MAXIMAL RESERVE TRIP GENERATION FOR TRANSIT NETWORKS

Qiang MENG
Assistant Professor
Department of Civil Engineering
National University of Singapore
BLK E1, #07-10,1 Engineering Drive 2
Singapore 117576
Fax: +65 6779 1635
Email: cvemq@nus.edu.sg

Der-Horng LEE
Associate Professor
Department of Civil Engineering
National University of Singapore
BLK E1A, #07-16,1 Engineering Drive 2
Singapore 117576
Fax: +65 6779 1635
Email: dhl@nus.edu.sg

Lan WU
Lecturer
Faculty of Civil Engineering
Nanjing Forestry University,
Jiangsu, P. R. China
Email: Lannywu_sg@yahoo.com

Abstract: This paper addresses a maximal transit reserve trip generation which quantitatively reflects interaction between land-use and public transit system. It firstly proposes a variational inequality and a diagonalization method for the combined transit trip distribution/logit-based stochastic user equilibrium transit assignment with elastic frequencies of transit lines. Secondly, this paper develops a generalized bilevel programming model for the maximal transit trip generation, which consists of two level problems – upper and lower level problems. The upper level problem aims to maximize the reserve transit trip for each origin subject to the transit equity constraints, which is formulated as a multi-objective maximization problem. The lower level problem is the proposed variational inequality for the combined trip distribution/logit-based stochastic user equilibrium transit assignment. A genetic algorithm embedded with the diagonalization method is designed for solving the generalized bilevel programming model. Finally, a numerical example is carried out.

Key Words: Public transit network, Stochastic user equilibrium, Reserve trip generation

1. INTRODUCTION

Public transit is an important transport mode in cities with high population densities. There have been a number of transit related researches conducted in the past decades. For example, estimating passengers’ waiting, boarding and alighting times at a transit stop (Lam et al., 1998); transit scheduling and routing (e.g., Turnquist, 1981; Ceder, 1984); transit network design (e.g., Ceder and Wilson, 1986; Wan and Lo, 2003; Chakroborty, 2003); estimation of transit origin-destination matrices (Wong and Tong, 1998) transit assignment for uncongested transit networks (Spiess and Florian, 1989). Since early 1980s, transit assignment problems for congested transit networks have been received more attentions. The deterministic user equilibrium (DUE) transit assignment problems have been investigated (Nguyen and Pallottino, 1988; De Cea and Fernández, 1993; Wu et al., 1994; Cominetti and Correa, 2004; Babazadeh and Aashtiani, 2005; Cepeda et al., 2006). Lam et al. (1999) examined the logit-based stochastic user equilibrium (SUE) transit assignment problem for congested transit networks. Gao et al. (2004) proposed a transit network design problem with logit-based SUE transit constraints. The preceding studies demonstrate that DUE and SUE
principles arising from road transportation networks (Shiffi, 1985) have been successfully generalized to characterize the passengers’ route choices in transit networks. Although the combined trip distribution/traffic assignment problem has been well explored for road transportation networks (Oppenheim, 1994), it has not yet been extended for congested transit networks. Such an extension is needed to examine transit trip generation. Hence, the first goal of this paper is to study a combined transit trip distribution/logit-based SUE transit assignment problem.

Having had a combined transit trip distribution/logit-based SUE transit assignment model, we can seek the maximal reserve transit trip generation for a transit network that is vital for evaluating a land-used and development project. It is well-known that any decision of land-use and development, such as residential and/or commercial developments, inevitably affects traffic generation and mode choice, and therefore the degree of congestion of the transportation network. There have been notable efforts in modeling interactions between land-use and road transportation (Oppenheim, 1994). However, little attention has been placed to address interaction between land-use and public transit. It is no doubt that a land-use development project will induce added transit trips. These induced transit trips may cause a redistribution of transit network flow pattern due to passengers’ behavior in transit route choice. Consequently, the average origin-destination (O-D) travel times are changed and most likely increased. Note the degree of the increase in O-D travel times may vary differently across O-D pairs. Thus, the transit equity represented by the changes of O-D travel times becomes a must addressed issue. It should be pointed out that the equity issue for road network design problems has been studied by Meng et al. (2002). The second goal of this paper is to formulate the maximal reserve trip generation problem for transit networks subject to the transit equity constraints, while taking into account passengers’ stochastic user-equilibrium (SUE) behavior in their route choices. Note that the maximal reserve trip generation problem for road transportation networks has been addressed by Yang et al. (2000) and Lee et al. (2006), which are different from the maximal reserve capacity proposed by Wong and Yang (1997).

This paper first proposes a variational inequality (VI) model for the combined transit trip distribution/logit-based SUE transit assignment problem with elastic frequencies of transit lines. To certain extent, it is an extension of the conventional combined travel forecasting model for road transportation networks (see, Oppenheim, 1994). A diagonalization method for solving the proposed VI model is then developed. In view of the equity issue for road transportation network by Meng et al. (2002), the concerned transit equity issue is quantified by the ratio of changes in average O-D travel times before and after the implementation of a land-use development project. By incorporating the transit equity constraints and physical limitations in land-use and zonal developments, a generalized bilevel programming model is formulated to find the maximal reserve transit trip generation. The generalized bilevel programming model consists of two problems – upper and lower level problems. The upper level problem aims to maximize the reserve transit trip for each origin subject to the transit equity constraints, which can be formulated as a multi-objective maximization problem. The lower level problem is the VI model for the combined transit trip distribution/logit-based SUE transit assignment. Since the generalized bilevel programming model is a non-differentiable optimization problem, a genetic algorithm (GA) embedded with the proposed diagonalization method is designed for solving the generalized bilevel programming model.

This remainder of this paper is organized as follows. Section 2 introduces concepts, notations and assumptions in modeling transit networks. Subsequently, a VI model and its solution
method for the combined transit trip distribution/logit-based SUE transit assignment problem with elastic frequency of transit lines are presented. Based on the VI model, a generalized bilevel programming model for the maximal reserve transit trip generation will be developed. A GA embedded with the diagonalization method is presented for solving the generalized bilevel programming model. Finally, a numerical example is conducted to demonstrate the proposed models and algorithms followed by the conclusions.

2. DEFINITIONS, NOTATIONS AND ASSUMPTIONS

2.1 Transit Network Defined by Route Sections

A transit network consists of a set of transit lines denoted by $L$, and a set of stations (nodes) denoted by $N$ where passengers can board, alight or transfer. A transit line can be described by its frequency, route, and vehicle types. A transit line segment, denoted by $e$, is a portion of a transit line between any two consecutive stations (nodes). Different transit lines may run parallel for part of their itineraries with some stations in common. The fact of common lines between a pair of transfer stations makes the transit assignment difficult (Lam et al., 1999) because passengers may transfer one or more than one time toward their destinations. It is a unique feature in transit networks that the passengers would usually prefer choosing the transfer nodes then determining the detailed sequence of transit lines to be traveled. This is why the concept of a path broadly used in road transportation networks cannot be directly applied to transit networks. De Cea and Fernández (1993) proposed the concepts of transit route and rout section for the need in modeling transit networks. A transit route is defined by a sequence of nodes including origins, destinations and all intermediate transfer nodes. A route section is defined by a segment between any two consecutive transfer nodes. In other words, a transit route can also be expressed by a series of transit route sections.

The transit network considered in this paper is a transit network defined by route sections. The definitions of route sections and transit routes are matching to the concepts of links and paths in road networks. A transit network can be represented by a directed graph $G = (N, S)$, where $S$ is the set of route sections. Let $I$ and $J$ be the sets of origins and destinations and let $R_{ij}$ be the set of transit routes between origin-destination (O-D) pair $(i, j)$, $i \in I, j \in J$; $d_{ij}$ denotes the passenger demands between O-D pair $(i, j)$; $h^r_{ij}$ denotes the passenger flows on transit route $r \in R_{ij}$; $v_s$ denotes the passenger flows on route section $s \in S$. Let $v$ and $h$ denote two column vectors of all route section flows and transit route flows, i.e., $h = (\cdots, h^r_{ij}, \cdots)^T$ and $v = (\cdots, v_s, \cdots)^T$. According to the conservation law of passenger flows, it follows that

$$\sum_{i \in I} \sum_{j \in J} \sum_{r \in R_{ij}} h^r_{ij} = d_{ij}, i \in I, j \in J$$

$$v_s = \sum_{i \in I} \sum_{j \in J} \sum_{r \in R_{ij}} \delta^r_{sr} h^r_{ij}, s \in S$$

$$h^r_{ij} \geq 0, r \in R_{ij}, i \in I, j \in J$$

Where $\delta^r_{sr}$ is the transit route /route section incidence indicator. That is, $\delta^r_{sr} = 1$ if transit route $r \in R_{ij}$ passes route section $s$; otherwise, $\delta^r_{sr} = 0$. 

572
2.2 Elastic Frequency of a Transit Line

The frequency of a transit line is defined by the journey time of the transit line divided by the number of available transit vehicles (fleet size) of the line. Let \( t_m^l \) and \( \lambda_n^l \) denote the vehicle’s travel time and dwelling time on line segment \( m \) and node \( n \) for transit line \( l \). The journey time for a transit line can be calculated by summing the dwelling time at terminals, travel time on each line segment, and dwelling time at each line station (Lam et al., 2002), that is:

\[
T_j = \alpha_l t_0^l + \sum_{m \in M_l} t_m^l + \sum_{n \in N_l} \lambda_n^l
\]  

(4)

Where \( t_0^l \) is the terminal time of line \( l \), \( M_l \) and \( N_l \) are the set of line segments and stations of line \( l \), parameter \( \alpha_l \in \{1, 2\} \). When transit line \( l \) has only one terminal in its journey, then \( \alpha_l = 1 \); otherwise, \( \alpha_l = 2 \). The latter case means that there are two terminals in the transit line.

Wirasinghe and Szplett (1984) pointed out that the vehicle’s dwelling time at a station is governed by the volumes of boarding and alighting passengers, i.e., the number of all interchanging passengers. The numbers of alighting and boarding passengers at a station \( n \in N_l \) for transit line \( l \) can be calculated as below (Lam et al., 1998):

\[
Al_n^l = \sum_{s \in S} \delta^-_{ns} \xi_{ls} x^l_s v_n, n \in N_l, l \in L
\]  

(5)

\[
Bo_n^l = \sum_{s \in S} \delta^+_{ns} \xi_{ls} x^l_s v_n, n \in N_l, l \in L
\]  

(6)

Where \( Al_n^l \) and \( Bo_n^l \) are the volumes of alighting and boarding passengers; \( x^l_s \) is the proportion of passengers in route section \( s \) choosing line section \( l \); \( \delta^-_{ns} \), \( \delta^+_{ns} \) and \( \xi_{ls} \) are the spatial relations between stations, route sections and transit lines, namely:

\[
\delta^-_{ns} = \begin{cases} 1, & n = i^- (s) \\ 0, & n \neq i^- (s) \end{cases}
\]  

(7)

\[
\delta^+_{ns} = \begin{cases} 1, & n = i^+ (s) \\ 0, & n \neq i^+ (s) \end{cases}
\]  

(8)

\[
\xi_{ls} = \begin{cases} 1, & l \in A_s \\ 0, & l \notin A_s \end{cases}
\]  

(9)

Where \( i^- (s) \) and \( i^+ (s) \) are tail and head of route section \( s \).

According to Lam et al. (1999), dwelling time \( \lambda_n^l \) for a transit vehicle on line \( l \) at node \( n \) can be described by a continuous function with respect to both boarding and alighting passenger volumes:

\[
\lambda_n^l = \varphi (Al_n^l, Bo_n^l), n \in N_l, l \in L
\]  

(10)

The dwelling time is also a function of passenger flow pattern \( v \):

\[
\lambda_n^l = \lambda_n^l (v), n \in N_l, l \in L
\]  

(11)

Eqn. (4) indicates that the journey time of a transit line is a function of passenger flow pattern \( v \), i.e.,

\[
T_j = T_j (v), l \in L
\]  

(12)

Hence, the frequency of transit line \( l \) can be calculated by using formula:
\[ f_i = \frac{N_i}{T(v)}, i \in L \]  

Where \( N_i \) is number of available vehicles for transit line \( i \). Eqn. (13) exhibits that the frequency of transit line is a function of passenger flows on route sections rather than constants.

### 2.3 Passengers’ Travel Time Functions on Route Sections

For route section \( s \in S \), the average passenger travel time, denoted by \( t_s \), consists of two components: average in-vehicle travel time \( t_s^0 \) and average combined waiting time \( u_s \) (reflecting waiting time at the tail node of the route section). Without loss of generality, it is assumed that the average in-vehicle travel time is fixed, but the average combined waiting time is a function of the frequencies of transit lines traversing that route section. For example, if we assume that passengers board the first arriving transit vehicle from transit lines passing route section \( s \), then the average combined waiting time can be evaluated as follows (see, Spiess and Florian, 1989):

\[ u_s(f) = \frac{\alpha}{f_i}, s \in S \]  

where \( A_s \) is the set of transit lines traversing route section \( s \), \( f \) is a column vector of all transit line frequencies, i.e., \( f = (\cdots, f_i, \cdots)^T \), and parameter \( \alpha \in (0,1] \). Various values of parameter \( \alpha \) have different physical meanings in charactering passengers’ average waiting times at stations. For instance, \( \alpha = 1 \) corresponds to a case that the distribution of arriving vehicles of a transit line follows the exponential distribution with mean of \( 1/f_i \) and a uniform passenger arrival rate at a station. Hence, the average travel time for passengers on a route section is a function of the frequency of transit services, namely,

\[ t_s(f) = t_s^0 + u_s(f), s \in S \]  

### 2.4 Assumptions

It is assumed that passengers follow logit-based SUE criterion to travel from an origin to a destination in a transit network (Lam et al., 1999). Namely, the passenger flows on a transit route between an O-D pair should satisfy the logit-based SUE conditions as follows:

\[ h_{ij}^\gamma = d_{ij} \exp\left(-\theta_i c_{ij}^\gamma(f)\right) \sum_{k \in R_j} \exp\left(-\theta_j c_{kj}^\gamma(f)\right) r \in R_j, i \in I, j \in J \]  

Where \( \theta > 0 \) is the travel time dispersion coefficient in the logit model, and \( c_{ij}^\gamma(f) \) is the passengers’ travel time on transit route \( r \in R_j \) between O-D pair \( (i, j) \), which is defined by

\[ c_{ij}^\gamma(f) = \sum_{s \in S} t_s(f) B_{aw} \]  

The probability of passengers using transit route \( r \in R_j \) can be calculated by

\[ P_{ij}^\gamma(f) = \exp\left(-\theta_i c_{ij}^\gamma(f)\right) \sum_{k \in R_j} \exp\left(-\theta_j c_{kj}^\gamma(f)\right), r \in R_j, i \in I, j \in J \]  

The probability of passengers traversing a route section can be calculated as follows.
Thus, the logit-based SUE passenger flow on a route section can be evaluated by formula:

\[ v_s = \sum_{i \in I} \sum_{j \in J} d_{ij} p_s^j (f), s \in S \]  

(20)

For each O-D pair \( \langle i, j \rangle \), the expected minimum travel time can be calculated as follows (Sheffi, 1985).

\[ \pi_{ij} = -\frac{1}{\theta} \ln \sum_{r \in R} \exp \left( c_{ij} (f) \right), i \in I, j \in J \]  

(21)

3. A COMBINED TRANSIT TRIP DISTRIBUTION/SUE TRANSIT ASSIGNMENT MODEL WITH ELASTIC FREQUENCIES OF TRANSIT LINES

3.1 A Variational Inequality Model

At each origin \( i \in I \), it is assumed that there are induced trips, denoted by \( a_i \), generated by land-use development projects. Each destination \( j \in J \) is associated with a continuously differentiable function, denoted by \( \eta_j (d_j) \), where \( d_j \) is the number of passengers destined to destination \( j \), to describe the attractiveness of this destination (Oppenheim, 1994). Furthermore, it is assumed that these induced trip generations choose their destinations following the logit model with the expected minimum O-D travel times (Oppenheim, 1994). Therefore, transit trip distributions for passengers should satisfy the following conditions in the logit model.

\[ q_{ij} = a_i \frac{\exp \left( -\theta_2 \pi_{ij} + \eta_j (d_j) \right)}{\sum_{j \in J} \exp \left( -\theta_2 \pi_{ij} + \eta_j (d_j) \right)}, i \in I, j \in J \]  

(22)

Where \( \theta_2 > 0 \), a parameter in the logit model for trip distribution, and \( \pi_{ij} \), defined in eqn. (21), is the expected minimum travel time between O-D pair \( \langle i, j \rangle \). However, the logit-based SUE flows for route sections \( v \) in eqns. (14) must aggregate passenger flows from the existing O-D demands \( \{ d_{ij}, i \in I, j \in J \} \) and the induced trip generations \( \{ a_i, i \in I \} \), i.e.,

\[ v_s = \sum_{i \in I} \sum_{j \in J} \sum_{r \in R} \delta_{ir} \left( h_{ir}^{1j} + h_r^{2j} \right), s \in S \]  

(23)

\[ \sum_{i \in I} \sum_{j \in J} h_{ij}^{1j} = d_j, i \in I, j \in J \]  

(24)

\[ \sum_{i \in I} \sum_{j \in J} h_r^{2j} = q_j, i \in I, j \in J \]  

(25)

\[ h_{ij}^{1j}, r \in R, i \in I, j \in J \]  

(26)

\[ h_r^{2j} \geq 0, r \in R, i \in I, j \in J \]  

(27)

Both transit route flow patterns \( \{ h_{ij}^{1j}, r \in R, i \in I, j \in J \} \) and \( \{ h_r^{2j}, r \in R, i \in I, j \in J \} \) should fulfill the logit-based SUE condition as depicted by eqn. (16). Nevertheless, \( d_j \), the number of passengers with destination \( j \in J \), can be calculated by

\[ d_j = \sum_{i \in I} \sum_{r \in R} \left( h_{ij}^{1j} + h_r^{2j} \right) = \sum_{i \in I} \left( d_{ij} + q_j \right), j \in J \]  

(28)
Let $h^1$ and $h^2$ be two vectors of all transit route flows with respect to the fixed O-D demands and the induced trip generations, i.e., $h^1 = (\cdots, h^{ij}_1, \cdots)^T$ and $h^2 = (\cdots, h^{ij}_2, \cdots)^T$; $q$ is the vector of trip distribution associated with the induced trip generations, namely, $q = (\cdots, q_j, \cdots)^T$. Thus, the combined transit trip distribution/SUE transi assignment with elastic transit frequencies is to seek a set of transit route flows $h^1, h^2$, trip distribution $q$ and transit frequency $f$ such that they satisfy eqns (13), (16) and (22) simultaneously.

Define three vector functions:

$$F_i(h^i, f) = \left(\cdots, \frac{1}{\theta_i} \ln h^{ij}_i + \sum_{s \in S} (t_s + u_s(f)) \delta^{ij}_w, \cdots\right)^T \quad (29)$$

$$F_2(h^2, f) = \left(\cdots, \frac{1}{\theta_i} \ln h^{2ij}_i + \sum_{s \in S} (t_s + u_s(f)) \delta^{ij}_w, \cdots\right)^T \quad (30)$$

$$F_3(q) = \left(\cdots, \frac{1}{\theta_i} \ln q_j + \eta_j \left(\sum_{i \in l} (d_{ij} + q_{ij})\right), \cdots\right)^T \quad (31)$$

By using the generalized Kurash-Kuhn-Tucker (KKT) conditions for variational inequalities, it is easy to verify that $(h^*, h^{2*}, q^*, f^*)$ is the solution of the combined transit trip distribution/SUE transit assignment with elastic frequencies if and only if it is the solution for the variational inequality as follows:

Finding a $(h^*, h^{2*}, q^*, f^*) \in \Omega$ such that

$$F_1(h^*, f^*)^T (h^1 - h^*) + F_2(h^{2*}, f^*)^T (h^2 - h^{2*}) + F_3(q^*)^T (q - q^*) \geq 0, \quad \forall (h^1, h^2, q, f) \in \Omega \quad (32)$$

Where $\Omega$ is the set of all the feasible solutions:

$$\Omega = \left\{(h^1, h^2, q, f) \mid (h^1, h^2, q, f) \text{ fulfills eqns. (13) and (22) - (27)}\right\} \quad (33)$$

It can be seen that set $\Omega$ includes two kinds of constraints: passenger flow conservation equations expressed by eqns. (23)-(27) and the nonlinear relation between transit line frequencies and passenger flows on route sections. Considering the simplicity and imposed length limitation of this paper, here we do not get into the details of the existence and uniqueness of the solution of the proposed VI model.

### 3.2 A Solution Algorithm

In the case of fixed frequency of transit lines, the resultant VI model (32) is equivalent to the following strictly convex minimization problem due to the sufficient and necessary conditions for nonlinear programming problems (Bazaar et al., 1993).

$$\min g(h^1, h^2, q) = \sum_{s \in S} (t_s + u_s) v_s + \frac{1}{\theta_1} \sum_{i \in I} \sum_{j \in J} \sum_{r \in R_{ij}} h^{ij}_i (\ln h^{ij}_i - 1) + \frac{1}{\theta_1} \sum_{i \in I} \sum_{j \in J} \sum_{r \in R_{ij}} h^{2ij}_i (\ln h^{2ij}_i - 1)$$

$$+ \left(\frac{1}{\theta_2} - \frac{1}{\theta_1}\right) \sum_{i \in I} \sum_{j \in J} q_j (\ln q_j - 1) + \sum_{j \in J} \int_0^{\tau_j + \delta_j} \eta_j(y) dy$$

subject to constraints (23)-(27).

The above convex minimization problem can be efficiently solved by the method of successive average (MSA) for the SUE traffic assignment problems (Sheffi, 1985). Having
the route section flows from model defined in eqn. (34), transit line frequency \( f \) can be calculated by eqn (13). Repeating this process can form an algorithm to solve the foregoing VI model, which is named as the diagonalization method—a variation of conventional diagonalization method for solving a variational inequality.

**Diagonalization Method**

Step 0: Given an initial transit line frequency \( f^{(m)} \), let \( m = 0 \).

Step 1: For given \( f^{(m)} \), solve the convex minimization problem (34) to find the route section flows denoted by \( v^{(m)} \).

Step 2: Calculate the frequency of transit lines by eqn. (13). i.e.,

\[
N_i / T_i(v^{(m)}), l \in L
\]

Step 3: If \( \max_{i \in k} |f_i^{(m+1)} - f_i^{(m)}| \leq \varepsilon \), for a predetermined tolerance \( \varepsilon > 0 \), then stop; otherwise, set \( m = m + 1 \) and go to Step 1.

4. A GENERALIZED BILEVEL PROGRAMMING MODEL AND AN ALGORITHM FOR THE MAXIMAL RESERVE TRIP GENERATION

4.1 Two Types of Physical Constraints for Induced Trip Generation

Let \( o_i^{\text{max}} \) and \( d_j^{\text{max}} \) be the upper bounds for transit trips generated from origin \( i \in I \) and attracted to destination \( j \in J \). These two bounds reflect the physical capacities in trip generations and attractions at origins and destinations for land-use development. For example, the number of households will basically set the maximal possible number of transit trips generated from the zone, and the number of factories and their sizes may set an upper bound for the number of transit trips that may be attracted. Given an induced transit trip generation pattern arising from a land-use development project, denoted by a column vector \( o = (\cdots, o_i, \cdots)^T \) where \( i \in I \), transit trip distribution \( q(o) \) obtained from VI model (32) is a function of these trip generations. Hence, this induced trip generation should fulfill the following two kinds of physical constraints:

\[
o_i \leq o_i^{\text{max}} - \sum_{j \in I} d_{ij}, i \in I
\]

\[
\sum_{i \in I} q_{ij}(o) \leq d_{ij}^{\text{max}} - \sum_{i \in I} d_{ij}, j \in I
\]

Note that right hand sides of eqns. (35)-(36) give the upper bounds of reserve transit trip generation at origins and trip attractions to destinations, respectively, given the existing transit OOD demand pattern \( \{d_{ij}, i \in I, j \in J\} \).

4.2 Capacity Constraints for Passenger Flow on a Route Section

As vehicles operating for transit lines have physical capacities in accommodating the number of passengers, the total number of passengers of a route section can not exceed the sum of vehicle capacities from transit lines passing the route section when an induced trip generation pattern \( o \) is given, i.e.,

\[
v_s(o) \leq C_s(f(o)), s \in S
\]

Where \( (v(o), f(o)) \) can be obtained by solving the VI model (32); \( C_s(f(o)) \) is the capacity of route section \( s \) and it is a function of the frequency of transit lines. A typical capacity of a route section can be defined by
\[ C_i (f(o)) = \sum_{l \in L} \xi_{il} f_i (o) K_l \]  

Where \( K_l \) is the maximal number of passengers that a vehicle of transit line \( l \) can accommodate.

### 4.3 Transit Equity Constraints

As previously discussed, land-use and development projects will induce extra transit trip generations that may result in an equity issue in the form of average travel times of passengers between different O-D pairs in a transit network. Let \( \bar{c}_{ij} \) be the average travel time between O-D pair \( \langle i, j \rangle \) before any land-use development project is implemented. It can be calculated by the following equations:

\[ \bar{c}_{ij} = \sum_{s \in S} p_{ij}^s (f^* (t_s + u_s (f^*))) \in I, j \in J \]  

Where \((h^*, f^*)\) is the solution of the logit-based SUE transit assignment problem formulated as a VI problem:

Find a \((h^*, f^*)\) \( \in \Omega \) such that

\[ F_i (h^*, f^*) (h^1 - h^{1*}) \geq 0, \text{ for any } (h^1, f) \in \Omega_i \]  

Where \( \Omega_i \) is the set of feasible transit route flows and line frequencies, namely,

\[ \Omega_i = \left\{ (h^1, f) \right\} \text{ fulfills eqns. (13), (24) and (26)} \]  

The aforementioned diagonalization method can be adopted for solving VI model (40). The probability of route section usage defined in eqn. (19) can be calculated without path enumerations by using Dial’s or Bell’s stochastic loading methods (Bell, 1995).

After implementation of a land-use development project, further transit trip generation \( o \) will be induced. The resulting average travel time for O-D pair \( \langle i, j \rangle \) is thus a function of the induced trip generation, denoted by \( c^o (\gamma) \), which can be expressed by

\[ c^o (\gamma) = \sum_{s \in S} p_{ij}^s (f^* (\gamma)) (t_s + u_s (f^* (\gamma))) \in I, j \in J \]  

Where \((h^o (\gamma), h^{2*} (\gamma), q^{*} (\gamma), f^* (\gamma))\) is the solution of VI model (32). It can be expected that the range of ratios of average travel times for different O-D pairs before and after the implementation of the land-use development project may be very wide. In other words, it is unfair for those passengers between some O-D pairs to suffer from longer average O-D travel times because of the new land-use development. To alleviate such an inequity problem, eqn. (43) is introduced to prevent the land-use development project from causing inequity in average O-D travel time.

\[ \gamma_1 \leq \frac{c^o (\gamma)}{\bar{c}_{ij}} \leq \gamma_2, i \in I, j \in J \]  

Where \( \gamma_1 \) and \( \gamma_2 \) are two parameters to set the bounds of changes of O-D average travel time in order to meter the issue of equity. In fact, eqn. (43) guarantees that the ratio of change in average O-D travel time for passengers between each O-D pair should be within the interval \([\gamma_1, \gamma_2]\) after extra transit trips are induced. It is the worst case for an O-D pair that its average O-D travel time increases \( (\gamma_2 - 1) \) percents as compared with the average O-D travel time without the induced trip generations.
4.4 A Generalized Bilevel Programming Model

The maximal transit trip generation problem is to find the maximal addition transit trip generation \( o \) subject to the above physical, capacity and transit equity constraints while taking account passengers’ SUE route choice behavior and logit-based trip distribution for the induced transit trip generation. It can be formulated as a generalized bilevel programming model:

\[ \begin{align*}
\max & \quad o = (..., o_i, ...)^T \\
\text{subject to} & \quad \text{constraints (35), (36), (37) and (43).}
\end{align*} \]  

(44)

where \( \{c_{ij}(o), i \in I, j \in J\} \), \( q(o) \), \( v(o) \) and \( f(o) \) involved in constraints (36), (37) and (43) can be obtained by solving VI model (32) for any given transit trip generation \( o \).

In this model, the upper-level problem (44) is a multi-objective optimization problem, and the lower-level problem is a VI model for the combined transit trip distribution/SUE transit assignment with elastic frequencies. To solve the above generalized bi-level programming model, we adopt a weighted summation of the induced trip generation for each O-D pair to convert the multi-objective function (44) into a single objective function:

\[ F(o) = \sum_{i \in I} w_i o_i \]  

(45)

Where \( \{w_i, i \in I\} \) are weights associated with O-D pairs, which represent the strength of zonal land-use developments. A bilevel programming model for the maximal reserve transit trip generation is formulated as:

\[ \max_{o} F(o) = \sum_{i \in I} w_i o_i \]  

(46)

subject to

\[ o_i \leq o_i^\max - \sum_{j \in J} d_{ij}, i \in I \]  

(47)

\[ \sum_{i \in I} q_{ij}(o) \leq d_{ij}^\max - \sum_{i \in I} d_{ij}, j \in J \]  

(48)

\[ v_s(o) \leq C_s(f(o)), s \in S \]  

(49)

\[ \gamma_s \leq \frac{c_{ij}(o)}{v_{ij}} \leq \gamma_2, i \in I, j \in J \]  

(50)

Where \( \{c_{ij}(o), i \in I, j \in J\} \), \( q(o) \), \( v(o) \) and \( f(o) \) are obtained by solving VI model (32) for any given induced transit trip generations \( o \).

4.5 A Genetic Algorithm Embedded with the Diagonalization Method

Because of the inherent non-differentiability, it is usually hard to have efficient algorithms for solving the generalized bilevel programming problems. The average O-D travel times \( \{c_{ij}(o), i \in I, j \in J\} \), trip distributions \( q(o) \), route section flows \( v(o) \) and transit frequencies \( f(o) \) are all continuous, non-convex, and non-differentiable functions with respect to the induced trip generations \( o \) in the bilevel programming model as presented in eqn. (46)-(50). Hence, we employ a genetic algorithm (GA) because GA can solve the bi-level programming problems. Although there is no analytical expression for nonlinear constraints (48)-(50), their values can be calculated using the diagonalization method. The fitness function, shown in eqn. (51), is constructed by using the penalty function approach from nonlinear programming problems:
\[
\hat{F}(o) = F(o) - \rho \left\{ \sum_{i \in I} \max \left( 0, o_i - o_i^{\max} + \sum_{j \in J} d_{ij} \right) + \sum_{j \in J} \max \left( 0, \sum_{i \in I} q_{ij}(o) - d_j^{\max} + \sum_{i \in I} d_{ij} \right) \right. \\
\left. \sum_{i \in S} \max \left( o_i, v_s - C_s(f(o)) \right) + \sum_{i \in I} \sum_{j \in J} \max \left( 0, \gamma_{ij} - c_{ij} / \bar{c}_{ij} \right) + \sum_{i \in I} \sum_{j \in J} \max \left( 0, c_{ij} / \bar{c}_{ij} - \gamma_2 \right) \right\} 
\]  

(51)

Where \( \rho \) is a sufficient large penalty parameter such that the value of the fitness function will be small enough when any of the constraints (48)-(50) is violated. Thus, the upper-level optimization problem in the above bilevel programming model has become an unconstrained minimization problem:

\[
\max \hat{F}(o) 
\]  

(52)

A GA consists of four main operations: chromosome coding, selection, mutation and crossover. The induced trip generation, \( o = (\cdots, o_i, \cdots) \), is defined as a chromosome in real number. A GA embeded with the diagonalization method for solving the above generalized bilevel programming model can be stated as follows:

Step 0: (Initialization) Generate random population of \( P \) chromosomes.

Step 1: (Calculation of the fitness) Evaluate the fitness function \( \hat{F}(o) \) of each chromosome \( o \) in the population after executing the preceding diagonalization method for solving VI (32) with respect to the trip generation pattern \( o \).

Step 2: (Generation of s New population) Create a new population by repeating the following steps until the new population is complete.

Step 2.1: (Selection) According to the fitness function values calculated in Step 1, use the rank selection method to choose two parent chromosomes from the current population.

Step 2.2: (Crossover) With a crossover probability, \( \tilde{p}_c \), cross over the parents to form new offspring (children). If no crossover was performed, offspring is the exact copy of parents.

Step 2.3: (Mutation) With a mutation probability, \( \tilde{p}_m \), mutate new offspring at each locus (position in chromosome).

Step 2.4: (Accepting) Place new offspring in the new population.

Step 3: (Replace) Use new generated population for a further run.

Step 4: (Stopping criterion) If a stopping condition is satisfied, stop, and return the best solution in current population. Otherwise, go to Step 2.
5. A NUMERICAL EXAMPLE

To illustrate the proposed models and algorithms, a test network consists of 7 transit lines, shown in Figure 1, extracted from the real transit network in Singapore is used. In this example, we consider the transit trips from Hougang New Town and Jurong East to Pasir Panjing and Queenstown via three intermediate transfer nodes, i.e., Clementi, NUS and Queenstown. That is, $I = \{1, 2\}$ and $J = \{3, 4\}$. Table 1 gives the number of available transit vehicles, the vehicle capacity and travel times of line segments for each transit line.

![Figure 1 A transit sub-network in Singapore](image)

Table 1 Basic data for transit lines

<table>
<thead>
<tr>
<th>Transit line</th>
<th>Description</th>
<th>Number of available vehicles and their capacities</th>
<th>Travel time for line segments</th>
</tr>
</thead>
<tbody>
<tr>
<td>$L_1$</td>
<td>Bus line 188</td>
<td>20 (80 passengers)</td>
<td>10 minutes (nodes 1 to 5), 12 minutes (nodes 5 to 6), 14 minutes (nodes 6 to 3)</td>
</tr>
<tr>
<td>$L_2$</td>
<td>Mass Rapid Transit Line I</td>
<td>19 (900 passengers)</td>
<td>8 minutes (nodes 1 to 5), 10 minutes (nodes 5 to 4)</td>
</tr>
<tr>
<td>$L_3$</td>
<td>Bus line 96</td>
<td>11 (80 passengers)</td>
<td>10 minutes (nodes 5 to 6)</td>
</tr>
<tr>
<td>$L_4$</td>
<td>Bus line 10</td>
<td>11 (80 passengers)</td>
<td>11 minutes (nodes 6 to 3)</td>
</tr>
<tr>
<td>$L_5$</td>
<td>Mass Rapid Transit Line II</td>
<td>15 (900 passengers)</td>
<td>31 minutes (nodes 2 to 5)</td>
</tr>
<tr>
<td>$L_6$</td>
<td>Mass Rapid Transit Line III</td>
<td>16 (900 passengers)</td>
<td>35 minutes (nodes 2 to 4)</td>
</tr>
<tr>
<td>$L_7$</td>
<td>Bus line 33</td>
<td>17 (80 passengers)</td>
<td>24 minutes (nodes 5 to 4)</td>
</tr>
</tbody>
</table>
Table 2 Transit routes and route sections

<table>
<thead>
<tr>
<th>O-D pair</th>
<th>Transit paths</th>
<th>Transfer nodes</th>
<th>Route sections</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-3</td>
<td>R1 3</td>
<td>S1 (L1)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>R2 5</td>
<td>S3 (L1, L2), S4 (L1)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>R3 6</td>
<td>S2 (L1), S6 (L1, L4)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>R4 5,6</td>
<td>S3 (L1, L2), S5 (L1, L3), S6 (L1, L4)</td>
<td></td>
</tr>
<tr>
<td>1-4</td>
<td>R5 5</td>
<td>S3 (L1, L2), S8 (L2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>R6 6</td>
<td>S2 (L1), S9 (L7)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>R7 5,6</td>
<td>S3 (L1, L2), S5 (L1, L3), S9 (L7)</td>
<td></td>
</tr>
<tr>
<td>2-3</td>
<td>R8 5</td>
<td>S7 (L5), S4 (L1)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>R9 5,6</td>
<td>S7 (L5), S5 (L1, L3), S6(L1,L4)</td>
<td></td>
</tr>
<tr>
<td>2-4</td>
<td>R10</td>
<td>S10 (L6)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>R11 5</td>
<td>S7 (L5), S8 (L2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>R12 5,6</td>
<td>S7 (L5), S5 (L1, L3), S9 (L7)</td>
<td></td>
</tr>
</tbody>
</table>

Table 3 O-D Demand

<table>
<thead>
<tr>
<th>Origin</th>
<th>Destination</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2,000</td>
<td>1,000 passengers/hour</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>500 passengers/hour</td>
<td>1,500 passengers/hour</td>
<td></td>
</tr>
</tbody>
</table>

Table 2 lists the transit network represented by route sections and transit routes. In addition, Table 3 presents the existing O-D demands for passengers. The upper bounds for the number of trips generated from origins and attracted to destinations are assumed to be $o_{i}^{\text{max}} = 5000$ passengers/hour, $o_{2}^{\text{max}} = 4500$ passengers/hours, $d_{3}^{\text{max}} = 5000$ passengers/hour and $d_{4}^{\text{max}} = 5000$ passengers/hour.

Let us assume that $\alpha = 0.5$ as passenger waiting time defined by eqn. (14), and that $\alpha_{j} = 2$ for all seven transit lines in calculating transit line journey time defined by eqn. (4). The transit vehicles’ dwelling time at a station is of a linear function (50):

$$d_{i}^{l} = 1.5 + 0.083B_{i}^{l} + 0.05A_{i}^{l}, i \in L, n \in N$$

(53)

The attractive function for a destination has the form (Yang and Meng, 1998):

$$\eta_{j}(d_{j}) = \beta_{j}^{0}d_{j}^{\nu} - \beta_{j}^{0}, j \in J$$

(54)

Where parameters $\beta_{j}^{0}$, $\beta_{j}^{1}$ and $\kappa_{j}$ are given in Table 4.

Table 4 Coefficients in attractive function at a destination

<table>
<thead>
<tr>
<th>Destination</th>
<th>$\beta_{j}^{0}$</th>
<th>$\beta_{j}^{1}$</th>
<th>$\kappa_{j}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>0.5</td>
<td>0.25</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>1.2</td>
<td>0.25</td>
<td>2</td>
</tr>
</tbody>
</table>

This example adopts that $\theta_{1} = 1.5$ and $\theta_{2} = 1.0$ in the logit model, and that $\varepsilon = 0.01$ for the preceding diagonalization method. Further, we use uniform weight $w_{i} = 1.0, i \in I$ for the
weighted function defined in eqn. (45). Regarding the equity constraint (43), we set $\gamma_1 = 0$ and we will examine the impact of various values of parameter $\gamma_2$ in the maximal reserve trip generation problem.

Given any trip generation pattern $(o_1, o_2)$, applying the preceding diagonalization method for this example can yield route section flows, trip distributions and transit line frequencies. Due to the length limitation of this paper, we only present the changes of transit line frequencies with respect to different levels of total trip generations, i.e., $o_1 + o_2$, as shown in Figure 2.

![Figure 2](image-url) Changes of transit line frequencies versus different levels of total trip generations

According to Figure 2, it can be seen that the transit line frequencies are declined when the total trip generations are increased. This is because that the travel time of each line is increased in this case. It is also noted that the frequencies of transit lines $L_2$, $L_5$ and $L_6$ change evenly as they are indeed mass rapid transit (MRT) lines. However, the frequency of Line 3 is lower than the frequency of others due to its short journey distance.

For different values of parameter $\gamma_2$, Figure 3 shows the maximal trip generation at each origin by solving the bi-level programming model (46)-(50) using the preceding GA embed with the diagonalization method. According to Figure 3, it can be observed that the transit trip generation at Hougang New Town increases rapidly when parameter $\gamma_2$ becomes larger. Figure 3 also exhibits the land-use potential and projects to what extent the residents could be accommodated by the prevailing transit services. This information could help the transportation planners to evaluate the interactions between land-use and transportation, hence to manage future urban growth and re-development.
6. CONCLUSIONS

This paper has proposed and solved a generalized bilevel programming model that explicitly addresses the equity issue in terms of the changes in the equilibrium O-D travel times between each O-D pair in the transit network. It has also projected the maximum additional trips that each origin could further induce. It is a piece of original work to model the maximal reserve transit trip generations with equity constraints with the consideration of interactions residing between land-use developments and transit services. A VI model for the combined trip distribution/logit-based SUE transit assignment problem with elastic frequencies was first proposed. Then, a generalized bilevel programming model for the maximal reserve transit trip generation with the inclusion of equity constraints is formulated. This paper designed a diagonalization method and GA based algorithm for solving the proposed VI and bilevel programming models. As demonstrated in the presented example, the proposed model and algorithms are proved to be useful tools for planning equity-based transportation and land-use scheme.

ACKNOWLEDGEMENTS

This study was supported by Singapore Ministry of Education through an Academic Research Fund of National University of Singapore (R-264-000-200-112).

REFERENCES


