (as shown in Figure 9), the expected total system cost first increases and then decreases as the frequency of S5 increases (i.e., L5 is improved). Within the paradox region shaded in Figure 9, any improvement (i.e., an increase in line frequency) of L5 leads to an increase in expected total system cost. This result shows that an *improvement* of a part of the network can cause an increase in expected total system cost.

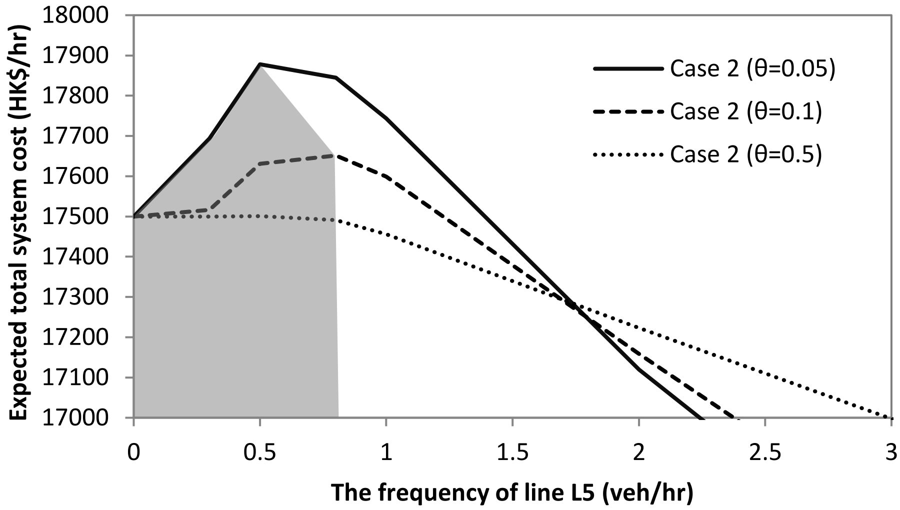


Figure 9. The paradox region of the second scenario

For both scenarios, the paradox is affected by passengers' perception of travel cost (measured by the value of θ). As shown in Figure 8 and Figure 9, the paradoxical phenomenon is obvious when the value of θ is small (i.e., passengers have little information about their travel costs). When the value of θ is sufficiently large (i.e., passengers have perfect information about their travel costs), the SUE condition approaches the UE condition, and such paradoxical phenomenon does not occur.

This situation is analogous to the paradox in traffic assignment presented by Sheffi and Daganzo (1978). This paradoxical phenomenon is mainly caused by the stochastic nature of the problem. When a worse path is added, passengers can make the "wrong" choice. The stochastic nature of the problem ensures that some passengers select the worse path, and thus the expected total system cost increases.

4.2.2. The paradox due to passengers' non-cooperative behavior

Following the network capacity defined by Yang and Bell (1998), the network capacity is the maximum throughput of the network at which all of the bottleneck links just reach their capacities under the equilibrium condition, and a bottleneck link is the link with the lowest capacity on a path. The network throughput against the frequency of S5 (L5) is shown in Figure 10. As mentioned in the previous example, in case 1, since the two available paths (i.e., S1-S2 and S3-S4) are symmetric, the network throughput remains unchanged (and equals 180 pass/hr) irrespective of the value of θ. In case 2, the network throughput first increases and then decreases as the frequency of L5 increases. The network throughput in case 2 depends on the value of θ.

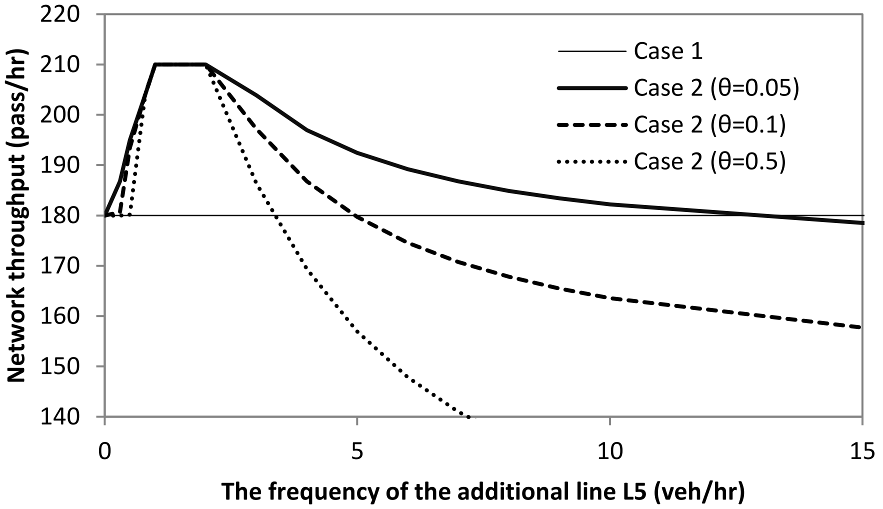


Figure 10. Network throughput against the frequency of L5

For the first scenario, comparing the curves of case 1 and case 2 in Figure 10, when a new line L5 is added to the network, the network throughput can be improved. However, when the additional link S5 is "good enough" (i.e., the cost of the path using S5 is significantly lower than the costs of the competitive paths using S1-S2 and S3-S4), the obtained network throughput in case 2 is smaller than that obtained in case 1 (i.e., 180 pass/hr) and the paradox occurs (i.e., the shaded area in Figure 11). This result shows that when a new line is *added* to the network, the network throughput can decrease.

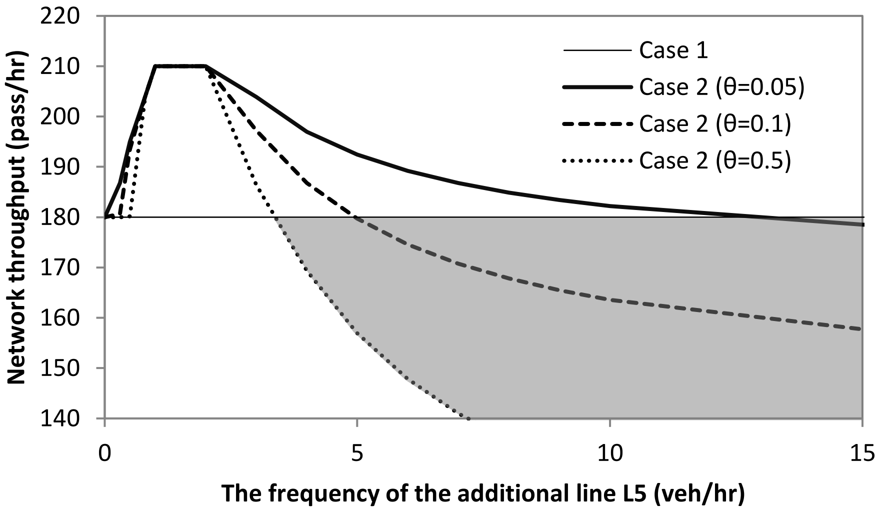


Figure 11. The paradox region of the first scenario

For the second scenario, considering the curves of case 2 only (as shown in Figure 12), the network throughput first increases and then decreases as S5 is improved (i.e., the frequency of L5 increases). Within the paradox region shaded in Figure 12, any improvement (i.e., an increase in line frequency) of L5 leads to a decrease in network throughput. This result shows that an *improvement* of a part of the network can cause a decrease in network throughput.

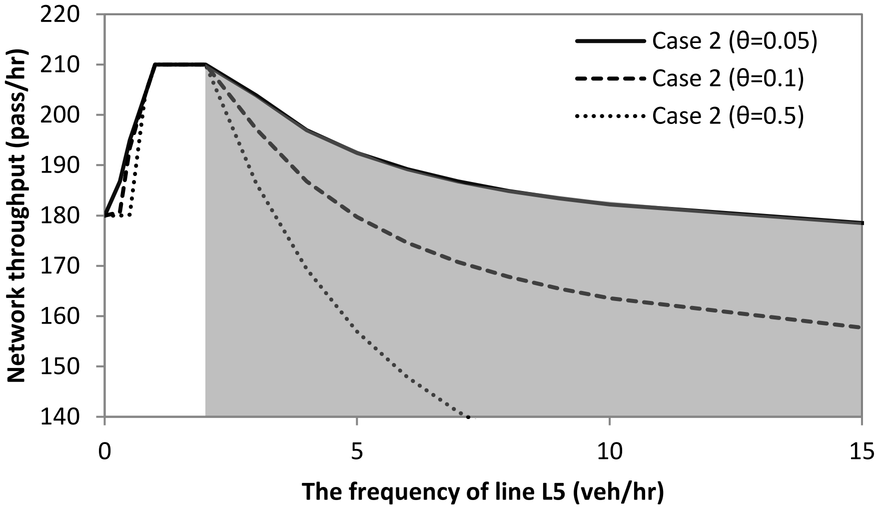


Figure 12. The paradox region of the second scenario

This paradoxical phenomenon is mainly caused by passengers' non-cooperative route choice behavior. The addition of link S5 with a significantly low cost attracts passengers to use the new path S1-S5-S4. The passenger flow on the new path *simultaneously* occupies some capacities of the sharing bottleneck links in the original two paths (i.e., links S1 and S4), implying that the residual capacities of the two bottleneck links are not independent. As a result, the network throughput is not equal to the sum of the bottleneck capacity of each path. When the flow on the new path is sufficiently high, the network throughput is reduced.

For both scenarios, the capacity paradox is affected by passengers' perception of travel cost (measured by the value of θ). As shown in Figure 11 and Figure 12, the network throughput decreases more sharply within the paradox region when the value of θ is larger (i.e., passengers have more information about their travel costs). When passengers have more information about their travel cost, they are more willing to switch to the new better path. It can also be observed from Figure 11 and Figure 12 that, for the first scenario, the range of frequency under which the paradox occurs is affected by the value of θ, while for the second scenario, the range of frequency is independent of the value of θ.

4.2.3. *The occurrence of the two paradoxes*

The aforementioned paradoxes may not occur simultaneously. Figure 13 shows the expected total system cost and network throughput against the frequency of L5 when $\theta = 0.1$. As shown in this figure, the paradox regions of the two paradoxes are not overlapping. Sheffi and Daganzo's paradox occurs when the frequency of the additional line (i.e., L5) is low, while the network throughput paradox occurs when the frequency of the additional line is high. This result implies that, when considering adding new transit lines to the network or improving frequencies of existing lines in the network, the frequencies should be carefully designed to avoid the occurrences of the two paradoxes. Addressing one paradox issue does not imply that the other paradox issue has also been addressed. A bi-objective bilevel transit network design model is needed for such a purpose.

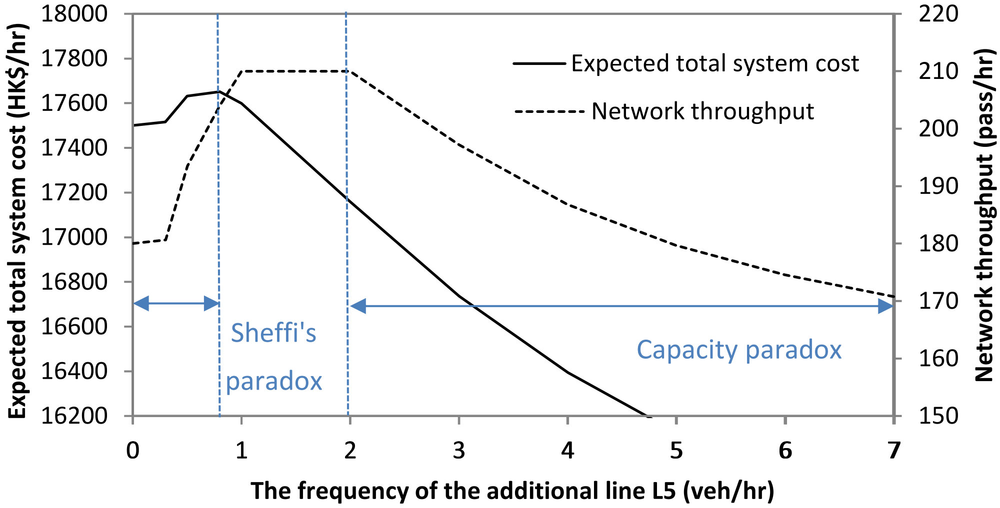


Figure 13. The occurrences of the two paradoxes

4.3. The performance of the cost-averaging SRAM

4.3.1. *The effect of step size parameters in the cost-averaging SRAM*

The effects of different algorithmic parameters are tested using the Sioux-Falls network as shown in Figure 14. Each number in the figure denotes the in-vehicle travel time of the link in seconds. Table 5 shows the frequency and stop sequence settings of the prefixed route structure. It is assumed that transit routes can pass a node without stopping at that node in this example. There are 16 O-D pairs in the network with demand data shown in Table 6. θ is set to be 0.5 and $\mu_{\mathrm{T}} = \mu_{\mathrm{W}} = 1$ HK$/min in Section 4.3 unless otherwise specified. The algorithm was coded and compiled by Bloodshed Dev-C++.

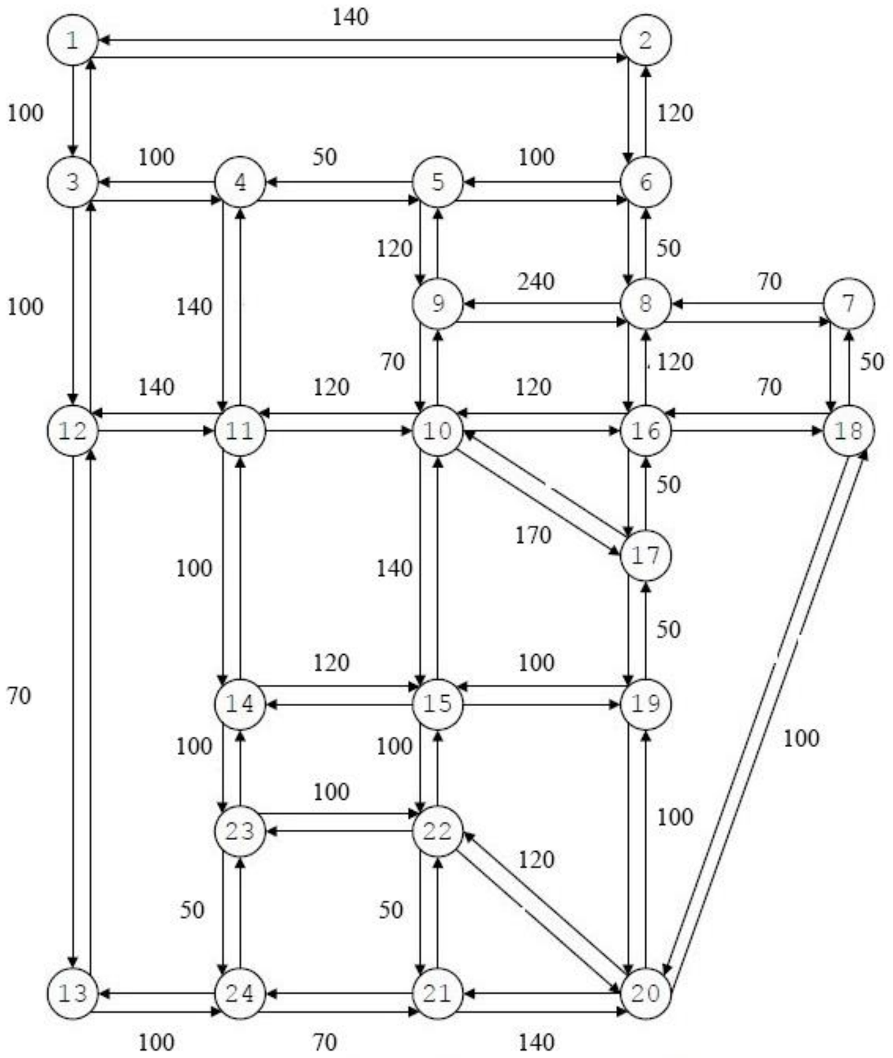


Figure 14. The Sioux-Falls network

Table 5. The transit route setting of the Sioux-Falls network

Route No.	Headways (min)	Stop sequence								
1	6	4	11	23	24					
2	6	1	3	12	13	24				
3	6	11	14	23	24	13				
4	5	8	20	21	22	23				
5	6	7	8	16	18	20				
6	6	14	15	19	20	22	23			
7	3	2	6	8	9	10	11	12		
8	3	4	5	9	10	17	19	20		
9	3	10	16	17	19	20	21	24		
10	3	1	3	4	5	9	10	15	19	20

Table 6. The demand setting (pass/hr) of the Sioux-Falls network example

Origin \ Destination	13	20	21	24
1	500	500	500	500
2	400	400	400	400
3	500	500	500	500
4	400	400	400	400

In order to investigate the effect of the step size parameters in the cost-averaging SRAM on the convergence speed, the Sioux-Falls network scenario was solved under different combinations of η and γ, and the number of iterations is used as the criterion for measuring computation speed. The results are shown in Table 7, with the convergence tolerance ε set to be 0.0001. As shown in the table, the smallest number of iterations evaluated at convergence occurs when $\eta = 3$ and $\gamma = 0.3$. It should be noted that when $\eta = \gamma = 1$, the SRAM is equivalent to the conventional MSA, and the number of iterations evaluated at convergence equals 83 using the conventional MSA. As we can see from the table, most of the combinations of η and γ in the SRAM give a smaller number of iterations than the conventional MSA. Therefore, the SRAM is more efficient than the MSA under the condition that the step size parameters are chosen properly.

Table 7. The number of iterations at the convergence

η / γ	1	1.5	2	2.5	3	3.5	4
1	**83**	75	72	58	64	75	91
0.9	65	56	57	46	50	59	71
0.8	52	45	46	37	40	65	80
0.7	42	38	38	32	32	53	65
0.6	36	33	31	37	26	43	53
0.5	32	29	34	33	21	33	44
0.4	30	27	29	31	20	28	37
0.3	28	25	25	28	**19**	25	29
0.2	32	28	22	26	21	21	25
0.1	38	29	27	35	28	21	24

4.3.2. A comparison of the cost-averaging SRAM, the cost-averaging MSA, the flow-averaging SRAM, and the flow-averaging MSA

To illustrate the efficiency of the cost-averaging SRAM (c-SRAM), the convergence rate of the cost-averaging SRAM (c-SRAM) is compared with the cost-averaging MSA (c-MSA), the flow-averaging SRAM (f-SRAM), and the flow-averaging MSA (f-MSA) based on the Sioux-Falls network (Figure 14). For the c-SRAM, we set η = 3 and γ = 0.3 (the best values obtained in the last example); for the f-SRAM, we set η = 1.8 and γ = 0.05 (the best values obtained from several trials). In this example, the error in expected total system cost is defined to be the difference between expected total system cost obtained at each iteration and that at convergence and is used as the convergence measure.

The convergence curves are presented in Figure 15. As shown in the figure, the proposed c-SRAM, c-MSA, and f-SRAM converge quickly, but the curve of the f-MSA has a very long tail. The f-MSA fails to meet the convergence criterion within 10000 iterations. As shown in the figure, the convergence rate of the c-SRAM is the fastest among the four algorithms examined. Furthermore, the performance of cost-averaging algorithms is better than that of the flow-averaging algorithms. At each iteration, the dimension of decision variables in both the flow-based versions of the SRAM and the MSA is $|S|\times|D|$, but the dimension of decision variables in both the cost-based versions of the SRAM and the MSA is only $|S|$. Therefore, the cost-averaging methods are more efficient than the flow-averaging methods, especially when the network is large. Moreover, the two SRAM algorithms are significantly faster than the two MSA algorithms. Since the f-MSA is a special case of the f-SRAM, we can always find out a step size setting such that the convergence rate of the f-SRAM is not worse than that of the f-MSA. A similar argument holds for the c-SRAM.

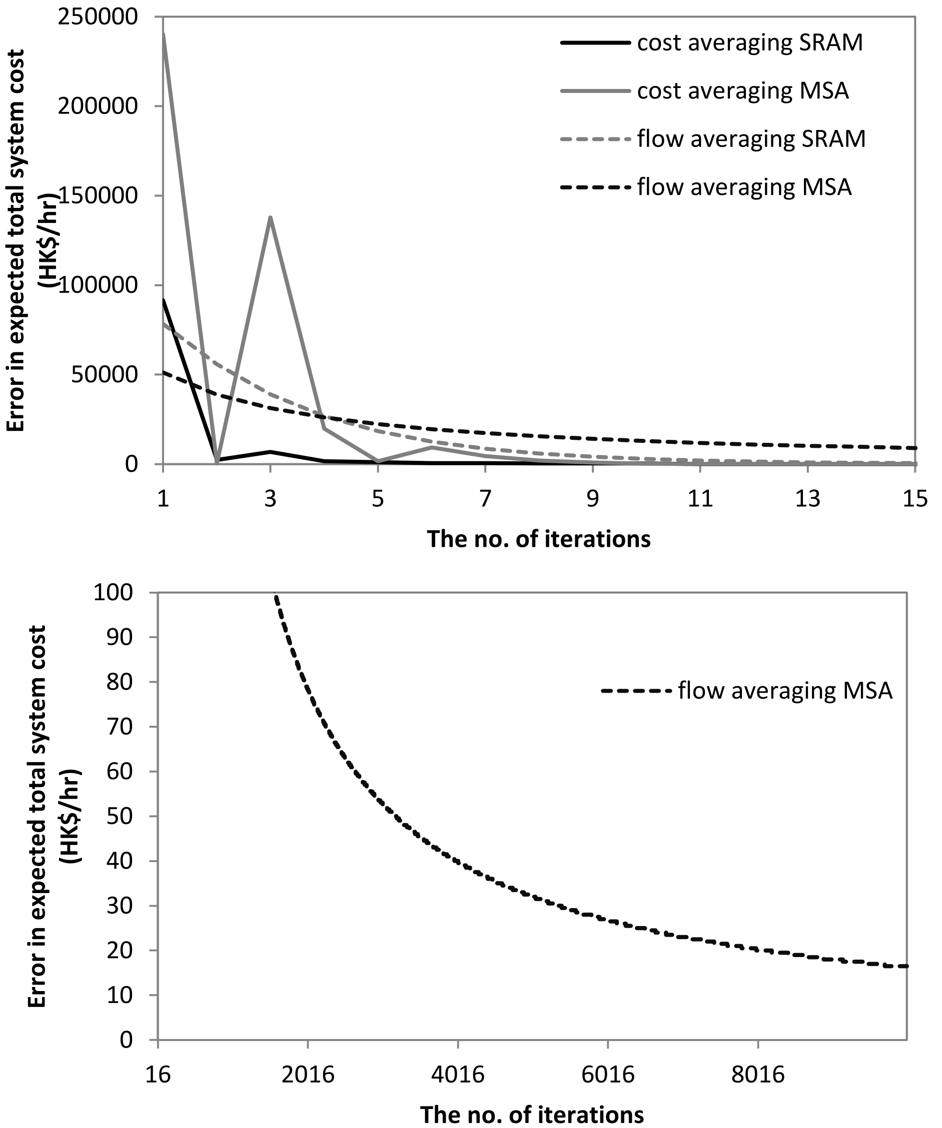


Figure 15. The comparison of the convergence of the 4 algorithms

4.3.3. The effect of demand levels

It is also interesting to test the effect of demand levels on the solution speed of the proposed model. The demand level reflects the degree of in-vehicle congestion. The demand of the Sioux-Falls network is uniformly varied by multiplying the base demand level by a constant demand scaling factor, where the factor represents the demand level relative to the base demand. In this example, the demand factor is set to be 0.2-2.0. The model was solved using the cost-averaging SRAM with the parameter setting as follows: $\eta = 3$ and $\gamma = 0.3$.

The results of the number of iterations evaluated at convergence obtained by varying the demand scaling factor are shown in Figure 16. As shown in the figure, the number of iterations evaluated at convergence increases as the demand scaling factor increases. This result implies that, as the network gets crowded, the computation time required for convergence is long. It is mainly because when the network gets crowded, the changes in congestion cost between successive iterations increase, and therefore the convergence rate decreases. Moreover, as the demand level changes, the most suitable step size parameters also change, and the attractive path set changes as well. Therefore, the curve shown in Figure 16 is not smooth.

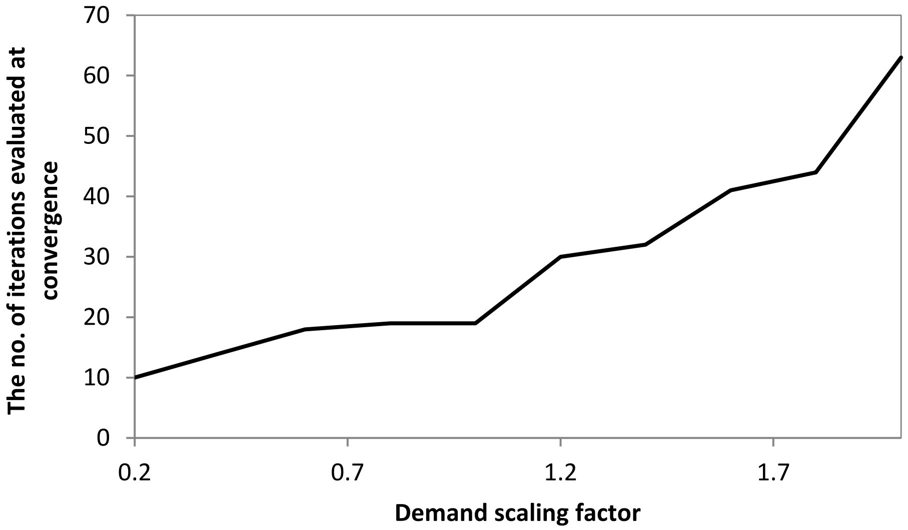


Figure 16. The number of iterations evaluated at convergence against the demand scaling factor

4.3.4. A large network experiment

The performance of the proposed c-SRAM is also tested using the Winnipeg transit network data provided in the EMME/4 software. The Winnipeg transit network as shown in Figure 17 consists of 691 transit nodes and 133 transit lines. The total demand over the network is 77130 pass/hr. The number of links and the number of approaches of the network are both 43283. The algorithm was coded and compiled using Microsoft Visual Studio 2010 and run on a PC with a 3.6-GHz Core processor and 32 GB RAM.

We set $\eta = 2$ and $\gamma = 0.1$ (the best values obtained from several trials) for this example.

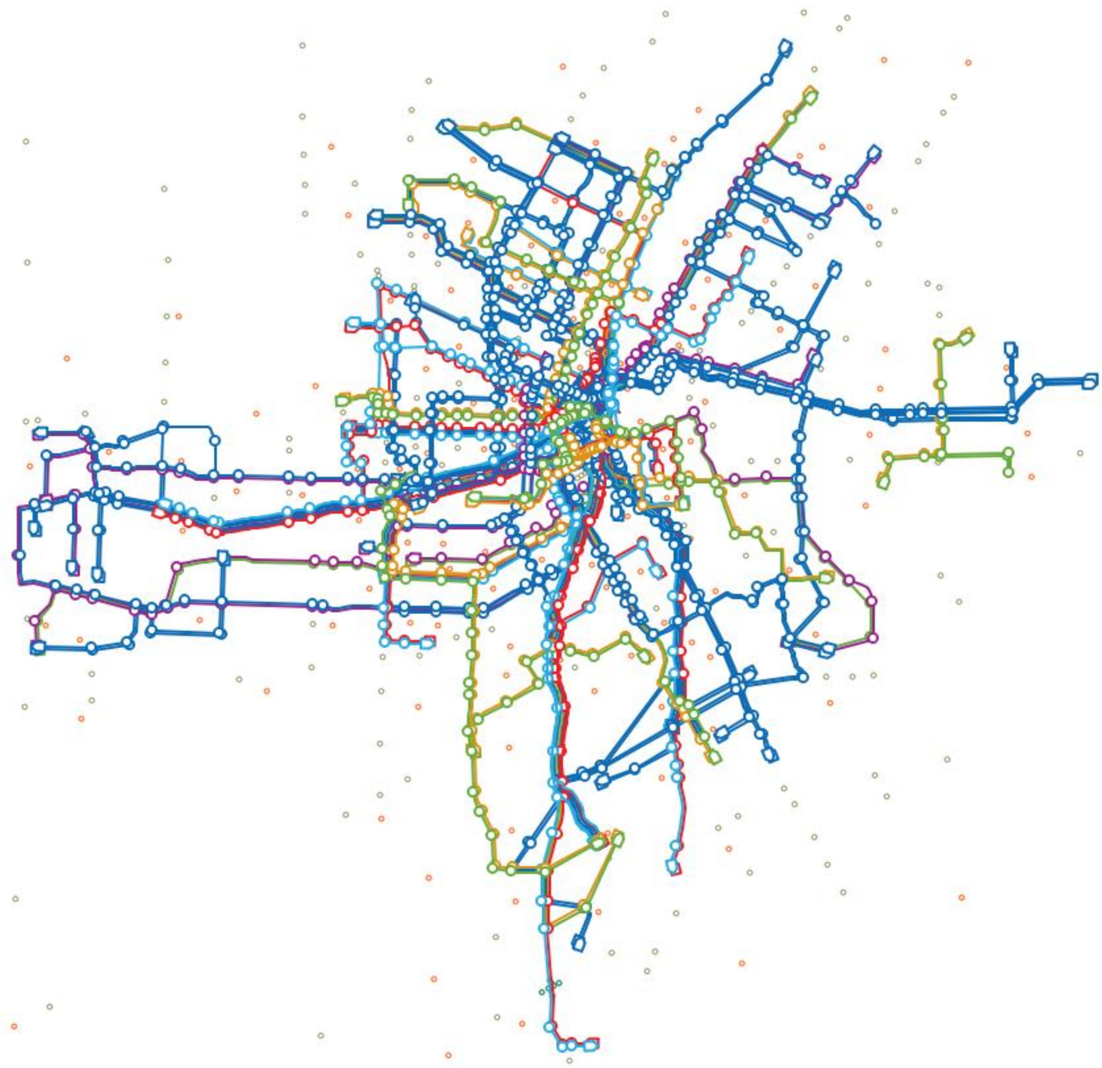


Figure 17. The Winnipeg transit network

The length of the vector of descent directions $\left\|\mathbf{h}^k\right\|$ at each iteration in the c-SRAM given by

$$\left\|\mathbf{h}^k\right\| = \sqrt{\sum_{s \in S}\left(c_s^k - c_s^{k-1}\right)^2}, \quad \forall k > 1, \tag{32}$$

as well as the difference in expected total system cost between two adjacent iterations at each iteration are shown in Figure 18 and Figure 19, respectively. As shown in these figures, the proposed c-SRAM converges efficiently. The difference in expected total system cost between two adjacent iterations is reduced to below 0.5% within the first 20 iterations. The average CPU time of each iteration is 607 seconds. To sum up, the proposed c-SRAM is capable of solving large-scale transit networks.

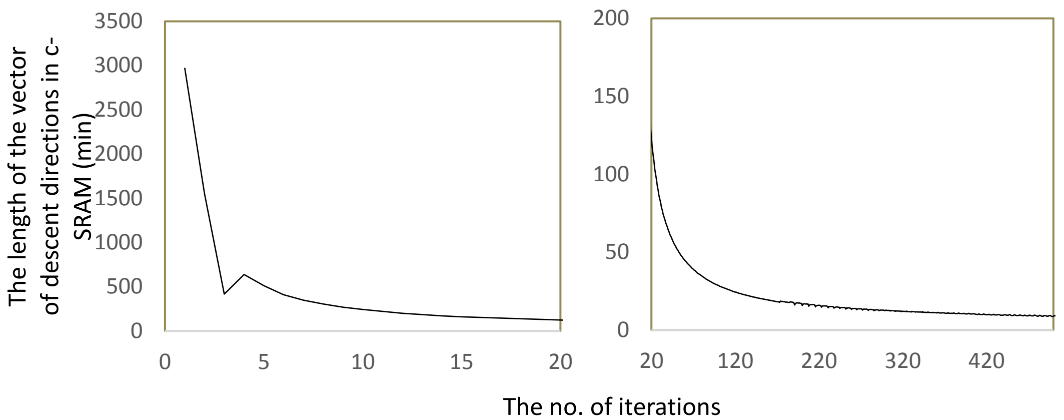


Figure 18. The length of the vector of descent directions in the c-SRAM

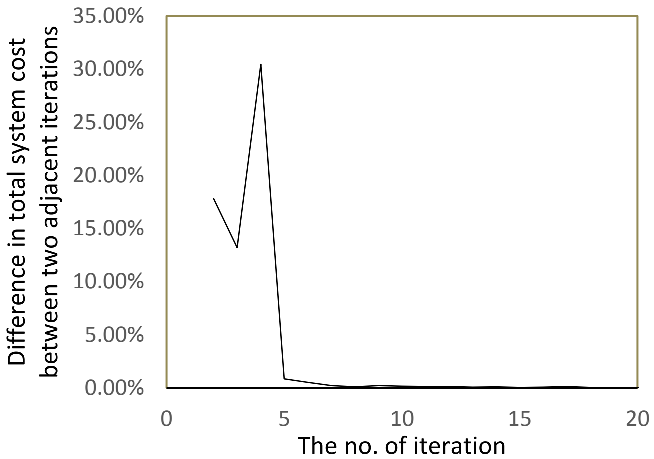


Figure 19. The difference in expected total system cost between two adjacent iterations

5. CONCLUSION

In this paper, an approach-based transit assignment model under the assumption of logit-based SUE with fixed demand is proposed. In-vehicle congestion cost is taken into account in the proposed model. It is proven that the model has exactly one solution. A cost-averaging SRAM is developed to solve the proposed model. The convergence of the algorithm is proven.

Various numerical tests are carried out to examine the model properties and paradoxical phenomena. It is shown that the solution of the approach-based SUE transit assignment model is equivalent to that of the path-based SUE transit assignment model. It is also shown that Sheffi and Daganzo's paradox or capacity paradox may occur when an additional transit line is introduced to the network, or the frequency of an existing transit line is improved. Sheffi and Daganzo's paradox is caused by the stochastic nature of the model and demonstrated to be obvious when the value of θ is small, while the capacity paradox is caused by passengers' non-cooperative route choice behavior and demonstrated to be obvious when the value of θ is large. The two paradoxes may not occur simultaneously. The occurrences of the two paradoxes regarding expected total system cost and network throughput highly depend on the value of θ. It is therefore important to accurately calibrate the θ value for transit network design applications. The results also show that the two paradoxes can be avoided under certain frequency settings. These findings may call for a bi-objective bilevel transit network design formulation to determine the optimal frequency to improve the system performance.

The performance of the cost-averaging SRAM, the c-SRAM, is illustrated and discussed. We find the following: 1) The speed of the c-SRAM can be optimized by choosing suitable values of step size parameters. 2) The c-SRAM is faster than the flow-averaging SRAM because the former has a smaller number of decision variables. 3) The c-SRAM is faster than the traditional MSA when suitable values of step size parameters are chosen. 4) The c-SRAM converges efficiently and can solve large-scale transit networks such as the Winnipeg transit network. 5) The c-SRAM convergence rate decreases as demand or the in-vehicle congestion level increases.

ACKNOWLEDGMENTS

The work described in this paper was partially supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region of China (HKU17218916), and a grant from the University Research Committee of the University of Hong Kong (201611159067). The authors are grateful for the two reviewers for their constructive comments.

REFERENCES

Binder, S., Maknoon, Y., and Bierlaire, M., 2017. Exogenous priority rules for the capacitated passenger assignment problem. *Transportation Research Part B: Methodological, 105*, pp. 19–42.

Cantarella, G. E., 1997. A general fixed-point approach to multimode multi-user equilibrium assignment with elastic demand. *Transportation Science, 31*(2), pp. 107–128.

Cats, O., West, J., and Eliasson, J., 2016. A dynamic stochastic model for evaluating congestion and crowding effects in transit systems. *Transportation Research Part B: Methodological, 89*, pp. 43–57.

Cepeda, M., Cominetti, R., and Florian, M., 2006. A frequency-based assignment model for congested transit networks with strict capacity constraints: characterization and computation of equilibria. *Transportation Research Part B: Methodological, 40*(6), pp. 437–459.

Chriqui, C. and Robillard, P., 1975. Common bus lines. *Transportation Science, 9*(2), pp. 115–121.

Codina, E. and Rosell, F., 2017. A heuristic method for a congested capacitated transit assignment model with strategies. *Transportation Research Part B: Methodological, 106*, pp. 293–320.

Cominetti, R. and Correa, J., 2001. Common-lines and passenger assignment in congested transit networks. *Transportation Science, 35*(3), pp. 250–267.

Daganzo, C. F. and Sheffi, Y., 1977. On stochastic models of traffic assignment. *Transportation Science, 11*(3), pp. 253–274.

De Cea, J. and Fernández, E., 1993. Transit assignment for congested public transport systems: an equilibrium model. *Transportation Science, 27*(2), pp. 133–147.

Dial, R. B., 1971. A probabilistic multipath traffic assignment model which obviates path enumeration. *Transportation Research, 5*(2), pp. 83–111.

Fisk, C., 1980. Some developments in equilibrium traffic assignment. *Transportation Research Part B: Methodological, 14*(3), pp. 243–255.

Florian, M. 1998. Deterministic time table transit assignment. Preprints of PTRC seminar on National models, Stockholm.

Florian, M., 2004. Finding shortest time-dependent paths in Schedule-Based transit networks: a Label Setting algorithm. In: *Schedule-based Dynamic Transit Modeling: Theory and Applications*. Springer, pp. 43–52.

Hamdouch, Y., Ho, H. W., Sumalee, A., and Wang, G., 2011. Schedule-based transit assignment model with vehicle capacity and seat availability. *Transportation Research Part B: Methodological, 45*(10), pp. 1805–1830.

Hamdouch, Y. and Lawphongpanich, S., 2008. Schedule-based transit assignment model with travel strategies and capacity constraints. *Transportation Research Part B: Methodological, 42*(7), pp. 663–684.

Hamdouch, Y., Szeto, W. Y., and Jiang, Y., 2014. A new schedule-based transit assignment model with travel strategies and supply uncertainties. *Transportation Research Part B: Methodological, 67*, pp. 35–67.

Jiang, Y. and Szeto, W. Y., 2016. Reliability-based stochastic transit assignment: Formulations and capacity paradox. *Transportation Research Part B:*

Methodological, 93, pp. 181–206.

Kurauchi, F., Bell, M. G. H., and Schmöcker, J. D., 2003. Capacity constrained transit assignment with common lines. *Journal of Mathematical Modelling and Algorithms, 2*(4), pp. 309–327.

Lam, W. H. K., Gao, Z. Y., Chan, K. S., and Yang, H., 1999. A stochastic user equilibrium assignment model for congested transit networks. *Transportation Research Part B: Methodological, 33*(5), pp. 351–368.

Lam, W. H. K. and Zhou, J., 1999. Stochastic transit assignment with elastic demand. *Journal of Eastern Asia Society for Transportation Studies, 3*(2), pp. 75–87.

Li, Z. C., Lam, W. H. K., Wong, S. C., and Sumalee, A., 2010. An activity-based approach for scheduling multimodal transit services. *Transportation, 37*(5), pp. 751–774.

Liu, H. X., He, X., and He, B., 2009. Method of successive weighted averages (MSWA) and self-regulated averaging schemes for solving stochastic user equilibrium problem. *Networks and Spatial Economics, 9*(4), pp. 485–503.

Liu, Z. and Meng, Q., 2014. Bus-based park-and-ride system: a stochastic model on multimodal network with congestion pricing schemes. *International Journal of Systems Science, 45*(5), pp. 994–1006.

Long, J. C., Huang, H. J., Gao, Z. Y., and Szeto, W. Y., 2013. An intersection-movement-based dynamic user optimal route choice problem. *Operations Research, 61*(5), pp. 1134–1147.

Long, J. C., Szeto, W. Y., and Huang, H. J., 2014. A bi-objective turning restriction design problem in urban road networks. *European Journal of Operational Research, 237*(2), pp. 426–439.

Long, J. C., Szeto, W. Y., Huang, H. J., and Gao, Z. Y., 2015. An intersection-movement-based stochastic dynamic user optimal route choice model for assessing network performance. *Transportation Research Part B: Methodological, 74*, pp. 182–217.

Nielsen, O. A., 2000. A stochastic transit assignment model considering differences in passengers utility functions. *Transportation Research Part B: Methodological, 34*(5), pp. 377–402.

Nielsen, O. A. and Frederiksen, R. D., 2006. Optimisation of timetable-based, stochastic transit assignment models based on MSA. *Annals of Operations Research, 144*(1), pp. 263–285.

Nuzzolo, A., Crisalli, U., Comi, A., and Rosati, L., 2016. A mesoscopic transit assignment model including real-time predictive information on crowding. *Journal of Intelligent Transportation Systems: Technology, Planning, and Operations, 20*(4), pp. 316–333.

Poon, M. H., Wong, S. C., and Tong, C. O., 2004. A dynamic schedule-based model for congested transit networks. *Transportation Research Part B: Methodological, 38*(4), pp. 343–368.

Ran, B. and Boyce, D., 1996. Network Flow Constraints and Definitions of Travel Times. In: *Modeling Dynamic Transportation Networks*. Springer, pp. 69–85.

Rasmussen, T. K., Watling, D. P., Prato, C. G., and Nielsen, O. A., 2015. Stochastic user equilibrium with equilibrated choice sets: Part II–Solving the restricted SUE for the logit family. *Transportation Research Part B: Methodological, 77*, pp. 146–165.

Schmöcker, J. D., Fonzone, A., Shimamoto, H., Kurauchi, F., and Bell, M. G., 2011. Frequency-based transit assignment considering seat capacities. *Transportation Research Part B: Methodological, 45*(2), pp. 392–408.

Sheffi, Y. and Daganzo, C. F., 1978. Another "paradox" of traffic flow. *Transportation Research, 12*(1), pp. 43–46.

Spiess, H. and Florian, M., 1989. Optimal strategies: a new assignment model for transit networks. *Transportation Research Part B: Methodological, 23*(2), pp. 83–102.

Sumalee, A., Tan, Z., and Lam, W. H. K., 2009. Dynamic stochastic transit assignment with explicit seat allocation model. *Transportation Research Part B: Methodological, 43*(8), pp. 895–912.

Sun, L., Meng, Q., and Liu, Z., 2013. Transit assignment model incorporating bus dwell time. *Transportation Research Record, 2352*, pp. 76–83.

Szeto, W. Y. and Jiang, Y., 2014. Transit assignment: Approach-based formulation, extragradient method, and paradox. *Transportation Research Part B: Methodological, 62*, pp. 51–76.

Szeto, W. Y., Jiang, Y., Wong, K. I., and Solayappan, M., 2013. Reliability-based stochastic transit assignment with capacity constraints: Formulation and solution method. *Transportation Research Part C: Emerging Technologies, 35*, pp. 286–304.

Szeto, W. Y., Solayappan, M., and Jiang, Y., 2011. Reliability□Based Transit Assignment for Congested Stochastic Transit Networks. *Computer□Aided Civil and Infrastructure Engineering, 26*(4), pp. 311–326.

Teklu, F., 2008. A stochastic process approach for frequency-based transit assignment with strict capacity constraints. *Networks and Spatial Economics, 8*(2–3), pp. 225–240.

Trozzi, V., Gentile, G., Bell, M. G. H., and Kaparias, I., 2013. Dynamic user equilibrium in public transport networks with passenger congestion and hyperpaths. *Transportation Research Part B: Methodological, 57*, pp. 266–285.

Verbas, Ö., Mahmassani, H. S., and Hyland, M. F., 2016. Gap-based transit assignment algorithm with vehicle capacity constraints: Simulation-based implementation and large-scale application. *Transportation Research Part B: Methodological, 93*, pp. 1–16.

Watling, D. P., Rasmussen, T. K., Prato, C. G., and Nielsen, O. A., 2015. Stochastic user equilibrium with equilibrated choice sets: Part I–Model formulations under alternative distributions and restrictions. *Transportation Research Part B: Methodological, 77*, pp. 166–181.

Wu, J. H., Florian, M., and Marcotte, P., 1994. Transit equilibrium assignment: a model and solution algorithms. *Transportation Science, 28*(3), pp. 193–203.

Wu, Z. X. and Lam, W. H. K., 2003. Network equilibrium for congested multi□mode networks with elastic demand. *Journal of Advanced Transportation, 37*(3), pp. 295–318.

Yang, H. and Bell, M. G., 1998. A capacity paradox in network design and how to avoid it. *Transportation Research Part A: Policy and Practice, 32*(7), pp. 539–545.

APPENDIX

This appendix gives the proofs of Propositions 1-4.

Proposition 1. Under assumption A8, $a_s^{id} = a_s^{jd}$ for $\forall s \in S^{id} \cap S^{jd}, i, j \in N, d \in D$.

Proof: Following Long et al. (2015), it can be proved that under assumption A8, the topological distances of a common node in two sub-networks are the same. Therefore, under definition D1, for the common node of the two sub-networks formed by S^{id} and S^{jd}, their topological orders are the same.

According to Eq. (9), when $s \in S^{id} \cap S^{jd}$, we have

$$L_s^{id} = L_s^{jd} = \exp\left(\theta(\pi^{t(s)d} - \pi_s^{t(s)d})\right), \quad \forall i, j \in N, d \in D. \tag{33}$$

The following equation can be proved using mathematical induction.

$$W_s^{id} = W_s^{jd}. \tag{34}$$

Let the sequence of the common nodes of sub-networks formed by S^{id} and S^{jd} be $\{i_n, i_{n-1}, \ldots, i_k, \ldots, i_2, i_1\}$, and $i_1 = d$, where the subscript of node i is the topological order of that node. We assume that $W_s^{id} = W_s^{jd}, \forall s \in A_{i_k}^{+}$ is true for the kth node and its succeeding nodes (i.e., $i_{k'}, k' \le k$) in the sequence. Then for the (k+1)th node in the sequence, according to Eq. (33), we have

$$W_{s'}^{id} = L_{s'}^{id}\left(\delta_{h(s')}^{d} + \sum_{m \in A_{h(s')}^{+}} W_{m}^{id}\right) = L_{s'}^{jd}\left(\delta_{h(s')}^{d} + \sum_{m \in A_{h(s')}^{+}} W_{m}^{jd}\right) = W_{s'}^{jd}, \quad \forall i, j \in N, d \in D,$$

where $h(s') = i_{k'}, k' \le k$ and $t(s') = i_{k+1}$. Therefore, it is also true for the (k+1)th node and its succeeding nodes in the sequence.

When $k = 1$, we have $W_{s^*}^{id} = L_{s^*}^{id} = L_{s^*}^{jd} = W_{s^*}^{jd}, \ \forall i, j \in N, d \in D$, where $h(s^*) = i_1 = d$.

Therefore, $W_s^{id} = W_s^{jd}$ is true for $\forall s \in S^{id} \cap S^{jd}$.

By Eqs. (12) and (34), we have

$$a_s^{id} = \frac{W_s^{id}}{\sum_{m \in A_{t(s)}^{+}} W_m^{id}} = \frac{W_s^{jd}}{\sum_{m \in A_{t(s)}^{+}} W_m^{jd}} = a_s^{jd}, \quad \forall i, j \in N, d \in D, s \in S^{id} \cap S^{jd}. \tag{35}$$

This completes the proof. □

<u>Proposition 2.</u> The FP problem (23) has exactly one solution.

<u>Proof</u>: As discussed in Section 2.5.1, a solution to the FP problem (18) exists. Moreover, according to (17), v_s^d is a function of a_s^d. Thus, we can let a solution to the FP problem (18) be $\mathbf{v}^*(\mathbf{a}^*)$ with the elements v_s^{d*}, where $\mathbf{a}^* = \left(a_s^{d*}\right)$; let $\boldsymbol{\alpha}^*$ be the corresponding approach probability vector with the elements α_b^{d*} calculated by $\alpha_b^{d*} = a_{u(b)}^{d}{}^{*}$. Then, $\boldsymbol{\alpha}^*$ satisfies Eqs. (19) to (21) and $\mathbf{v}^*\left(\mathbf{a}^*(\boldsymbol{\alpha}^*)\right)$ satisfies Eq. (22). Thus, $\boldsymbol{\alpha}^*$ is a solution to the FP problem (23); i.e., a solution to the FP problem (23) also exists.

We assume that there is more than one solution to the FP problem (23). Let $\boldsymbol{\alpha}^\# = \left(\alpha_b^{d\#}\right)$ be another solution to the FP problem (23) ($\boldsymbol{\alpha}^* \ne \boldsymbol{\alpha}^\#$) and $\mathbf{v}^\#(\boldsymbol{\alpha}^\#)$ with the elements v_s^{d*} be the *corresponding* link flow vector calculated by Eq. (22); let $\mathbf{a}^{\#}$ be the corresponding link choice probability vector with the elements $a_s^{d\#}$ calculated by $a_s^{d\#} = \sum_{b \in A_{t(s)}^{+}} \delta_s^b \cdot \alpha_b^{d\#}$, where $\delta_s^b = 1$ if $u(b) = s$; $\delta_s^b = 0$ otherwise. Then, $\mathbf{a}^{\#}$ satisfies Eqs. (14) to (16) and $\mathbf{v}^{\#}\left(\boldsymbol{\alpha}^{\#}(\mathbf{a}^{\#})\right)$ satisfies Eq. (17). Thus, $\mathbf{a}^{\#}$ is a solution to the FP problem (18). This solution is different from $\mathbf{v}^{*}(\mathbf{a}^{*})$ because we assume ($\boldsymbol{\alpha}^{*} \neq \boldsymbol{\alpha}^{\#}$) and the relationship between link flow and approach probability is bijective. However, $\mathbf{v}^{*}(\mathbf{a}^{*}) \neq \mathbf{v}^{\#}\left(\boldsymbol{\alpha}^{\#}(\mathbf{a}^{\#})\right)$ contradicts the solution uniqueness of the FP problem (18) as discussed in Section 2.5.1. As a result, the FP problem (23) has only one solution. Therefore, the FP problem (23) has exactly one solution. This completes the proof. □

<u>Proposition 3.</u> The solution to the approach-based SUE problem (23) satisfies the logit-based SUE condition: $\alpha_p^{rd} = \dfrac{\exp\left(-\theta \cdot c_p\right)}{\sum_{p' \in P^{rd}} \left(\exp\left(-\theta \cdot c_{p'}\right)\right)}$, $\forall p \in P^{rd}, r \in R, d \in D$.

<u>Proof</u>: Combining Eqs. (19) to (21), when $u(b^*) = s^* \in S^p$, $p \in P^{rd}, r \in R, d \in D$, and $t(s^*) = r$, we have

$$\alpha_{b*}^{d} = \frac{\exp\left(-\theta(c_{s*} + \pi^{h(s^*)d} - \pi^{t(s^*)d})\right)\left(\delta_{h(s^*)}^{d} + \sum_{n \in A_{h(s^*)}^{+}} W_n^d\right)}{\sum_{m \in A_{t(s^*)}^{+}} W_m^d}$$

$$= \frac{\exp\left(-\theta(c_{s*} + \pi^{h(s^*)d} - \pi^{rd})\right)\left(\sum_{n \in A_{h(s^*)}^{+}} W_n^d\right)}{\sum_{m \in A_r^{+}} W_m^d}, \tag{36}$$

where $\delta_{h(s^*)}^{d} = 0$.

When $u(b^{\#}) = s^{\#} \in S^p$, $p \in P^{rd}$, $r \in R$, $d \in D$, and $h(s^{\#}) = d$, we have

$$\alpha_{b^\#}^d = \frac{\exp\left(-\theta(c_{s^\#} - \pi^{t(s^\#)d})\right) \cdot \delta_{h(s^\#)}^d}{\sum_{m \in A_{t(s^\#)}^+} W_m^d}$$

$$= \frac{\exp\left(-\theta(c_{s^\#} - \pi^{h(s^\#')d})\right)}{\sum_{m \in A_{h(s^\#')}^+} W_m^d}, \tag{37}$$

where $u(b^\#') = s^\#' \in S^p$ is the preceding link of link $s^\#$ and $t(s^\#) = h(s^\#')$; $\delta_{h(s^\#)}^d = 1$.

For all the other approaches b using intermediate links $u(b) = s \in S^p$, $p \in P^{rd}$, $r \in R$, and $d \in D$, we have

$$\alpha_b^d = \frac{\exp\left(-\theta(c_s + \pi^{h(s)d} - \pi^{t(s)d})\right)\left(\delta_{h(s)}^d + \sum_{n \in A_{h(s)}^+} W_n^d\right)}{\sum_{m \in A_{t(s)}^+} W_m^d}$$

$$= \frac{\exp\left(-\theta(c_s + \pi^{h(s)d} - \pi^{h(s')d})\right)\left(\sum_{n \in A_{h(s)}^+} W_n^d\right)}{\sum_{m \in A_{h(s')}^+} W_m^d}, \tag{38}$$

where $u(b') = s' \in S^p$ is the preceding link of link s and $t(s) = h(s')$; $\delta_{h(s)}^d = 0$.

Combining Eqs. (24), (25), and (36) to (38) and canceling out the repeated terms, we have

$$
\begin{aligned}
\alpha_p^{rd} &= \alpha_{b*}^{d} \cdot \ldots \cdot \alpha_{b'}^{d} \cdot \alpha_{b}^{d} \cdot \ldots \cdot \alpha_{b^{\#'}}^{d} \cdot \alpha_{b^{\#}}^{d} \\
&= \frac{\exp\left(-\theta(c_{s*} \cancel{+\pi^{h(s*)d}} - \pi^{rd})\right)\left(\cancel{\sum_{n\in A^{+}_{h(s*)}} W_n^d}\right)}{\sum_{m\in A_r^{+}} W_m^d} \cdot \ldots \\
&\cdot \frac{\exp\left(-\theta(c_{s'} \cancel{+\pi^{h(s')d}} \cancel{-\pi^{h(s'')d}})\right)\left(\cancel{\sum_{n\in A^{+}_{h(s')}} W_n^d}\right)}{\cancel{\sum_{m\in A^{+}_{h(s'')}} W_m^d}} \cdot \frac{\exp\left(-\theta(c_{s} \cancel{+\pi^{h(s)d}} \cancel{-\pi^{h(s')d}})\right)\left(\cancel{\sum_{n\in A^{+}_{h(s)}} W_n^d}\right)}{\cancel{\sum_{m\in A^{+}_{h(s')}} W_m^d}} \cdot \ldots \\
&\cdot \frac{\exp\left(-\theta(c_{s} \cancel{+\pi^{h(s^{\#'})d}} \cancel{-\pi^{h(s^{\#''})d}})\right)\left(\cancel{\sum_{n\in A^{+}_{h(s^{\#'})}} W_n^d}\right)}{\cancel{\sum_{m\in A^{+}_{h(s^{\#''})}} W_m^d}} \cdot \frac{\exp\left(-\theta(c_{s^{\#}} \cancel{-\pi^{h(s^{\#'})d}})\right)}{\cancel{\sum_{m\in A^{+}_{h(s^{\#'})}} W_m^d}} \\
&= \frac{\exp\left(-\theta(c_{s*} - \pi^{rd})\right)}{\sum_{m\in A_r^{+}} W_m^d} \cdot \ldots \cdot \exp\left(-\theta(c_{s'})\right) \cdot \exp\left(-\theta(c_{s})\right) \cdot \ldots \cdot \exp\left(-\theta(c_{s^{\#'}})\right) \cdot \exp\left(-\theta(c_{s^{\#}})\right) \\
&= \frac{\exp\left(-\theta(c_{s*} + \ldots + c_s + \ldots + c_{s^{\#}} - \pi^{rd})\right)}{\sum_{m\in A_r^{+}} W_m^d} \\
&= \frac{\exp\left(-\theta(c_p - \pi^{rd})\right)}{\sum_{m\in A_r^{+}} W_m^d}, \quad \forall p \in P^{rd}, r \in R, d \in D.
\end{aligned}
\tag{39}
$$

Moreover, by definition and Eq. (39), we have

$$
\sum_{p\in P^{rd}} \alpha_p^{rd} = \frac{\sum_{p\in P^{rd}} \left(\exp\left(-\theta(c_p - \pi^{rd})\right)\right)}{\sum_{m\in A_r^{+}} W_m^d} = 1, \quad \forall r \in R, d \in D. \tag{40}
$$

Therefore,

$$
\sum_{m\in A_r^{+}} W_m^d = \sum_{p\in P^{rd}} \left(\exp\left(-\theta(c_p - \pi^{rd})\right)\right), \quad \forall r \in R, d \in D. \tag{41}
$$

Combining Eqs. (39) and (41), we have

$$
\alpha_p^{rd} = \frac{\exp\left(-\theta(c_p - \pi^{rd})\right)}{\sum_{p'\in P^{rd}} \left(\exp\left(-\theta(c_{p'} - \pi^{rd})\right)\right)} = \frac{\exp\left(-\theta \cdot c_p\right)}{\sum_{p'\in P^{rd}} \left(\exp\left(-\theta \cdot c_{p'}\right)\right)}, \quad \forall p \in P^{rd}, r \in R, d \in D. \tag{42}
$$

This completes the proof. □

Proposition 4. The approach-based SUE problem (23) is equivalent to the link-based SUE problem (18).

Proof: First, we can let the solution to the FP problem (18) be $\mathbf{v}^*(\mathbf{a}^*)$ with the elements v_s^{d*}, where $\mathbf{a}^* = \left(a_s^{d*}\right)$; let $\boldsymbol{\alpha}^*$ be the corresponding approach probability vector with the elements α_b^{d*} calculated by $\alpha_b^{d*} = a_{u(b)}^{d\;*}$. Then, $\boldsymbol{\alpha}^*$ satisfies Eqs. (19) to (21) and $\mathbf{v}^*\left(\mathbf{a}^*(\boldsymbol{\alpha}^*)\right)$ calculated using Eq. (13) satisfies Eq. (22). Thus, $\boldsymbol{\alpha}^*$ is a solution to the FP problem (23).

Secondly, let the solution to the FP problem (23) be $\mathbf{v}^\#(\boldsymbol{\alpha}^\#)$ with the elements $v_s^{d\#}$, where $\boldsymbol{\alpha}^\# = \left(\alpha_b^{d\#}\right)$; let $\boldsymbol{\alpha}^\#$ be the corresponding link choice probability vector with the elements $a_s^{d\#}$ calculated by $a_s^{d\#} = \sum_{b \in A_{t(s)}^+} \delta_s^b \cdot \alpha_b^{d\#}$, where $\delta_s^b = 1$ if $u(b) = s$; $\delta_s^b =$ 0 otherwise. Then, $\mathbf{a}^\#$ satisfies Eqs. (9) to (12), and $\mathbf{v}^\#\left(\boldsymbol{\alpha}^\#(\mathbf{a}^\#)\right)$ calculated using Eq. (22) satisfies Eq. (13). Thus, $\mathbf{a}^\#$ is a solution to the FP problem (18).

Based on the above and the solution existence and uniqueness of the FP problems (18) and (23), we can conclude that the approach-based SUE problem (23) is equivalent to the link-based SUE problem (18).