Journal of Transportation Technologies, **2012, 2, 220-229**
http://dx.doi.org/10.4236/jtts.2012.23024 Published Online July 2012 (http://www.SciRP.org/journal/jtts)

Optimisation of a Bus Network Configuration and Frequency Considering the Common Lines Problem

Hiroshi Shimamoto[1], Jan-Dirk Schmöcker[1], Fumitaka Kurauchi[2]
[1]Department of Urban Management, Kyoto University, Kyoto, Japan
[2]Department of Civil Engineering, Gifu University, Gifu, Japan
Email: {shimamoto, schmoecker}@trans.kuciv.kyoto-u.ac.jp, kurauchi@gifu-u.ac.jp

Received May 6, 2012; revised June 10, 2012; accepted June 25, 2012

ABSTRACT

Public transportation network reorganisation can be a key measure in designing more efficient networks and increasing the number of passengers. To date, several authors have proposed models for the "transit route network design problem" (TRNDP), and many of them use a transit assignment model as one component. However, not all models have considered the "common lines problem," which is an essential feature in transit network assignment and is based on the concept that the fastest way to get to a destination is to take the first vehicle arriving among an "attractive" set of lines. Thus, we sought to reveal the features of considering the common lines problem by comparing results with and without considering the problem in a transit assignment model. For comparison, a model similar to a previous one was used, formulated as a bi-level optimisation problem, the upper problem of which is described as a multi-objective problem. As a result, although the solutions with and without considering the common lines showed almost the same Pareto front, we confirmed that a more direct service is provided if the common lines problem is considered whereas a less direct service is provided if it is not. With a small network case study, we found that considering the common lines problem in the TRNDP is important as it allows operators to provide more direct services.

Keywords: Transit Network Configuration and Frequency Design; Bi-Level Optimisation Formulation; Transit Assignment Model; Common Lines Problem

1. Introduction

To entice travellers to shift from private cars to public transportation, many public transport operators have taken measures such as reducing off-peak fares. Several researchers have proposed models for determining optimal fares [1,2] or optimal transit frequencies [3]. Another measure could be to make the network configuration more efficient because there are many inefficient bus networks worldwide. Therefore, an optimal public transportation route configuration is necessary as a benchmark to determine the design of a new public transportation route configuration. The problem of designing such a network is referred to as the "transit route network design problem" (TRNDP); it focuses on the optimisation of bus routes or frequencies in order to optimise a number of objectives representing the efficiency of public transportation networks (such as minimising passengers' cost or maximising profit) under operational and resource constraints such as the number and length of public transportation routes, allowable service frequencies, and the number of available buses [4].

Several researchers have proposed models for the TRNDP. For example, [5] also proposed a bus network optimisation model, the objective of which was to maximise the proportion of passengers travelling without transfers, and solved the model by a parallel ant colony algorithm. However, passenger behaviour principles seem not to be described clearly in their model. [6] proposed a model for optimising feeder bus routes, where the transfer point from the railway to a feeder bus was fixed and transferring between feeder buses was not allowed. [7] formulated a simultaneous optimisation problem for railway line configurations and passenger assignments as a linear binary integer problem. Because line frequencies were not determined in their model, they charged a given transfer penalty as an additional waiting time, but any additional waiting time due to a transfer should be defined related to the service frequency (passenger waiting time is short if the service frequency is high). Another feature of their model is that a branch-and-bound method was used as a solution algorithm to obtain an exact solution whereas all previous models had been solved with heuristic algorithms. However, they simplified the network to solve the model within a reasonable time. It would be difficult to apply a strict solution algorithm to a bus network optimisation problem,

which is, generally, more complex than a railway network optimisation problem.

The literature review has so far revealed that many researchers do not describe passenger route choice behaviour accurately. Thus, several papers include a transit assignment model within the TRNDP to consider the passenger route and transfer choice behaviour. [8] combined a line planning model and a traffic assignment model, and demonstrated a solution algorithm based on a column generation method. However, their assignment model was very similar to the traditional assignment model and did not consider the "common lines problem", which is an essential feature in a transit network assignment and is based on the concept that the fastest way to get to a destination is to take the first vehicle arriving among an "attractive" set of lines (the "attractive" set of lines is referred to as the "hyperpath"). [6] demonstrated a model framework for combining the TRNDP and a transit assignment model considering the common lines problem. [9] extended their model to determine the allocation of a limited number of environmentally friendly vehicles and applied the model to a real-size network (although the details of the transit assignment model are not described in either paper). [10] expanded the model of [2] to optimise both the routes and frequencies of public transportation, and they applied their model to a real network.

As described so far, some TRNDPs have considered the common lines problem, while others have not. Considering the common lines problem implies that passengers can consider the complex route set perfectly, which so far was almost impossible in dense networks. However, due to personalised information technology, such as smart phones, passengers can nowadays obtain better knowledge of the complex route set. Also, passengers will tend to use common lines if transit agencies aggregate separate bus stops. Therefore, it is important for transit agencies to know how the optimal network configuration differs if passengers come to know the complex route set. Thus, in this study, we explore the importance of the common lines problem in the TRNDP by comparing results with and without considering common lines. The model used in this study is similar to that of [10], which is formulated as a bi-level optimisation problem, the decision variables of which are the route and frequency of each line, but the following aspects were modified:

- The assumption of a fixed origin/destination for each bus line was relaxed by introducing a dummy origin and destination node.
- Vehicle number constraints were considered more accurately by combining a vehicle assignment procedure and a frequency setting procedure in the solution algorithm.
- Capacity constraint conditions were not considered in this paper to save computational costs ([10] confirmed that capacity constraint conditions did not affect the output of the TRNDP in a real network).

The remainder of the paper is organised as follows. Section 2 describes briefly the minimum cost hyperpath searching problem that is one component of the transit assignment model proposed by [11]. Section 3 describes a mathematical formulation of the bus network optimisation model, and Section 4 shows the solution algorithm. Section 5 illustrates a case study with a simple network, comparing the model with and without considering the common lines problem. Finally, Section 6 provides conclusions and identifies future research.

2. Minimum Cost Hyperpath Searching Problem

In this chapter, the minimum cost hyperpath searching problem, one component of a transit assignment model [11] that is used in the lower problem of the proposed model, is presented briefly.

2.1. Network Representation

To consider the capacities of transit lines together with the common lines problem, the transit network shown in **Figure 1(a)** was transformed into the graph model shown in **Figure 1(b)**. An *origin node* represents a trip start node. A *destination node* represents a trip end node. A *stop node* represents a platform at a station. Any transit lines stopping at the same platform are connected via boarding arcs and failure nodes. At stop nodes, passengers can either take a bus or walk to neighbourhood bus stops, and if they take a bus, they 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. Note that boarding and alighting node are separetely defined in order to allow consideration of dwell time and capacity constraints. *line arc* represents a transit line connecting two stations. A *boarding arc* denotes an arc connecting a stop node to a boarding node. 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 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 (access), a platform to a destination (egress), and neighboring platforms (walk to neighboring platforms).

Generally, the network representation used in public transit assignment models requires more computer memory compared to that used in road traffic assignment

models because of the many arcs and nodes. However, this may be less of a problem considering recent progress in computer technology. Moreover, because the components of the graph network are simple, it is possible to convert automatically from **Figure 1(a)** to **Figure 1(b)**. The minimum cost hyperpath searching problem is now described. If the common lines problem is not considered, it is easy to obtain a minimum cost path by applying a shortest path searching algorithm, such as the Dijkstra method, with the graph network shown in **Figure 1(b)**.

2.2. Notation

We use the following notation regarding the transit assignment model. Other notation will be shown as appropriate.

A_p: Set of arcs on hyperpath p;
L: Set of line arcs;
L_l: Set of line arcs on line l;
U_l: Set of platforms on transit line l;
WA: Set of walking arcs;
BA: Set of boarding arcs;
S_p: Set of stop nodes on hyperpath p;
V_p: Set of elementary paths on hyperpath p;
$OUT_p(i)$: Set of arcs that lead out of node i on hyperpath p;
w_{kl}: Stopping arc of line l on platform k;

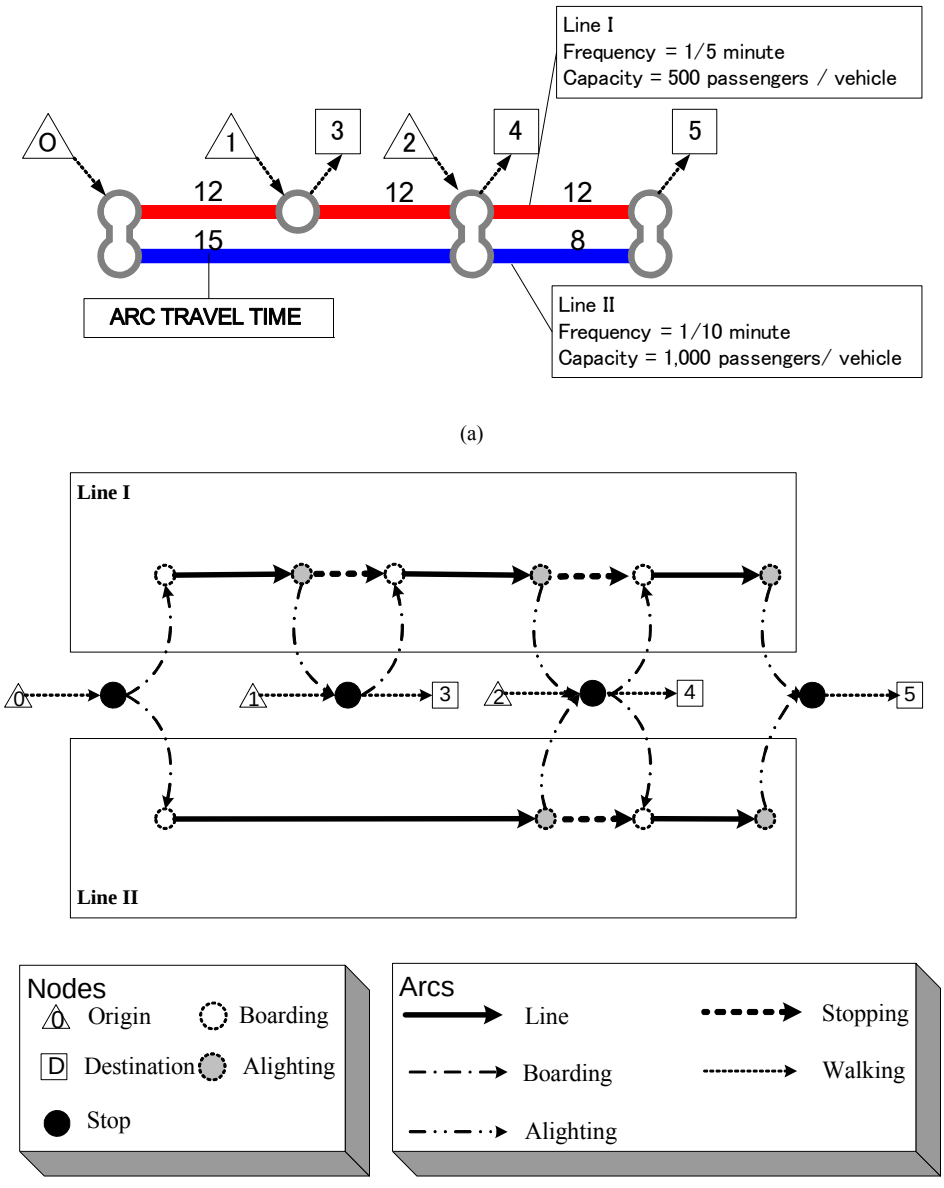

Figure 1. Network representation. (a) Example transit network; (b) Graph network.

b_{kl}: Boarding arc of line l on platform k;
τ_{ap}: Arc split probability on hyperpath p;
$l(a)$: A transit line that is included in arc a;
g_p: Cost of hyperpath p;
c_a: Arc cost on arc $a \in A$;
t_a: Travel time on arc $a \in A$;
ξ: On-board value of time;
ζ: Value of time for walking;
η: Value of time for waiting;
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}: Dummy variable, equal to 1 if arc a is included in l, otherwise 0;
ε_{il}: Equal to 1 if elementary path l traverses node i, otherwise 0;
Y: Vector of hyperpath flows;
X: Vector of arc flow;
f_l: Frequency of line l (1/min).

2.3. Assumptions

We adopted the following assumptions regarding the common lines problem, similar to previous studies (See, [11,12]):

- All bus lines operate with given exponentially distributed headways, and a mean equal to the inverse of the line frequency. The distributions of the lines are assumed to be independent of each other.
- Passengers arrive randomly at every stop node and decide whether to take a bus or walk. If they take a bus, they always board the first arriving vehicle of their choice set.

2.4. Arc Split Probabilities

Where there are several arcs leading out of nodes on a hyperpath, traffic is split according to τ_{ap}. As shown in **Figure 1**, passengers may be split at stop, failure, or alighting nodes. At stop nodes, because passengers cannot simultaneously choose between taking a bus and walking to other platforms, the arc split probability is defined with boarding arcs and walking arcs separately, as shown in Equation (1). If boarding arcs are included in hyperpaths, the arc spilt probability is proportional to the line frequency,

$$\tau_{ap} = \begin{cases} f_{l(a)} / F_{ip} & \left(\forall a \in OUT_p(i) \cap BA\right) \\ 1 & \left(\forall a \in OUT_p(i) \cap WA\right) \end{cases}, \ \forall i \in S_p \quad (1)$$

$$F_{ip} = \sum_{a \in OUT_p(i)} f_{l(a)} \quad (2)$$

2.5. Cost of Hyperpaths

In this paper, the cost of a hyperpath is represented as a generalised cost that consists of three elements: 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. We admit that passengers may walk to another bus stop by creating a walking arc between all stop nodes. Thus, the cost for each arc, c_a, is defined as:

$$c_a = \begin{cases} \xi t_a & \left(a \in L\right) \\ \zeta t_a & \left(a \in WA\right) \\ 0 & \left(\text{else}\right) \end{cases} \quad (3)$$

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

$$g_p = \sum_{a \in A_p} \alpha_{ap} c_a + \eta \sum_{k \in S_p} \frac{\beta_{kp}}{F_{kp}} \quad (4)$$

where

$$F_{ip} = \sum_{a \in OUT_p(i)} f_{l(a)} \quad (5)$$

$$\lambda_{lp} = \prod_{a \in A_p} \tau_{ap}^{\delta_{al}}, \forall l \in V_p \quad (6)$$

$$\sum_{l \in V_p} \lambda_{lp} = 1 \quad (7)$$

$$\alpha_{ap} = \sum_{l \in V_p} \delta_{al} \lambda_{lp}, \forall a \in A_p \quad (8)$$

$$\beta_{ip} = \sum_{l \in V_p} \varepsilon_{il} \lambda_{lp}, \forall i \in I_p \quad (9)$$

The first term of Equation (4) represents the "moving cost," which consists of the monetary value of the in-vehicle time and the walking cost. The second term represent the monetary value of the expected waiting time. Note that α_{ap} and β_{ip} in Equation (4) represent the probability that passengers traverse arc a and node k, respectively, both of which are derived from the probability that the elementary path l within hyperpath p is chosen. As Equation (4) can be separated by the subsequent node, Bellman's principle can be applied to find the minimum cost hyperpath. For simplicity, we treat t_a and f_l as constants.

3. Bus Network Configuration and Frequency Optimisation Model

3.1. Outline of the Model

In the proposed model, we consider two stakeholders, the operator and passengers, and both are assumed to wish to minimise the total travel time and movement cost. Additionally, if it is assumed that the operator knows the pas-

sengers' route choice norm (*i.e.*, minimising the movement cost) and that the operator can influence, but not control, passenger route-choice behaviour, then the proposed model can be formulated as a Stackelberg game or a bi-level optimisation problem, where the operator is the leader and the passengers are followers. If the transit operator tries to minimise the total travel time, the level of service will decrease, causing an increase in passenger cost; thus, the objectives of the stakeholders often conflict. Thus, the upper problem is formulated as a multiple objective optimisation problem. In addition to the assumptions shown in Section 2.3, we make the following assumptions in the proposed model.

First, regarding the bus operation service:

- The position of bus stops is given and fixed, but not all the bus stops have to be used.
- Express service is not considered (*i.e.*, all buses stop at all stops they pass en-route).
- Travel time between bus stops is constant.
- The maximum number of lines is fixed.
- Dwell time is not considered.

Second, regarding passenger movement:

- Passengers choose the minimum cost hyperpath for a given bus network configuration, and they can walk to a different bus stop from an origin or intermediate bus stop if it is cheaper (see the definition of hyperpath cost in Equation (4)).
- The OD demand is fixed regardless of the bus network configuration.

3.2. Model Formulation

The decision variables in the proposed model are the route and frequency of each line, denoted as $r = (r_1, r_2, \cdots, r_{|L|})$ and $f = (f_1, f_2, \cdots, f_{|L|})$, respectively. The proposed model is formulated as:

$$\min_{r,f} \psi_m(y, r, f),\ m = 1,2,\cdots M \tag{10}$$

such that

$$y_p = \begin{cases} d_{rs} & \left(g_p = m_{rs}^*\right) \\ 0 & \left(\text{otherwise}\right) \end{cases},\ p \in H_{rs} \tag{11}$$

$$\sum_{l=1}^{|L|} f_l C_l(r_l) \le NV \tag{12}$$

where

M: Number of objective functions in the upper problem;

$|L|$: Number of lines (fixed);

m_{rs}^*: Minimum cost from the origin of the hyperpath (r) to the destination (s);

d_{rs}: Travel demand between OD pair rs;

$C_l^{\max}$: Upper value of travel time on line l;

NV: Number of available vehicles.

Equation (10) represents the objective function, defined later, and Equation (11) describes the case where all the passengers between each OD pair choose the shortest hyperpath. Passengers are assigned to the shortest hyperpath based on Markov Chain assignments (see [11]). Equation (12) concerns vehicle number constraints. The number of vehicles required to operate a certain line is assumed to be proportional to the line length and frequency, implicitly neglecting turning time or waiting time at the depot. The objective function of the operator is to minimise total operational costs (ψ_1), and the objective function of the passengers is to minimise total movement cost (ψ_2), formulated as:

$$\psi_1(r, f) = \sum_{l=1}^{|L|} f_l C_l(r_l)^2 \tag{13}$$

$$\psi_2(y, r, f) = \sum_{rs \in W} \sum_{p \in H_{rs}^*} y_p \cdot g_p(y) \tag{14}$$

where

W: Set of OD pairs;

h_{rs}^*: Set of hyperpaths between OD pair rs.

Equation (13) represents the total travel time for the operator because the left-hand side of Equation (12) represents the number of vehicles required to operate line l.

4. Solution Algorithm

As shown in Section 3, the proposed model is formulated as the upper problem of a multi-objective optimisation problem. To solve the upper problem, we use the elitist non-dominated sorting genetic algorithm (NSGA-II) proposed by [13], which is an expanded genetic algorithm (GA) that requires fewer parameters than other methods. In this section, only a solution algorithm within one generation is described, which consists of "Vehicle Assignment", "Route Design", and "Frequency Setting". A frequency for each line is determined by combining a "Vehicle Assignment" procedure and a "Frequency Setting" procedure.

4.1. Vehicle Assignment

Figure 2 illustrates the chromosome for vehicle assignment. As shown in the figure, two types of genes are defined: genes representing vehicles (A genes) and genes representing boundaries between neighbourhood lines (B genes). The number of A and B genes are equivalent to the number of available vehicles and maximum number of operated lines, respectively. Using these genes, the number of vehicles for each line is equivalent to the number of A genes sandwiched between two B genes. **Figure 2** illustrates an example of vehicle assignment. In this example, two vehicles are assigned to line 1, three vehicles are assigned to line 2, but no vehicle is assigned to line 3.

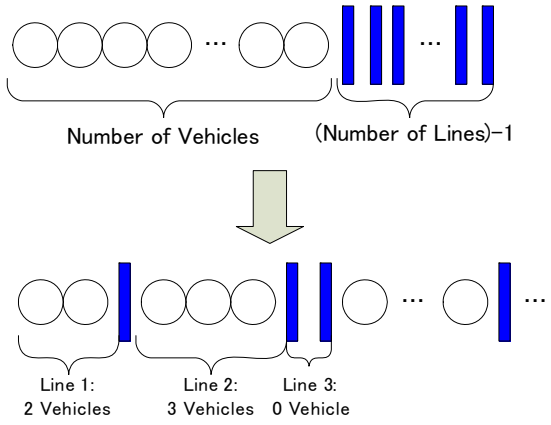

Figure 2. Chromosome illustration for vehicle assignment.

4.2. Route Design

Although many researchers examining the TRNDP use a genetic algorithm (GA) as a solution algorithm, the original GA has some shortcomings for solving the TRNDP, such as generating indirect routes. In this section, a modification of the GA procedure for route search under fixed origin and destination nodes is described following Inagaki *et al*. [14], using the example network shown in **Figure 3(a)**. In the modified GA procedure, the number of genes in a chromosome is the same as the number of nodes in a network N. Each gene m can only take the values of the nodes to which direct links from the node m exist; that is, a link connecting nodes m and n is represented by assigning node ID n to the m^{th} gene. Thus, the alignment of the genes in a chromosome can provide the ID of nodes that make up a route, if one keeps moving ("jumping") from gene m to gene n (in **Figure 3(b)**, these are the genes with a square).

Thus, the chromosome defined here consists of two types of genes, those contributing to the representation of the route and those not. The Proposition in the Appendix 1 shows that we can always obtain a valid route from the genes that contribute to the route description, unless a cyclic route is obtained, which occurs if the same node ID appears in at least two of these genes. **Figure 3(b)** represents the route (0 → 1 → 4 → 6 → 7), with predetermined origin and destination nodes as 0 and 7, respectively. For the genes that are not needed for the route description, a random node ID among the available node IDs is selected. With this chromosome definition, one can trace a unique route from the origin to destination nodes. There are also the well-known elements of crossover and mutation within GA optimisation, as described in (14) (See, Appendix 2).

Note that the modified GA procedure might not create all the possible routes with equal probability. For example in **Figure 3(a)**, the probability of creating route 0-1-3-7 would be higher than that of creating route 0-1-4-6-7 because with above definition, the probability of connecting from node 3 to 7 equals to one (there is only one arc connecting nodes 3 and 7) whereas the probability of connecting node 4 to 6 is 0.33. However, the latter route overlaps also with a number of would have a route with higher overlapping rate (e.g. 0-1-4-5-6-7) than the former route. Therefore, since the modified GA procedure implicitly generates routes that overlap less with higher probability, this procedure is expected to create more variety of routes.

4.3. Frequency Setting

As a result of vehicle assignment and route design procedures, the number of vehicles assigned to each line and the line length are known. Then, the frequency of each line is defined as:

$$f_l = \frac{V_l}{2T_l} \tag{15}$$

where

T_l: Travel time of line l;

V_l: Number of vehicles assigned to line l.

Different from [10], who chose the frequency of each line from four options, the proposed procedures, combining vehicle assignment and frequency setting, can consider vehicle number constraints more accurately.

5. Numerical Example

The proposed model was applied to the simple network shown in **Figure 4**. Bus stops were assumed to exist at each node, and passenger demand was generated at each node. The travel time between bus stops was 4 min by bus and 12.5 min by walking, and the value of time parameters were 13 yen/min (about 0.1 €/min), 26 yen/min,

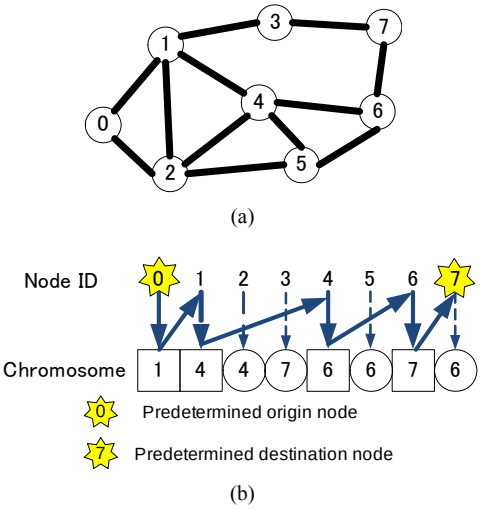

Figure 3. Alignment of genes in a chromosome. (a) Example network; (b) Alignment of chromosomes.

and 50 yen/min, respectively, for boarding time, waiting time, and walking time based on the SP-mode choice survey of [15]. Also, to relax the assumption that the origin and destination of each bus line is fixed, a dummy origin node (Node 25) is introduced, which is an origin of all lines of buses. Further, dummy links with zero cost connect the dummy origin node to all of the bus stops. Similarly, a dummy destination node (Node 26) and zero-cost dummy links connecting all the bus stops with the dummy destination node are added.

With this network, all the demands were assumed to be generated from node 22, with destinations spread to all the other nodes. The demand volume was generated with a uniform random number, taking the value from 0 to 1; 200, 100, and 50 times the random number was defined as the demand to node 12, the grey nodes in **Figure 4**, and other nodes, respectively. As a result, the demand pattern shown in **Table 1** was obtained. Also, the number of available vehicles and maximum number of operated lines were set as 30 and 10, respectively. Note that although the proposed model can be applied to a real network [10], we assume above demand distribution on a small grid network in order to better understand the effect of considering the common lines in TRNDP.

Figure 5 shows the Pareto solutions with and without considering the common lines problem. In total, 96 solutions with and 77 solutions without considering the common lines problem were obtained. The passenger cost for the upper left solutions is lower, whereas the operator cost of those solutions is higher. This relationship between the two stakeholders is reversed for solutions shown in the lower right side of the figure. The extreme case is a solution with zero operator cost, where the operator does not provide any bus services and all the passengers have to walk to their destination. Furthermore, contrary to expectations, solutions with and without considering the common lines problems showed almost the same Pareto front. We suspect that this is due to a fairly low number in GA iterations (1000 iterations with 100 individuals) exploring only a very small range of the possible solutions. Nevetheless we find some interesting differences in the network structures as described in the following. The reason for the fairly low number in GA iterations is due to the computational cost of the lower problem. Especially when considering common lines is the lower problem becomes computationally relatively expensive.

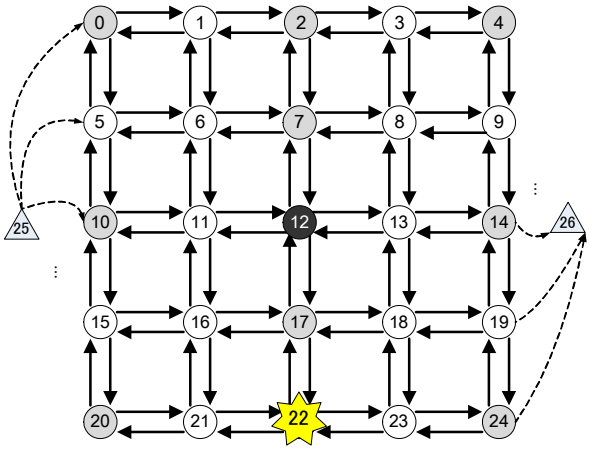

Figure 4. Test network.

Table 1. Assumed OD volume (Origin node is 22).

Destination	Demand	Destination	Demand	Destination	Demand
0	66	8	44	16	43
1	37	9	5	17	48
2	30	10	19	18	3
3	27	11	8	19	10
4	97	12	130	20	32
5	41	13	3	21	23
6	42	14	4	23	17
7	93	15	21	24	7

Because it is impractical to show the results of all the Pareto solutions, we compared the results when the operator's cost = 40. **Figure 6** illustrates the line configuration and headways (inverse of frequency) from the output with and without considering the common lines problem. The thickness and the print of each line was proportional to the operational frequency, as shown in the figure. When considering the common lines problem, many lines were concentrated in the centre of the network (22-17-12-7-2), a "trunk with feeders" network, whereas only one line ran in the centre of the network, a "trunk and branch" network when not considering the common lines problem. The reason for such a difference is that a "trunk with feeders" network brings benefit to those who consider the common lines problem because they can take a line coming first among the bundle. On the other hand, a "trunk with feeders" network does not bring benefit to those who do not consider the common lines problem, because they stick to a single line. Instead, a "trunk with branch" network is more attractive to them because they have more chance to transfer to a branch line from the feeder line. As a result, many direct services from node 22 to node 2 are provided if considering the common lines problem whereas only one direct service from node 22 to node 2 is provided and many passengers are forced to transfer if it is not considered. A transfer penalty was not considered in this study; a different result might have been obtained if we had considered one.

6. Conclusions

This study revealed the effect of considering the common

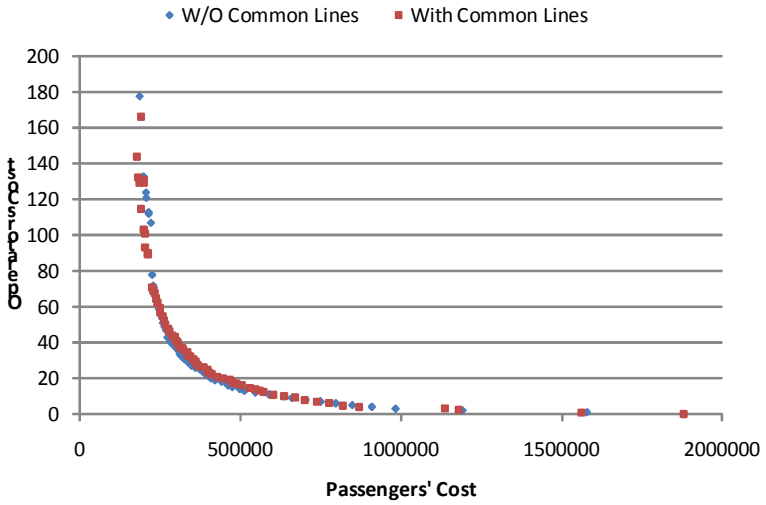

Figure 5. Pareto solutions and the corresponding Pareto front (With/without considering the common lines problem).

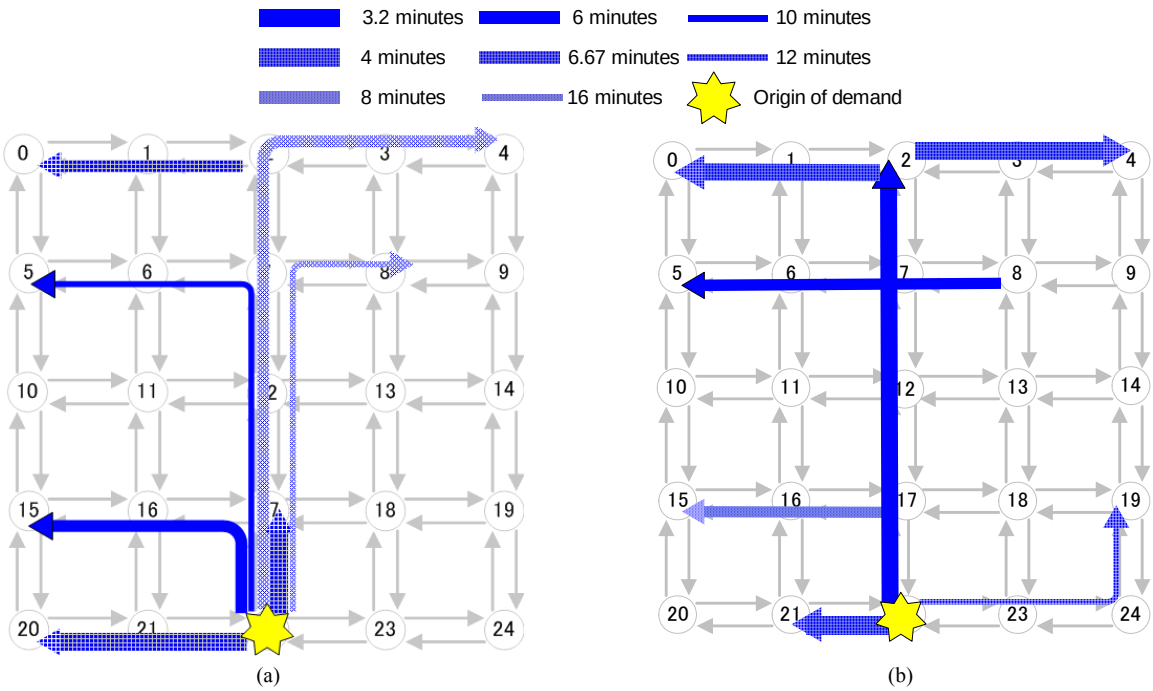

Figure 6. Bus routes and headways (Solution at operator's cost = 40). (a) Considering the common lines problem; (b) Without considering the common lines problem.

lines problem in the transit route network design problem (TRNDP). A similar model to previous ones, formulated as a bi-level optimisation problem, the upper problem of which was described as a multi-objective problem, was used, but several assumptions were relaxed. The lower problem of the model is a passenger assignment model and the output of the TRNDP with and without considering the common lines problem in passengers' route choice was compared. The output of the TRNDP in a simple network was compared with and without considering the common lines problem. It was confirmed that a more direct service is optimal if the common lines problem is considered. This has implications for transit planners and the design of bus stops. Our results suggests that, if bus stops for several lines are arranged in such a way as to make it easy for passengers to interchange between lines, this could result in very different network structures, possibly reducing operator as well as passenger costs.

In future work, it would be valuable to confirm whether this finding is generalisable with various demand patterns and other networks. Furthermore, more

computational efficient algorithms are future research topics, especially in connection with further applications to larger networks.

REFERENCES

[1] J. Zhou, and W. H. K. Lam., "A Bi-Level Programming Approach—Optimal Transit Fare under Line Capacity Constraints," *Journal of Advanced Transportation*, Vol. 35, No. 2, 2000, pp. 105-124. doi:10.1002/atr.5670350204

[2] H. Shimamoto, *et al*., "Evaluation of Public Transit Congestion Mitigation Measures Using Passenger Assignment Model," *Journal of Eastern Asia Transportation Studies*, Vol. 6, 2005, pp. 2076-2091.

[3] Z. Y. Gao, *et al*., "A Continuous Equilibrium Network Design Model and Algorithm for Transit Systems," T*ransportation Research Part B*, Vol. 38, No. 3, 2004, pp. 235-250. doi:10.1016/S0191-2615(03)00011-0

[4] K. Kepaptsoglou and M. Karlaftis, "Transit Route Network Design Problem: Review," *Journal of Transportation Engineering-ASCE*, Vol. 135, No. 8, 2009, pp. 491-505. doi:10.1061/(ASCE)0733-947X(2009)135:8(491)

[5] Z. Yang, *et al*., "A Parallel Ant Colony Algorithm for Bus Network Optimization," *Journal of Computer-Aided Civil and Infrastructure Engineering*, Vol. 22, No. 1, 2007, pp. 44-55. doi:10.1111/j.1467-8667.2006.00469.x

[6] M. Petrelli, "A Transit Network Design Model for Urban Areas," In: C. A. Brebbia and L. C. Wadhwa, Eds., *Urban Transport X*, WIT Press, Southampton, 2004, pp. 163-172.

[7] J. F. Guan, *et al*., "Simultaneous Optimization of Transit Line Configuration and Passenger Line Assignment," *Transportation Research Part B*, Vol. 40, No. 10, 2006, pp. 885-902. doi:10.1016/j.trb.2005.12.003

[8] K. Nachtigall and K. Jerosch, "Simultaneous Network Line Planning and Traffic Assignment," 2008. http://drops.dagstuhl.de/opus/volltexte/2008/1589

[9] B. Beltran, *et al*., "Transit Network Design with Allocation of Green Vehicles: A Genetic Algorithm Approach," *Transportation Research Part C*, Vol. 17, No. 5, 2009, pp. 475-483. doi:10.1016/j.trc.2009.04.008

[10] H. Shimamoto, *et al*., "Evaluation of an Existing Bus Network Using a Transit Network Optimisation Model: A Case Study of the Hiroshima City Bus Network," *Transportation*, Vol. 37, No. 5, 2010, pp. 801-823. doi:10.1007/s11116-010-9297-6

[11] F. Kurauchi, *et al*., "Capacity Constrained Transit Assignment with Common Lines," *Journal of Mathematical Modelling and Algorithms*, Vol. 2, No. 4, 2003, pp. 309-327. doi:10.1023/B:JMMA.0000020426.22501.c1

[12] H. Spiess and M. Florian, "Optimal Strategies: A New Assignment Model for Transit Networks," *Transportation Research Part B*, Vol. 23, No. 2, 1989, pp. 83-102. doi:10.1016/0191-2615(89)90034-9

[13] K. Deb, *et al*., "Fast Elitist Non-Dominated Sorting Genetic Algorithm for Multi-Objective Optimization: NSGA-II," *Parallel Problem Solving from Nature VI* (*PPSN-VI*), 2000, pp. 849-858.

[14] J. Inagaki, *et al*., "A Method of Determining Various Solutions For Routing Application with a Genetic Algorithm (in Japanese)," *Transactions of the Institute of Electronics*, *Information and Communication Engineers*, Vol. 82, No. 8, 1999, pp. 1102-1111.

[15] F. Kurauchi, *et al*., "Experimental Analysis on Mode Choice Behaviour for Merged Public Transport Systems (CD-Rom, in Japanese)," *Proceedings of Infrastructure Planning Conference on Civil Engineering*, 2004.

Appendix 1

1.1. Proposition

The alignment of the genes within a chromosome defined in Section 4.2 ensures that the destination is reached if the route is not cyclic.

1.2. Proof

Suppose none of the genes takes value d, which is the ID of the destination node. Let N_1 indicate the number of genes that describe the route until destination d is reached. Because no node takes value d, the search for a feasible route will continue until all bus stops in the network have been visited so that N_1 will be equal to N, the number of genes or bus stops. Because no cyclic route is included, this means that each bus stop is visited only once. However, because N equals the number of bus stops, this means that one gene must take value d or at least one bus stop must be visited twice, both of which contradict our assumptions.

Q.E.D

Appendix 2

The procedures of crossover and mutation in the modified GA procedure are described using following notations.

r: The predetermined origin node;

s: The predetermined destination node;

N: The set of nodes (representing for bus stops) in a network;

g[*]: Gene of parent;

h[*]: Gene of offspring;

M_{rate}: Mutation rate;

$K_{\max}$: Threshold number.

1) Crossover

The crossover procedure in the modified GA is shown as below. Note that the crossover is executed only when a repeatedly generated random number, is smaller than a given crossover rate.

Step 1 (Initialise)

Set $m \leftarrow r$ and $h[n] \leftarrow \phi$ for $n \in N$;

Step 2 (Roulette selection)

Select two parents randomly g_1 and g_2;

Step 3 (Crossover)

Select one parent i (i = 1, 2) randomly and $h[m] \leftarrow g_i[m]$, $m \leftarrow g_i[m]$;

Step 4 (Termination of loop)

Repeat Step 3 until $m = s$;

Step 5 (Interpolation of the empty genes)

If $g[m] = \phi$ ($m \in N$), then, select one node n going out from node m and $g[m] \leftarrow n$.

As the genes of the offspring will only take the value of a valid node, this crossover procedure always generates a connected route which either is cyclic or valid following the above proposition. **Figure 7(a)** illustrates this crossover. In this example, a route (0→1→4→5→6→7) is generated as a result of the crossover with predetermined origin and destination nodes as 0 and 7.

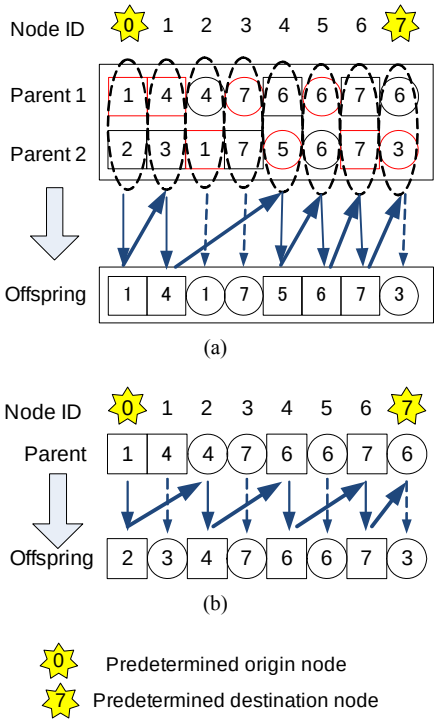

Figure 7. Illustration of mutation and crossover. (a) Example of crossover; (b) Example of mutation.

2) Mutation

The following procedure shows the mutation of the modified GA.

Step 1 (Initialise)

Set $m \leftarrow r$ and $h[n] \leftarrow \phi$ for $n \in N$;

Step 2 (Mutation)

Generate a random number *rnd*;

If *rnd* < *Mrate*, then, select one node n going out from node m and $h[m] \leftarrow n$ and $m \leftarrow n$;

Else

$h[m] \leftarrow g[m]$ and $m \leftarrow g[m]$;

Step 3 (Termination of loop)

Repeat Step 2 until $m = s$;

Step 4 (Interpolation of the empty genes)

If $g[m] = \phi$ ($m \in N$), then, select one node n going out from node m and $g[m] \leftarrow n$.

Figure 7(b) illustrates mutation of the parent chromosome. A route (0→2→4→6→7) is generated as a result of the mutation with predetermined origin and destination nodes as 0 and 7.

