Journal of the Eastern Asia Society for Transportation Studies, Vol. 6, pp. 2076 - 2091, 2005

EVALUATING PUBLIC TRANSIT CONGESTION MITIGATION MEASURES USING A PASSENGER ASSIGNMENT MODEL

Hiroshi SHIMAMOTO
Ph.D. Student
Graduate School of Engineering
Kyoto University
Yoshida-honmachi, Sakyo-ku, Kyoto
606-8501 Japan
Fax: +81-75-753-5907
E-mail:shimamoto@urbanfac.kuciv.kyoto-u.ac.jp

Fumitaka KURAUCHI
Assistant Professor
Graduate School of Engineering
Kyoto University
Yoshida-honmachi, Sakyo-ku, Kyoto
606-8501 Japan
Fax: +81-75-753-5907
E-mail:kurauchi@urbanfac.kuciv.kyoto-u.ac.jp

Yasunori IIDA
Emeritus Professor
Graduate School of Engineering
Kyoto University
Yoshida-honmachi, Sakyo-ku, Kyoto
606-8501 Japan
Fax: +81-75-753-5907
E-mail:iida@urbanfac.kuciv.kyoto-u.ac.jp

Michael.G.H.BELL
Professor
Centre for Transport Studies
Imperial College London
Exhibition Road, London,
SW7 2AZ, UK
Fax: +44-20-7594-6102
E-mail: m.g.h.bell@imperial.ac.uk

Jan-Dirk SCHMÖCKER
Ph.D. Student
Centre for Transport Studies
Imperial College London
Exhibition Road, London,
SW7 2AZ, UK
Fax: +44-20-7594-6102
E-mail: j-d.schmoecker@imperial.ac.uk

Abstract: This paper proposes a method for evaluating the effect of transit fare systems on passengers' behaviour using a fore-constructed, capacity-constrained, transit assignment model. The optimum fare setting model is formulated as a mathematical programming with equilibrium constraints (MPEC) problem in which passengers are assumed to choose the minimum cost paths where a user equilibrium condition is achieved. Both total cost and connectivity reliability, which is defined here as the *probability of arriving at the destination without failing to board at any station*, are adopted as the evaluation index of the problem. The equality of the congestion mitigation effect for each OD pair is considered explicitly utilising the idea of the *Gini coefficient*, which is the economic indicator. The proposed model is solved using NSGA-II (Non-dominated Sorting GA), which is one method for solving a multi-objective optimisation problem. Finally, a numerical example for a simple network is shown.

Key Words: transit assignment, MPEC, multi-objective optimisation, NSGA-II

1 INTRODUCTION

Daily, passengers world-wide experience the problem of overcrowded transit vehicles. Even where transit overcrowding is not endemic, events like football matches can lead to temporary overcrowding. However, passengers often use one particular line, while another, which is less attractive, has available space. One reason for this situation is the fare system, *i.e.*, passengers try to take a direct line whenever possible, to avoid transfer fares, as most bus networks in

Japan are not part of a flat fare system. Therefore, fare systems affect the congestion on a transit network. With the introduction of IC card charge systems by many transit companies, passengers do not have to pay a fare every time they board. Therefore, operators will adopt more flexible fare systems in the future. Based on these considerations, this study evaluated the effect of transit fare systems on passenger behaviour, focusing on setting the optimum fare in the transit network.

Generally, there are two major stakeholders in a transit system: operators and passengers. In this study, we assume that operators want to reduce the congestion on the network under the fixed capacity constraints for all links and have the option to charge more on heavily crowded lines, which is equivalent to congestion charging systems for automobile traffic. Conversely, passengers want to minimise travel costs, which consist of the fare and the monetary value of their travel time. Such a problem can be represented as a Stackelberg game and formulated as a bi-level programming problem, where the operator is the leader and the passengers are followers. It is assumed that the operator can influence, but cannot control, the passengers' route-choice behaviour. In addition, when the lower problem is defined as an equilibrium problem, the overall problem can be formulated as a mathematical programming with equilibrium constraints (MPEC) problem. To date, several attempts have been made to solve the transit-planning problem. Zhou and Lam (2000) proposed a bi-level programming model to decide the optimum fare. In their approach, the objective in the upper problem was to minimise the total network travel time. Gao *et al*. (2004) presented a bi-level programming model to decide the optimum frequency based on a transit assignment model proposed by De Cea and Fernandez (1993). They adopted the approach of minimising the total cost of the network in the upper problem. However, these two models did not consider the effect of congestion mitigation of each of the OD pairs. As Kurauchi *et al*. (2004) found, the effect of congestion mitigation is not spread throughout each OD pair equally: the congestion of some OD pairs is relieved, whereas that of other OD pairs becomes worse. For road transport planning, there are many papers discussing this kind of game. For example, Sumalee (2004) considered the equality of each OD pair in a road pricing scheme model. He formulated the model as a bi-level programming problem using multi-objectives and adopted social welfare improvement, revenues, and distribution equity impact as objective functions. Distribution equity impact is defined using the idea of the *Gini coefficient*, which is the economic indicator often used to measure income inequity.

By contrast, some passengers likely cannot get on a vehicle due to limited capacity. It is important to represent the capacity constraint explicitly because, as mentioned later, connectivity reliability, which is defined here as the *probability of arriving at the destination without failing to board at any station*, is adopted as the evaluation index in this study. De Cea and Fernandez (1993) advanced a model that considers limited line capacity. They concentrate congestion at the transfer nodes and a BPR-type function is used to estimate the wait at each transfer node in terms of the "through flow"-to-capacity ratio. However, this is not based on queuing theory and still allows capacity to be exceeded. Cominetti and Correa (2001) presented a framework for congested transit assignment that can incorporate congestion functions obtained from queuing models. The link waiting time is defined as equal to the inverse of the "effective frequency", which is itself expressed as a function of the vector of link flows. However, they assumed that the arrival rates of passengers are stable during the period modelled; hence if we take this period to be an entire day, demand is apparently smaller than the capacity, even if this is not true during peak periods. An alternative approach proposed by Schmöcker *et al*. (2002) loads passengers while considering capacity constraints explicitly. They proposed a time-dependent network loading approach so that passengers who failed to board in one period are carried forward to the next time interval. However, their model does not implement the common-lines problem, which is an essential feature in the transit network, and means that a passenger who takes the first vehicle to come on his/her "attractive" lines can get to his or her destination earliest (See Chiriqui *et al*. (1975)). Kurauchi *et al*. (2003) implemented the common-lines problem in the model proposed by Schmöcker *et al*. (2002). As presented later, they adopted the idea of the *hyperpath* and transformed a transit network into a graph model to consider the restrained capacity.

Based on these considerations, it is obvious that only a few papers have considered the equality of the effect of each OD pair. Therefore, using the capacity constraint transit assignment model proposed by Kurauchi *et al*. (2003), we attempt to construct a model with multi-objective functions that can set the optimum additional fare.

The transit network representation and notations are shown in the next section. Then, the capacity-constrained transit assignment model with common lines (named CapCon-CL) is presented. The evaluating transit fare systems model is formulated as a multi-objective problem. Finally, a solution algorithm is explored and a case study for a simple network is shown.

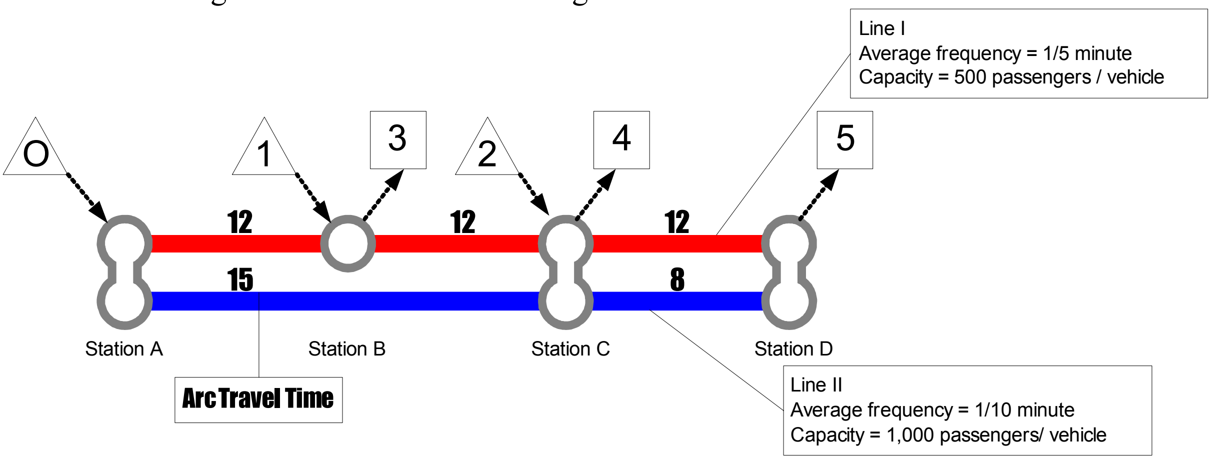
2 NETWORK REPRESENTATION AND NOTATIONS

2.1 Network Representation

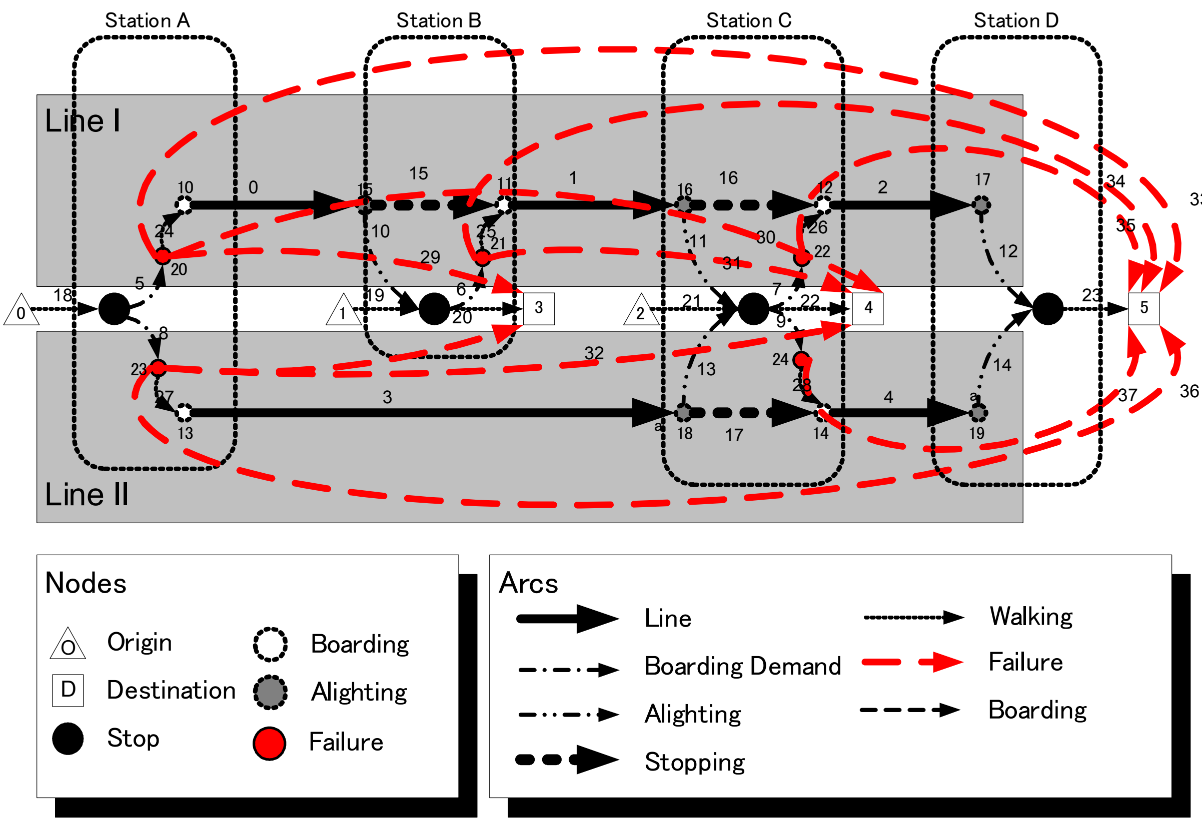
In order to consider the capacities of the transit lines together with the common-lines problem, the transit network shown in Figure 1(a) is transformed into the graph model in Figure 1(b). Bold numbers on the transit lines in Figure 1(a) represent the travel times between stations. The idea follows line-specific arcs proposed by Fearnside and Draper (1971), so that the capacity of transit lines is easily considered, whereas stop nodes are included to explain the common-lines problem in the manner that Nguyen and Pallotino (1988) proposed. Furthermore, to explain a failure to board, failure nodes and arcs are implemented. Nodes are categorised into six types and arcs are categorised into seven types.

An *origin node* represents a trip start node. It has no predecessors and at least one successor. A *destination node* represents a trip end node. It has no successors and at least one predecessor. A *stop node* represents a platform at a station. Any transit lines stopping at the same platform are connected via boarding demand arcs, failure nodes, and boarding arcs. At stop nodes, passengers are assigned to any of the attractive lines in proportion to the arc transition probabilities. A *boarding node* is a line-specific node at the platform where passengers board. An *alighting node* is a line-specific node at the platform where passengers alight. A *failure node* is a node that explains failure to board. When a transit line capacity is exceeded if all passengers board, some of them are forced to use the *failure arc*. It has one predecessor arc from the stop node and (1+ number of destinations) successors. One arc is connected to the corresponding boarding node and the others are connected to each destination node. Note that we assumed that those who fail to board at some stations do not have a priority to board in the next time step in order to deal with the model statically.

A *line arc* represents a transit line connecting two stations. A *boarding demand arc* denotes an arc connecting the stop node to the failure node. The flow on this arc represents the boarding demand for the transit line from a specific platform. An *alighting arc* denotes an arc from an alighting node to a stop node. A *stopping arc* denotes a transit line stopping on a platform after the passengers alight and before the new passengers board. This arc is created to express the available capacity on the transit line explicitly. A *walking arc* connects an origin to a platform or a platform to a destination. A *failure arc* denotes the demand that failed to board. This excess demand is sent directly to its respective destination via this arc. A *boarding arc* is an arc connecting a failure node to a boarding node.



(a) Example Transit Network



(b) Graph Network

Figure 1. Network Representation

Our network representation requires much computer memory because of the many arcs and nodes. However, this may not be a problem considering recent progress in computer technology. The creation of the network can be automated once we prepare the network as shown in Figure 1(a).

2.2 Notations

In order to describe this network representation and formulate transit planning model, we will denote:

L	:	Set of transit lines,
L_c	:	Set of arcs on which additional fare is charged,
A	:	Set of arcs,
A_p	:	Set of arcs on hyperpath p,
A_{ip}	:	Set of arcs on sub-hyperpath of hyperpath p from i to s,
I	:	Set of nodes,
I_p	:	Set of nodes on hyperpath p,
V_p	:	Set of elementary path on hyperpath p,
R	:	Set of origin nodes,
S	:	Set of stop nodes,
S_p	:	Set of stop nodes on hyperpath p,
S_{ip}	:	Set of stop nodes on sub-hyperpath of hyperpath p from i to s,
E	:	Set of failure nodes,
E_p	:	Set of failure nodes on hyperpath p,
E_{ip}	:	Set of failure nodes on sub-hyperpath of hyperpath p from i to s,
V_p	:	Set of elementary paths on hyperpath p,
D_s	:	Set of failure arcs destined to s,
U_l	:	Set of platforms on transit line l,
$i(a)$	:	A head node of arc $a \in A$,
$j(a)$	:	A tail node of arc $a \in A$,
$l(a)$	:	A transit line that is included in arc a,
$IN(i)$	:	Set of arcs that lead into node i,
$IN_p(i)$	:	Set of arcs that lead into node i on hyperpath p,
$OUT(i)$	:	Set of arcs that lead out of node i,
$OUT_p(i)$	:	Set of arcs that lead out of node i on hyperpath p,
w_{kl}	:	Stop arc of line l on platform k,
b_{kl}	:	Boarding demand arc of line l on platform k,
τ_{ap}	:	Arc split probability on hyperpath p,
f_l	:	Frequency of service on transit line $l \in L$,
z_l	:	Capacity of transit line $l \in L$,
t_a	:	Travel time on arc $a \in A$,
c_a	:	Arc cost on arc $a \in A$,
p_f	:	Transit fare (which does not depend on the distance in this study),
s_a	:	Additional cost charged on arc $a \in L$ due to the congestion,
ξ	:	Value for boarding time,
η	:	Value for waiting time,
q_i	:	Probability that a passenger fails to board at node $i \in E$,
θ	:	Parameter for risk of failing to board,
m_{rs}^*	:	Minimum cost from the origin of the hyperpath p (r) to the destination (s)
Ω	:	Set of feasible hyperpath flows satisfying flow conservation,
λ_{lp}	:	Probability of choosing any particular elementary path l of hyperpath p,

α_{ap}	:	Probability that traffic traverses arc *a*,
β_{ip}	:	Probability that traffic traverses node *i*,
δ_{al}	:	Takes 1 if arc *a* is included in *l*, otherwise 0,
ε_{il}	:	Takes 1 if elementary path *l* traverses node *I*, otherwise 0,
$\mathbf{x}$	:	A vector of arc flows,
$\mathbf{y}$	:	A vector of hyperpath flows,
$\mathbf{q}$	:	A vector of fail-to-board probabilities,
$\mathbf{s}$	:	A vector of additional fare charged on the transit line,
N_i	:	The number of passengers of OD pair *i*,
CR_i	:	Connectivity reliability of OD pair *i*,
I	:	The number of OD pair,
P_i^*	:	Optimal set of hyperpaths of OD pair *i*.

3 CAPACITY CONSTRAINED TRANSIT ASSIGNMENT MODEL with Common Lines (CapCon-CL)

3.1 Hyperpath

To obtain an attractive set of transit lines that minimises the expected travel time, we adopt the idea of the *hyperpath* proposed by Nguyen and Pallottino (1988). The hyperpath connecting an origin, *r*, to a destination, *s*, is defined as sets of stops, arcs, and arc transition probabilities, H_p=(I_p, A_p, T_p), where H_p is a hyperpath connecting *r* to *s*, if:
- H_p is acyclic with at least one arc;
- node *r* has no predecessors and *s* no successors;
- for every node $i \in I_p - \{r, s\}$, there is a path from *r* to *s* traversing *i*, and if node $i \notin R$, then *i* has at most one immediate successor;
- the vector, $\mathbf{t}_p$, contains the arc split probabilities that satisfy

$$\sum_{a \in OUT(i)} \tau_{ap} = 1, \ \forall i \in I_p, \tag{1}$$

and

$$\tau_{ap} \geq 0, \ \forall a \in A_p. \tag{2}$$

3.2 Arc Split Probability

Where there are several arcs leading out of nodes on a hyperpath, traffic is split according to τ_{ap}. As shown in Figure 1, traffic may be split at either stop, failure, or alighting nodes.

Stop nodes
We adopt the following assumptions regarding the common-lines problem:
- *Passengers arrive randomly at every stop node*, *and always board the first arriving carrier of their choice set*; and
- *All transit lines are independent statistically with given exponentially distributed headways*, *and a mean equal to the inverse of line frequency.*

Considering these assumptions, τ_{ap} is calculated as follows:

$$\tau_{ap} = f_{l(a)} / F_{ip}, \ \forall i \in S_p, \ \forall a \in OUT_p(i), \tag{3}$$

$$F_{ip} = \sum_{a \in OUT_p(i)} f_{l(a)} \tag{4}$$

Failure nodes
When q_i is not zero, some passengers fail to board. Traffic is split according to the failure-to-board probability, q_i, at the failure nodes.

$$\tau_{ap} = \begin{cases} 1 - q_i & \text{if } a \notin D_s \\ q_i & \text{if } a \in D_s \end{cases}, \forall i \in E_p . \tag{5}$$

Alighting nodes
There may be several arcs leading out of alighting nodes. However, more than one waiting and alighting arc are never included in an optimal hyperpath because getting on the next train on the same line as one gets off is irrational.

3.3 The Cost of Hyperpaths

In this paper, the cost of a hyperpath is represented as a generalised cost, which consists of four elements: the fare, the monetary value of the travel time, the monetary value of the expected waiting time, and the implicit cost associated with the risk of failing to board. Before discussing the cost of hyperpaths, travel time and travel fare are charged on arc a, as shown below, in order to represent the cost of a hyperpath as a generalised cost. Note that in this paper, passengers are charged another fare when they change trains.

$$c_a = \begin{cases} \xi \cdot t_a & (a \in \{L - L_c\}) \\ \xi \cdot t_a + s_a & (a \in L_c) \\ p_f & (a \in b_{kl}, k \in U_l, l \in L) \\ \infty & (a \in D_s) \\ 0 & (\text{else}) \end{cases} \cdot \tag{6}$$

Using the cost of arc a, c_a, the generalised cost of hyperpath p, g_p, can be written as follows:

$$g_p = \sum_{a \in A_p} \alpha_{ap} c_a + \eta \sum_{k \in S_p} \frac{\beta_{kp}}{F_{kp}} - \theta \ln\left(\prod_{k \in E_p} (1 - q_k)^{\beta_{kp}} \right). \tag{7}$$

where,

$$F_{ip} = \sum_{a \in OUT_p(i)} f_{l(a)}, \tag{4}$$

$$\lambda_{lp} = \prod_{a \in A_p} \tau_{ap}^{\delta_{al}}, \forall l \in V_p, \tag{8}$$

$$\sum_{l \in V_p} \lambda_{lp} = 1, \tag{9}$$

$$\alpha_{ap} = \sum_{l \in V_p} \delta_{al} \lambda_{lp}, \ \forall a \in A_p, \tag{10}$$

$$\beta_{ip} = \sum_{l \in V_p} \varepsilon_{il} \lambda_{lp}, \ \forall i \in I_p. \tag{11}$$

The first term of Equation (7) represents the "moving cost" which consists of the fare, any additional fare, and the monetary value of the in-vehicle time. The second and third terms represent the monetary value of the expected waiting time and the cost associated with the risk of failing to board, respectively. The parameter for the risk of failing to board, θ, denotes risk averseness. If $\theta \to \infty$, then passengers are absolutely risk averse, and they are not interested in travel time or expected waiting time; when $\theta = 0$ passengers do not care about failing to board. As Equation (7) can be separated by the subsequence node, Bellman's principle can be applied to find the minimum-cost hyperpath. For simplicity, we treat t_a and f_l as constants.

4 MATHMATICAL FORMULATION

As mentioned below, since the CapCon-CL problem can be formulated using the equilibrium problem, the transit-planning problem discussed here can be formulated using the MPEC (mathematical programming with equilibrium constraints) problem, as shown below.

$$\min_{\mathbf{s}} \psi_m(\mathbf{y}, \mathbf{q}, \mathbf{s}),\ m=1,\dots,M \tag{12}$$

such that
$(\mathbf{y}^*, \mathbf{q}^*)$ satisfies (User Equilibrium) (13)

where $\mathbf{s}$ is a vector of the congestion charging cost on the line arc and M is the number of objective functions. Since the transit fare influences the passengers' route choice, the objective functions of the upper level (Equation (12)) not only depend on the transit fare, but also on the link flow and failure-to-board probability at each platform. Equation (13) means that the flows and failure-to-board probability at each platform satisfy the equilibrium condition mentioned in the next section.

4.1 The CapCon-CL (Lower Problem)

Let us assume that passengers use a hyperpath of minimum cost in Equation (7). The cost of a hyperpath is a function of the failure-to-board probability for each transit line on each platform. By contrast, the failure-to-board probability depends on boarding demand, passengers already on board, and transit line capacity, which in turn depends on the failure-to-board probability. Therefore, the transit assignment model can be formulated as a fixed-point problem, which defines the equilibrium:

Find $(\mathbf{y}^*, \mathbf{q}^*)$ such that

$$\mathbf{y}^* \cdot \mathbf{u}(\mathbf{y}^*, \mathbf{q}^*) = 0,\ \mathbf{u}(\mathbf{y}^*, \mathbf{q}^*) \geq \mathbf{0},\ \mathbf{y} \in \Omega \tag{14}$$

$$\mathbf{q}^* \cdot \mathbf{v}(\mathbf{y}^*, \mathbf{q}^*) = 0,\ \mathbf{v}(\mathbf{y}^*, \mathbf{q}^*) \geq \mathbf{0},\ \forall \mathbf{0} \leq \mathbf{q} \leq \mathbf{1} \tag{15}$$

where

$$u_p(\mathbf{y}^*, \mathbf{q}^*) = g_p(\mathbf{y}^*, \mathbf{q}^*) - m_{rs}^* \tag{16}$$

$$v_{kl}(\mathbf{y}^*, \mathbf{q}^*) = f_l z_l - x_{w_{kl}} - (1 - q_{h_{kl}}) x_{b_{kl}},\ \forall k \in U_l, l \in L \tag{17}$$

As shown in Equation (16), $\mathbf{u}$ denotes a vector of the cost difference between $g_p(\mathbf{y}^*, \mathbf{q}^*)$ and the minimum cost from the origin of hyperpath p (r) to the destination (s). Therefore, Equation (14) represents the user equilibrium condition. Moreover, as shown in Equation (17), $\mathbf{v}$ denotes the vector of vacancies on the line arc on line l from platform k. Therefore, Equation

(15) represents the capacity constraint condition.

The existence of a fixed point is intuitive, since any excess demand simply implies non-zero failures to board. However, because of the non-linear relationship in Equation (15), there is a possibility of multiple fixed points. This fixed-point problem can be solved by combining the method of successive averages and absorbing Markov chains (Kurauchi *et al.*, 2003).

4.2 Formulation of Objective Functions (Upper Problem)

Needless to say, it is desirable to minimise the total cost. Therefore, many transit operators have implemented measures to achieve this objective. However, as Kurauchi *et al.* (2004) pointed out, the congestion-mitigation effect is not equal for each OD pair when the transit operator implements service improvement measures. The congestion level for each OD pair can be measured using connectivity reliability, which we define as the *probability of arriving at the destination without failing to board at any station*, since it is possible in CapCon-CL that some passengers cannot get to their destination because they fail to board at some station. In addition, it is important to secure equity for each OD pair in the transit planning, although this is not often considered in real transit systems. This study considers not only the total cost, (ψ_1), but also the equity of the connectivity reliability of each OD pair, (ψ_2), as shown below:

$$\psi_1(\mathbf{y},\mathbf{q},\mathbf{s}) = \sum_{i=1}^{I} \sum_{p \in P_i^*} y_p \cdot g_p(\mathbf{y},\mathbf{q}) \tag{18}$$

$$\psi_2(\mathbf{y},\mathbf{q},\mathbf{s}) = \frac{1}{2 \cdot N^2 \cdot \overline{CR}} \sum_{i=1}^{I} \sum_{j=1}^{I} N_i \cdot N_j \left| CR_i - CR_j \right| \tag{19}$$

where,

$$CR_i = \sum_{p \in i} y_p \cdot \prod_{k \in E_p} (1 - q_k)^{\beta_{kp}} \Big/ \sum_{p \in i} y_p \tag{20}$$

$$N = \sum_{i=1}^{I} N_i \tag{21}$$

$$\overline{CR} = \sum_{i=1}^{I} CR_i \Big/ I \tag{22}$$

In Equation (19), we formulate ψ_2 to represent the equity of the connectivity reliability of each OD pair using the concept of the *Gini coefficient*, which is often used to measure income inequity and recently has also found applications in operations research (Sumalee, 2004). The *Gini coefficient* takes a value between 0 and 1, where 0 corresponds to perfect equity and 1 to perfect inequity. It is illustrated in Figure 2. On the horizontal axis, a population is ordered by income from the lowest to the highest. The vertical axis is the cumulative share of total income. The most equitable situation is that everyone has the same income and this will produce a straight line (line *Equality*). In reality, the cumulative income distribution will not be a straight line. This is often shown by the '*Lorentz curve*', which illustrates the distribution of resources in a community. The *Gini coefficient* is defined as twice the area between these two curves.

Since the *Gini coefficient* is used as a measure of income equity, one might think that it is reasonable to consider the equity of generalised cost instead of that of the connectivity reliability. However, it is difficult to discuss the equity of generalised cost for each OD pair, since the generalised cost increases as the distance between ODs becomes longer.

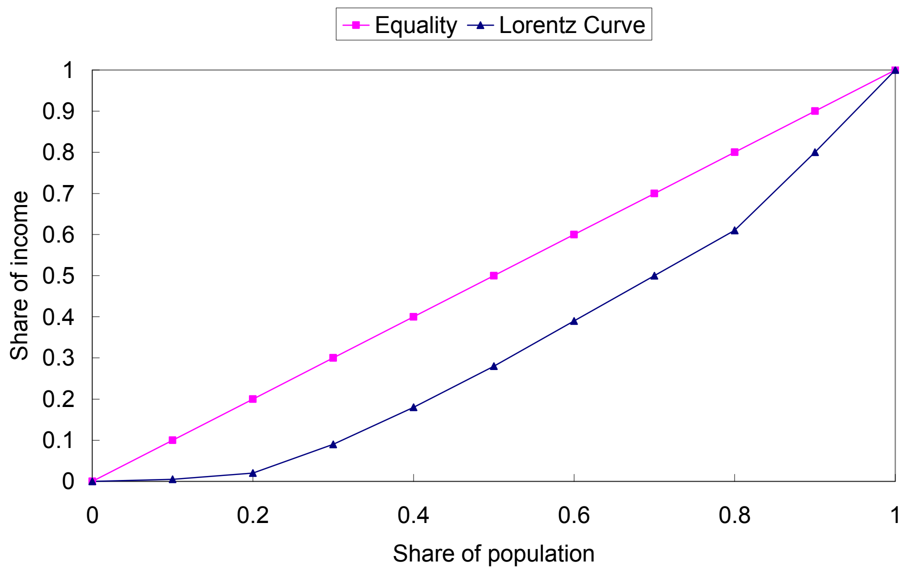


Figure 2. Lorenz Curve and the Gini Coefficient

5 NON-DOMINATED SORTING GENETIC ALGORITHM II (NSGA-II)

Many authors have proposed methods to tackle the multi-objective optimisation problem. To solve the upper problem formulated in the previous section, we utilise the elitist non-dominated sorting Genetic Algorithm (GA) proposed by Deb *et al*. (2000), which requires fewer parameters than other methods. Note that the result might be biased because of the possibility of multiple fixed points in the CapCon-CL. This bias is well known in MINLP (Mixed Integer Non-linear Programming) and is still a challenging problem.

5.1 Overview of the Procedure

NSGA-II is based on the GA (Genetic Algorithm) and the algorithm is outlined below.

(1) (Initialisation)

t:=0;

Set a random (parent) population (P_0) of charging level(**s**) of size N;

(2) (Create offspring population)

Create offspring population Q_t (size N) from P_t using binary tournament selection (with a crowded tournament operator described later) and crossover and mutation operators;

(3) (Combine P_t and Q_t)

Combine parent and offspring populations and create $R_t = P_t \cup Q_t$;

Perform non-dominated sorting (described in the next section) of R_t and identify fronts: F_i, i=1,2,…

(4) (Selection)

i:=1;

Set a new population $P_{t+1}=\phi$;

While $|P_{t+1}|+|F_i|<N$, perform $P_{t+1} = P_{t+1} \cup F_i$ and $i:=i+1$;

(5) (Crowding distance sorting)
if $|P_{t+1}|+|F_i|>N$ **then**
Perform the crowding-sort procedure (described later);
Eliminate $(|P_{t+1}|-N)$ solutions in worse order of crowding distance value;

(6) (Iteration)
$t:=t+1$;
Repeat (2) to (6) until t reaches the fixed number of iterations.

Since all previous and current members are included in R_t at procedure (3), elitism is ensured in NSGA-II.

5.2 Non-domination Sort

The goal for this multi-objective optimisation problem is to find the *Pareto front*, which is a set of non-dominated solutions. Given a set of objective functions, $\psi_1,\ldots,\psi_m$ (assuming the minimisation of all objectives without loss of generality), a solution, $x^{(1)}$, is said to dominate another solution, $x^{(2)}$, if both of the conditions below are satisfied:

a) Solution $x^{(1)}$ is no worse than $x^{(2)}$ in all objectives, or $\psi_m\left(x^{(1)}\right) \le \psi_m\left(x^{(2)}\right)$ for all m,

and

b) Solution $x^{(1)}$ is strictly better than $x^{(2)}$ in all objectives, or $\psi_m(x^{(1)})<\psi_m(x^{(2)})$ for at least one m.

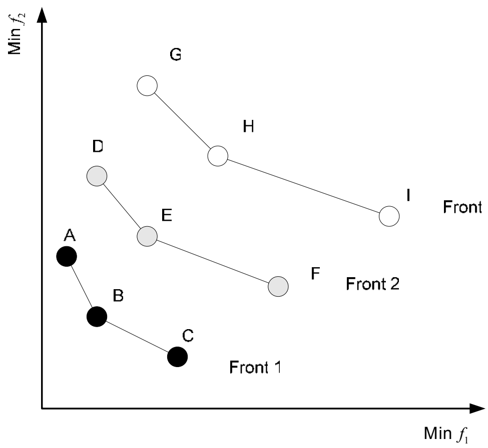


Figure 3. Non-dominating Levels

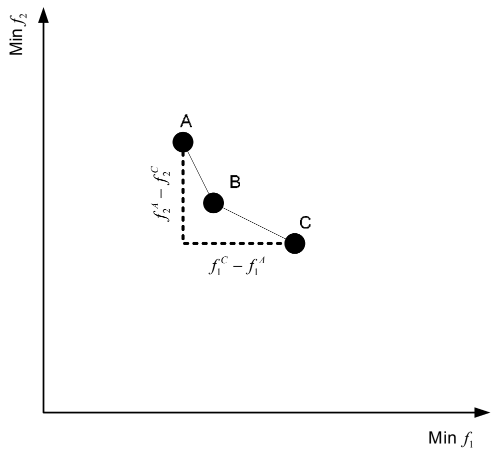


Figure 4. Crowding Distance

In NSGA-II, the population is sorted into different non-dominated levels. Figure 3 shows an example of this process for a problem with two objective functions. The axis shows the values of the objective functions. Since Chromosomes A, B, and C are not dominated by any other chromosomes, these three chromosomes are at the highest level of non-domination and are referred to as Front 1. Chromosomes D, E, and F in Figure 3 are the second level of non-domination and are referred to as Front 2. Similarly, Chromosomes G, H, and I are at the third level of non-domination and are referred to as Front 3.

5.3 Crowding Distance and Crowding-sort

Generally, it is important for heuristic solutions to search from a wide space. NSGA-II can define how spread out each population is without using any parameters, although other solutions need parameters to define them. In NSGA-II, the density of the population in the solution space is adopted as the main indicator of how well the solution performs on the *Pareto Front*. To estimate the density of solutions surrounding a particular solution, i, in the population, the average distance of two solutions on either side of solution i along each axis of the objectives is adopted. This measure, (d_i), represents an estimate of the perimeter of the cuboid formed by using the nearest neighbours as the vertices (called the *crowding distance*). Figure 4 shows an example of the crowding distance. The algorithm for calculating the crowding distance of each point in set F is as follows:

(1) (Initialisation)

Count the number of populations in F as l=|F|;

Set d_i:=0 for i:=1,2,…,l;

(2) (Calculate the crowding distance)

for each objective function m=1,2,…,M

Sort the set in worse order of f_m;

for each populations j=1,2,…,l

if j=0, or j=l **then**

$$d_{I_j^m} := \infty \, ;$$

else

$$d_{I_j^m} := d_{I_j^m} + \frac{f_m^{\left(I_{j+1}^m\right)} - f_m^{\left(I_{j-1}^m\right)}}{f_m^{\max} - f_m^{\min}} \, .$$

where,

f_m denotes the value of the objective function and ψ_m and I_j denote the solution index of the j-th member in the sorted list.

5.4 Crowded Tournament Selection Operator

The crowded comparison operator compares two solutions and returns the winner of the tournament. It assumes that a solution, i, wins a tournament with another solution, j, if and only if any of the following conditions are true:

(i) if solution i has a better rank than solution j, that is, r_i<r_j;

(ii) if two solutions have the same rank, but solution i has a better crowding distance than solution j, that is, r_i=r_j and d_i>d_j

6 NUMERICAL EXAMPLE

The proposed model is applied to the example network shown in Figure 5, which has four stations and three lines in total. The travel times between stations are shown in Figure 5. The frequency and capacity of each line are set to 1 service every 5 minutes and 10 (passengers/minute), respectively. The travel demand between each OD pair and the transit fare are set as shown in Table 1. We assume that it costs 200 yen (about €1.5) when transferring from one line to another. The parameter of risk for failure-to-board, θ, is set to

100 and the time values ξ and η, are 13 (yen/minute) and 26 (yen/minute), respectively. Since a heuristic solution is applied to the proposed model, the charging level ($s_a, a \in L$) should be decided discretely. In this study, the charging level can be decided from 0 to 150 yen, in increments of 10 yen. Calculations with different initial values of **q** reveal that we obtain a unique solution for this network.

Figure 6 shows the Pareto solutions and the Pareto front. As seen in the figure, the Pareto front forms a convex curve. There are ten Pareto solutions in total, as labelled in the figure, whose charged fares are shown in Table 2. For further analysis, we will divide the ten Pareto solutions into three groups: a cost-oriented solution group (group 1), a group that balances cost and equality constraints (group 2), and a congestion mitigation-oriented group of solutions (group 3), as shown in the figure. Table 3 represents the failure-to-board probability for solution 1 in which no additional fare is charged at each platform. This table shows that there is heavy congestion on Line I and no congestion on Lines II and III. Therefore, Table 2 reveals that the additional fare is mainly charged on the crowded line (between A and B, B and C, and C and D).

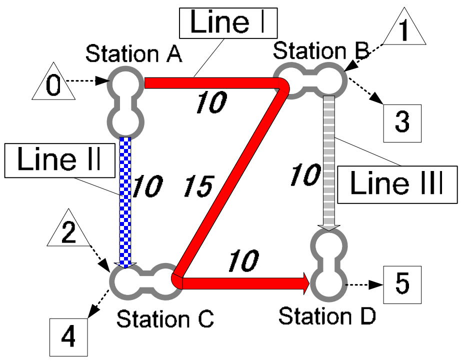


Figure 5. Example Transit Network

Table 1. Travel Demand and Transit Fare

Origin	Destination	Travel Demand (passenger/minutes)	Travel Fare (yen)
0	3	8	
0	4	8	
0	5	15	200
1	4	8	
1	5	8	
2	5	8	

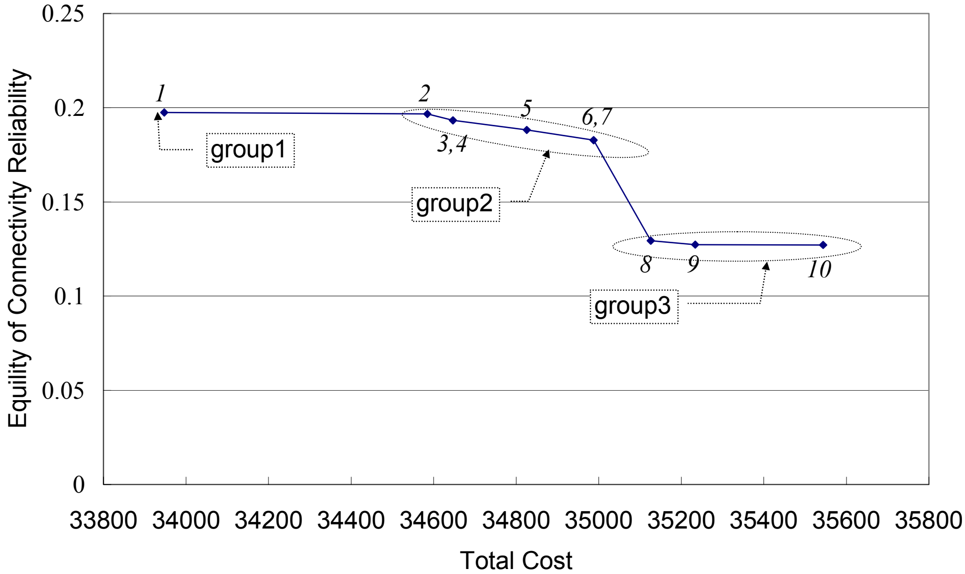


Figure 6. Pareto Solutions and the corresponding Pareto Front

Table 2. Optimal Congestion Fare

	Charged Fare						
Solution No.	AB	AC	BC	BD	CD	Total Cost	*Gini*
1	0	0	0	0	0	33947.3	0.1974
2	0	80	0	0	0	34584.8	0.1967
3	0	0	40	0	0	34646.9	0.1933
4	40	0	0	0	0	34646.9	0.1933
5	0	0	50	0	0	34825.9	0.1881
6	10	0	50	0	0	34987.7	0.1827
7	0	0	60	0	0	34987.7	0.1827
8	0	0	150	0	0	35126.0	0.1294
9	0	0	150	0	10	35234.0	0.1273
10	0	0	150	0	30	35543.9	0.1272

Table 3. The Failure-to-board Probability at each Station (Solution 1)

		Failure-to-board Probability (**q**)
Station A	Line I	0.565
	Line II	0.000
Station B	Line I	0.566
	Line II	0.000
Station C	Line I	0.566

Table 4. Flow of OD (0, 5) in Solutions 1, 5, and 8

Solution No.	Path	Flow
1	A⇒B⇒C⇒D	15.00
	A⇒B⇒D	0.00
	A⇒C⇒D	0.00
5	A⇒B⇒C⇒D	11.58
	A⇒B⇒D	0.00
	A⇒C⇒D	3.41
8	A⇒B⇒C⇒D	1.88
	A⇒B⇒D	5.61
	A⇒C⇒D	7.49

For more detailed analysis, three solutions (Solutions 1, 5, and 8) are selected from all the groups. Table 4 shows the path flow of OD (0, 5) in Solutions 1, 5, and 8. When no additional fare is charged (Solution 1), all the passengers choose a long in-vehicle time path (path A→B→C→D) to avoid the transfer fare. When a intermediate "congestion fare" is charged between B and C (Solution 5), some passengers choose path A→C→D to avoid passing between B and C. Finally, when a high congestion fare is charged (Solution 8), almost all the passengers choose path A→B→D or A→C→D with a short in-vehicle time. Figure 7 shows the connectivity reliability of each OD pair. In the do-nothing case (Solution 1), the connectivity reliability of OD (0, 3), (0, 5), (1, 4), and (2, 5) is low due to the concentration of the flow of OD (0, 5) on Line I. Then, as an additional fare is charged between B and C (Solutions 5 and 8), the connectivity reliability of these OD pairs becomes high because of the effect of flow dispersion of OD (0, 5). However, it should be mentioned that the connectivity reliability of OD (0, 4), which was high in the do-nothing case (Solution 1), decreases in Solutions 5 and 8. This is because some passengers of OD (0, 5) change their route choice and take now Line II.

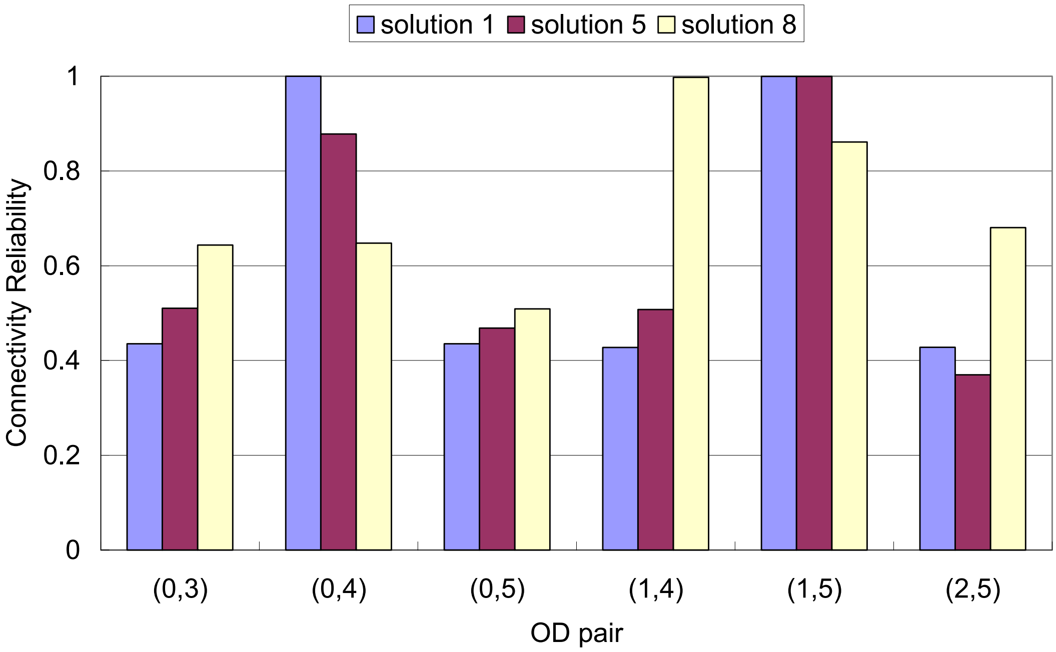


Figure 7. Connectivity Reliability of each OD pair

7 CONCLUSIONS

This paper presented a model formulation and solution algorithm for evaluating the effect of transit fare systems on passengers' behaviour using the capacity-constrained transit assignment model. The transit fare problem discussed here can be represented as a bi-level programming problem, where the transit operator is a leader and the passengers are followers. We assumed that the operators have the option of charging more on the heavily crowded line in order to mitigate congestion and that the passengers choose their travel route in the UE manner under the given transit fare system. The proposed model is formulated with MPEC, which is one of the most challenging problems in optimisation. We adopted two objectives in the upper problem, which are to minimise the total cost and to maximise the equality of the congestion mitigation effect for each OD pair. Furthermore, NSGA-II is used to solve the MPEC, whose upper level is described as a multi-objective problem. The numerical example shows that the additional fare on the crowded line can contribute to the equality of the congestion-mitigation effect; although some OD pairs may suffer from heavier congestion.

The objective of the upper problem needs more consideration, since the equity of the generalised cost for each OD pair is not included. In this context, the instrumental variable for the upper problem can be replaced, for example, with the line frequency or the line capacity. In addition, we need to consider the operation cost in the upper problem. Moreover, it is possible that some passengers will cancel their travel plans or change their travel mode. Therefore, an elastic transit demand should be incorporated into the lower problem. Another technique to overcome this problem is to incorporate a car mode in this model. Then, passengers are assumed to choose the minimum cost mode and route. Finally, application of the model to an actual network is needed, although it is still very time-consuming to express this as a graph model, as mentioned in this paper.

REFERENCES

Chiriqui, C. and Robilland, P. (1975) Common bus lines, **Transportation Science**, **9**, 115-121.

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

De Cea, J. and Fernandez, E. (1993) Transit assignment for congested public transport systems: An equilibrium model, **Transportation Science**, **27**, 133-147.

Deb, K., Agrawal, S., Pratap, A. and Meyarivan, T. (2000) A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-II, **the Parallel Problem Solving from Nature VI (PPSN-VI)**, 849-858.

Fearnside, K. and Draper, D.P. (1971) Public transport assignment - A new approach, **Traffic Engineering and Control**, **12**, 298-299.

Gao, Z.Y., Sun, H, and Shan, L.L. (2004) A continuous equilibrium network design model and algorithm for transit systems, **Transportation Research B**, **38**, 235-250.

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**,309-327.

Kurauchi, F., Shimamoto, H., Iida, Y. and Bell, M.G.H. (2004) Evaluation of public transportation connectivity reliability using capacity-constrained transit assignment model, **Proceedings 2nd International Symposium on Transportation Network Reliability**, Christchurch, 42-47.

Nguyen, S. and Pallotino, S. (1988) Equilibrium traffic assignment for large-scale transit networks, **European Journal for Operational Research***,* **37**, 176-186.

Schmöcker, J.-D., Bell, M.G.H. and Lee, Chi (2002) An application of congested transit network Loading with the Markov chain approach, **Paper presented at 9th Meeting of the EURO Working Group on Transportation "Intermobility, Sustainability and Intelligent Transport Systems"**, Bari.

Sumalee, A. (2004) An innovative approach to option generation for road user charging scheme design: Constrained and multi-criteria design, **Proceedings of the 10th World Conference on Transportation Research**, Istanbul, CD-ROM.

Zhou, J. and Lam. W. H. K (2000) A bi-level programming approach – optimal transit fare under line capacity constraints, **Journal of Advanced Transportation**, **35**, 105-124