Method for Identifying “Bottleneck” Stations on Inter-regional Public Transport Networks

Takuya WATANABE
Transport Planning and Marketing Laboratory, Signalling and Transport Information Technology Division

Munenori SHIBATA
Transport Planning and Marketing Laboratory, Signalling and Transport Information Technology Division

Takamasa SUZUKI
Transport Planning and Marketing Laboratory, Signalling and Transport Information Technology Division (Former)

This study attempts to develop a method for identifying “bottleneck” stations on the inter-regional public transport network. Investing in these bottleneck stations enables the network to be more convenient. First, a method for identifying bottleneck stations on the network was formulated based on multi-objective optimization with a generic algorithm. Second, a system was developed to automatically identify bottleneck stations in a manner which makes them visible. A network outlining the actual inter-regional network in Japan was constructed. The results of numerical experiments on this network suggest that bottleneck stations may be located not only in the metropolitan areas but also in principal cities in the countryside.

Keywords: inter-regional transport, network evaluation, multi-objective optimization, multi-objective genetic algorithm

1. Introduction

When taking trains, passengers require transfer(s) from other transport modes in order to access railway stations. Improving the accessibility of railway stations for transfer passengers helps improve the overall public transport network. The Railway Technical Research Institute has developed a method for quantitatively evaluating obstacles faced by passengers in transit, with a system which evaluates transfer convenience in stations [1].

Stations which require improvement measures because they have obstacles hindering passengers during transfer are defined as “bottleneck” stations. To the author’s knowledge however, no method for identifying bottleneck stations has yet been developed. A method designed to identify bottleneck stations and capable of predicting the volume of passengers using those stations would therefore be a useful tool to improve the convenience of overall inter-regional public transport.

This paper describes the development of a method to achieve this goal and the methodology and results obtained to identify bottleneck stations using a multi-objective genetic algorithm.

2. A method for identifying “bottleneck” stations

2.1 An approach for identifying “bottleneck” stations

This study focused on the inter-regional public transport network composed of high-speed railways, air travel, coaches, inter alia. An ideal network which combines convenience and produces a lower environmental load than current networks can be achieved by changing the current level of service supply, such as service frequency. In an ideal context, if a station has insufficient capacity for the expected passenger volumes it will form a “bottleneck”, because it will hinder passenger flows. There may be a large difference in the volume of transfer passengers at a bottleneck station between a current network and an ideal network.

The volume of transfer passengers passing through stations in both types of network was computed through the optimization of inter-regional public transport. Bottleneck stations were then identified, as stations where the volumes of passengers in an optimized network exceeded current network volumes. Investments targeting these stations to increase their capacity could then improve convenience to passengers using the public transport network.

2.2 Formulating a method for identifying bottleneck stations

In order to identify bottleneck stations, the optimal state of an inter-regional public transport network needs to be quantitatively evaluated according to certain evaluation criteria. An optimal state was found by searching for routes between any two cities. Passenger flows were then computed for each route, and the evaluation criteria of the network were calculated. Bottleneck stations were then identified by comparing volumes of transfer passengers through stations on the optimized network and the current network. Details of the method for identifying bottleneck stations are explained below, according to the research flow chart shown in Fig. 1.

2.2.1 Route search algorithm

The extended labeling method proposed by Kato et al. [2] was used for finding routes between any two cities. The generalized cost, the linear sum of cost of travel and travel time, were adopted as costs in the labeling method. The value of travel time was set at 46.28 yen per minute on the basis of a route choice model constructed in previous research [3, 4]. No more than 30 routes were found for each pair of cities to save computational resources.
2.2.2 Calculating the probability of choosing each route

Certain routes were multi-modal, whereby passengers used more than two inter-regional modes. The normal logit model tends to overestimate the possibility of choosing such routes. Therefore, the C-logit model was applied. This model was originally proposed by Cascetta et al. [5] to obtain more realistic results for each route by using overlapping ratios and commonality factors. The model parameters are shown in Table 1. Every parameter was statistically significant in this model.

Table 1 Estimated parameters for route choice model

| Explanatory variables       | Estimated value | Statistical t-value |
|----------------------------|-----------------|---------------------|
| Travel time (hour)          | -1.900          | -38.144             |
| Travel cost (10,000yen)     | -1.471          | -8.130              |
| The number of transfer     | -0.572          | -10.875             |
| The inverse of the frequency of services | -1.478 | -7.862 |

Number of samples: 20,048

2.2.3 The passenger flow distribution model

Passenger flows between any two cities were estimated using a gravity model, whose explanatory variables were the product of the populations of both cities, the distance in a straight line and the accessibility index. The accessibility index was obtained by quantifying the convenience of the trip between both cities using the utility function in the C-logit model. The resulting estimated parameters are shown in Table 2. Statistical t-values are the indices indicating the statistical significance of the parameter. Every parameter was judged significant because the absolute value of these t-values exceeded 2.58. The employment of the accessibility index makes it possible to express passenger flow changes associated with the change of the network. In addition, the demand in the future can be estimated by substituting the predicted population into the equation of this model.

Table 2 Estimation of the passenger flow distribution model

| Explanatory variables       | Estimated value | Statistical t-value |
|----------------------------|-----------------|---------------------|
| Constant term              | -22.427 **      | -92.08              |
| Product of the population of both cities | 8.057 × 10^1 ** | 122.78              |
| Accessibility index         | 3.465 × 10^-2 ** | 16.56               |
| Straight distance between points | 2.052 × 10^-1 ** | 10.76               |

Number of samples: 20,321 , **: significant at 1% significance level

2.2.4 Evaluation criteria of the network

Carbon dioxide emissions and consumer surpluses, representing customer gain from improved convenience in monetary terms, were employed as the network evaluation criteria. The objective was a computation of the optimized network by multi-objective optimization of these two criteria.

Consumer surpluses (UB, yen/day) were calculated with a consumer surplus approach [6] as shown in (1) and (2). These equations express the benefits resulting from network improvements indicated by the product of the passenger flow and the change in generalized cost per passenger.

\[
UB = \frac{1}{2} \sum_{OD} (\sum_{i=1}^{k} T_{OD}^i + T_{OD}^k)(C_{OD}^i - C_{OD}^k) 
\]

(1)

\[
C_{OD}^k = \log \left( \sum_{i} \exp(V_{OD}^i) \right) / \alpha_{cost} \]

(2)

where OD : pair of cities, 
\(T_{OD}^i\) : passenger flow between OD in network \(k\), 
\(C_{OD}^i\) : the \(i\)th route between OD, 
\(V_{OD}^i\) : utility of using route OD in network \(k\), 
\(C_{OD}^k\) : generalized cost between OD in network \(k\), 
\(\alpha_{cost}\) : estimated parameter for travel cost in the utility function, and 
\(k\) : state of the network (current network if \(k=0\); optimized network if \(k=1\)).

Total carbon dioxide emissions from the optimized network (\(CO_2\, tCO_2/\text{day}\)) shown in (3) were employed as another evaluation criterion.

\[
CO_2 = \sum_{e} x_{e,\text{freq}} \cdot x_{e,\text{dist}} \cdot UCC_e \]

(3)
where \( x_{e, \text{freq}} \): frequency of service on link \( e \) on the optimized network,
\( x_{e, \text{dist}} \): distance covered by link \( e \) on the optimized network, and
\( \text{UCC}_e \): CO\(_2\) emissions per kilometer from transport \( m \) for link \( e \).

### 2.2.5 Constraints and objective functions

The following four constraints were imposed on the network to optimize the objective functions according to capacity or operational safety.

- The maximum and minimum frequency of service for each link
- The maximum frequency of arrivals to and departures from each node
- The maximum total number of vehicles or aircraft kilometers for each journey
- The maximum transport capacity for each link

The numerical settings for each parameter in relation to these constraints are shown in Table 3 based on the frequency of services at peak hours on the current network or landing slots at each airport.

For the purpose of reducing the calculation load, the lagrangian relaxation method was used to deal with the fourth constraint. This method is an optimization method which incorporates a part of the constraint into an objective function. The objective functions that should be optimized are written in (4) and (5) by setting the penalty term related to passenger overflow, i.e. passengers who are left behind when passenger flows exceed capacity. Then, as shown in (5), the optimization problem with CO\(_2\) emissions was transformed into a maximization problem by applying a negative sign to total carbon dioxide emissions in the optimized network CO\(_2\) as shown in (3).

\[
\text{eval}_{UB} = UB - \lambda_{UB} \sum_e \text{OF}_e \quad (4)
\]
\[
\text{eval}_{CO2} = -\text{CO2} - \lambda_{CO2} \cdot \sum_e \text{OF}_e \quad (5)
\]

where \( \text{OF}_e \): passenger overflow volume at link \( e \),
\( \lambda_{UB} \): coefficient of the penalty term related to consumer surplus, and
\( \lambda_{CO2} \): coefficient of the penalty term related to CO\(_2\) emission.

The coefficient of the penalty term related to consumer surplus can be considered as the inconvenience caused by not being able to occupy a seat, therefore this coefficient \( \lambda_{UB} \) was set at 500 yen per passenger, which is equivalent to the usual seat reservation fee on inter-regional railways in Japan. The coefficient of the penalty term related to CO\(_2\) emissions \( \lambda_{CO2} \) was calculated by multiplying \( \lambda_{UB} \) by the monetary value of CO\(_2\) [4], which was \( 1.29 \times 10^4 \text{ gCO2/pas} \)senger.

### 2.2.6 Identifying "bottleneck" stations

The volume of transfer passengers in each station was calculated for both the optimized network and the current network. This calculation was carried out for each combination of transport modes before and after the transfer. The summation of the increases of the volume of transfer passengers brought about by changing from the current to the optimized network was calculated. The value for stations \( s \), ADTs, was regarded as the criterion for identifying bottleneck stations. A large ADT, value indicates that station \( s \) is a bottleneck station and needs to be improved.

### 3. System for identifying bottleneck stations

Based on the method for identifying bottleneck stations explained in the previous chapter, a system to do the same thing was developed. A target network was designed for evaluation. Below is a summary of the system.

#### 3.1 Formation of a network for evaluation

A network for evaluation consists of nodes and links.

| Table 3 Parameters set according to constraints in each transport mode |
|--------------------------|-----------------|-----------------|-------------------|-----------------|
| Transport Mode | Air | Coach | Shinkansen | Limited express train | Extra transport |
| Frequency of services for each link (/day) | 1 ~ 90 | 1 ~ 500 | 9 ~ 270 | 1 ~ 100 | 1 ~ 500 |
| Frequency of arrivals to and departures from each node (/day) | 1400 | 1500 | 600 | 400 | 2000 |
| Total train or air kilometers for each journey (10000 kilometers) | 150 | 120 | 60 | 40 | 20 |
| Transport capacity for each link (passengers) | 350 | 50 | 1200 | 500 | 2000 |

\( \dagger \): Explained in 3.1.2
Nodes indicate transport terminals such as stations, and links represent transport paths between nodes. The following sections explain in detail how the nodes and links were set.

### 3.1.1 Formation of nodes on the network

Japan is divided into 207 inhabited zones identified by the government [7], and one node is given to each zone as the representative transport terminal. The location of the main station in the representative city in each zone is considered as the location of the node. Additionally, several junction stations are also considered as nodes to replicate the actual railway network as much as possible in the simulation. A total of 217 nodes were designated, including 207 representative zone nodes and 10 junction nodes.

### 3.1.2 Formation of links on the network

Links are formed by connecting nodes to replicate the actual transport network. They are given mode by mode, i.e., different modes between the same nodes have different links. Modes considered were air, Shinkansens (high speed trains), limited express trains, and coaches. These modes were surveyed in Inter-regional Travel Survey (hereafter referred to as IRTS) [7]. Each link possessed attributes based on the level of services (hereafter referred to as LOS) data. The LOS of each link includes travel time, cost of travel, travel distance, and frequency of services. Data (both the LOS data and network topology) used to construct the network was mainly data from December 1, 2010, the weekday survey date of the 5th IRTS.

Travel time for each air link was calculated as the sum of the average flight time and the waiting time. A uniform waiting time of 80 minutes was set for all flights, which included time for check-in, security inspection, baggage claim, and waiting for egress transport, etc. In addition, travel time and cost of access and egress trips between airports and representative zone nodes were taken into consideration. This procedure was not applied to transits from air to air because no access or egress trips are necessary.

The cost of travelling on the limited express train was calculated based on a linear function which depends on travel distance in order to save computational resources. The regression coefficient of the linear function was calculated with the least-squares method by using the LOS data for 82 routes which originated and arrived in central stations for every prefectural capital where passengers can travel without transfers.

As for Shinkansen links, since the number of intermediate stops differs among train types, a “Super links” were employed. These represent major links connecting only major stations where only the fastest trains stop.

The fare calculation method applied to limited express train links was also applied to the Shinkansen network. Parameters were derived using fare data for the 85 routes which originated and arrived in central stations for every prefectural capital where passengers can travel without transfers.

Most coach routes operate a “closed door system”. This system limits stops where passengers can embark and disembark. Therefore, the network for finding routes and the network for deciding the frequency of services were developed separately. Travel time and cost of access and egress trips between stops and representative zone nodes were considered in the same manner as for air routes.

Even after developing the air, Shinkansen, limited express train and coach networks, several nodes remained off the network. In reality, however, these nodes can be reached with other local transport modes including local railways, regular-route buses, ferries and so on. Therefore, extra transport links were also provided to replicate the transport service between these “remote” and other nodes.

A holistic overview of the air network and inter-regional railway network described above are shown in Fig. 2.

![Image of the created air and railway network](image-url)
3.2 Network optimization based on multi-objective genetic algorithm and identification of “bottleneck” stations

The combination of the frequency of services of each link optimizing two objective functions described in (4) and (5) was obtained by multi-objective optimization. Here, the best solution is not always the obvious solution because there seems to be a trade-off between two objective criteria, inferred from a previous study by the authors [8]. In this study, plural optimized networks, which are called pareto optimal solutions, are obtained by using a multi-objective genetic algorithm (Fig. 3). The computational procedure is shown in Fig. 4. The fitness indicator of each solution was computed based on the function values of the two objective functions by using the pareto ranking method and the sharing method [9]. Then, solutions with a high fitness indicator were kept as possible solutions for the next generation by roulette selection. On repeating this procedure until reaching the preset number of iterations, the optimized network was produced as the output. Then, bottleneck stations were identified by comparing the optimized network and the current network.

3.3 Example of output from the system

The developed system has a function which displays passenger flows for each link and can highlight the bottleneck stations. It also possessed a function which displays information for a selected link or node, such as the frequency of services, the volume of transfer passengers, etc. A screen shot of how the highlighted bottleneck stations appear in the developed system is shown in Fig. 5.

Fig. 5 Screen shot of highlighted bottleneck stations

4. Case analysis using the developed system

The identification of bottleneck stations of the developed network described in chapter 3.1 was conducted as a system trial. The following parameters were set for the genetic algorithm: the population size in one generation was set at 50, the number of a generation was set at 1000, the mutation rate was set at 0.9, and the crossover rate was set at 0.1.

The distribution of pareto optimal solutions in the last generation are shown in Fig. 6. There were 34 pareto opti-

Fig. 6 Screen shot of highlighted bottleneck stations

QR of RTRI, Vol. 57, No. 1, Feb. 2016
The relationship between ADT, and the summation of transfer passenger volumes on the current network is shown in Fig. 7. The vertical axis of the figure indicates the size of the station. The larger the value, the larger the terminal station and the number of users.

The decision indicator for bottleneck stations tends to be large in proportion to the volume of passengers. However, there are some stations with relatively small ADT, encircled by the green ellipse in Fig. 7. Such stations are mostly large stations located in metropolitan areas, with relatively smooth passenger flows. Therefore, they are not considered not to be facing serious problems.

On the other hand, there are some stations with relatively large ADT, circled in red in Fig. 7. Such stations are mostly terminal stations in principal cities with fewer transfer passengers. This result suggests that small terminal stations may become bottleneck stations, and therefore improvements are required.

If bottleneck stations can be identified in this way, it will be possible to examine plans for measures to improve station facilities, taking into account surrounding local characteristics.

5. Conclusions and further research

This study developed methods for optimizing the interregional public transport network and identifying bottleneck stations which form an obstacle to smooth passenger flows. Specifically, consumer surplus and CO$_2$ emissions were employed simultaneously as evaluation criteria, and stations with a large difference in the volume of transfer passengers between an optimized network computed by multi-objective genetic algorithm and the current network were identified as bottleneck stations. Numerical analyses using the developed system suggest that small terminal stations in principal cities may become bottleneck stations and therefore require improvement action.

Further research will look at the quantitative evaluation of transfers to other modes of transport [1] taking into account the bottleneck stations identified through this system. Results from this phase of research would make it possible to obtain more detailed transfer paths for each station.

References

[1] Suzuki, T., Muto, M., Matsubara, H., and Yamamoto, M., “Development of a method of evaluation of railway transfer convenience considering characteristics of station surroundings,” RTRI Report, Vol. 27, No. 9, pp. 41-46, 2013 (in Japanese).
[2] Kato, N., Ibara, T., and Mine, H., “An O(Kn) algorithm for K shortest simple paths in an undirected graph with nonnegative arc length,” The Transactions of the Institute of Electronics, Information and Communication Engineers A, Vol. J61-A, No. 12, pp. 1199-1206, 1978 (in Japanese).
[3] Shibata, M., Suzuki, T., and Okuda, D., “Demand simulation of inter-regional railway passengers considering multi-mode route alternatives,” The Journal of the Institute of Electrical Engineers of Japan, Vol. 135, No. 4, pp. 322-326, 2015 (in Japanese).
[4] Ministry of Land, Infrastructure, Transport and Tourism, “Technical guidance for cost-benefit analysis in public project evaluation,” 2009 (in Japanese).
[5] Cascetta, E., Nuzzolo, A., Russo, F., and Vitetta, A., “A modified logit route choice model overcoming path overlapping problems. Specification and some calibration results for interurban networks,” Transportation and Traffic Theory. Proceedings of the 13th International Symposium on Transportation and Traffic theory, pp. 697-711, 1996.
[6] Ministry of Land, Infrastructure, Transport and Tourism, “Project evaluation manual of rail projects,” 2012 (in Japanese).
[7] Ministry of Land, Infrastructure, Transport and Tourism, “Development of a method of evaluation of railway transfer convenience considering characteristics of station surroundings,” RTRI Report, Vol. 27, No. 9, pp. 41-46, 2013 (in Japanese).
ism, “The Fifth Inter-regional Travel Survey,” 2010 (in Japanese).

[8] Watanabe, T., Shibata, M., and Suzuki, T., “Evaluation of inter-regional transportation network considering multi-mode route alternatives,” Proceedings of the Eastern Asia Society for Transportation Studies, Vol.10, in press, 2015.

[9] Gen, M. and Lin, L., Network models and multiobjective GA. Kyoritsu Shuppan Co., Ltd., Bunkyo, Tokyo, Japan, pp. 139-168, 2008 (in Japanese).

Authors

Takuya WATANABE
Researcher, Transport Planning and Marketing Laboratory, Signalling and Transport Information Technology Division
Research Areas: Geospatial information engineering, Transport planning

Munenori SHIBATA, Dr. Eng.
Senior Researcher, Transport Planning and Marketing Laboratory, Signalling and Transport Information Technology Division
Research Areas: Transport planning, Travel behavior analysis, Revenue management on train services, Transport economic evaluation

Takamasa SUZUKI, Dr. Eng.
Researcher, Transport Planning and Marketing Laboratory, Signalling and Transport Information Technology Division (Former)
Research Areas: Transport planning