Optimization System of Reserved/Non-reserved Seating Plans for Improving Convenience and Revenue on Inter-city Trains

Shingo NAKAGAWA
Transport Planning and Marketing Laboratory, Signalling and Transport Information Technology Division

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

Noriko FUKASAWA
Transport Planning and Marketing Laboratory, Signalling and Transport Information Technology Division

Seating plans for reserved/non-reserved seats on inter-city express trains do not necessarily correspond to passenger demand which may vary depending on many factors such as calendar day and/or service section. This mismatch may cause inconvenience to passenger due to congestion experienced on board trains and may cause revenue loss for railway operators. A prototype was developed for a seating allocation optimization system, which is expected to be an effective solution to overcome these problems. Case studies were conducted which showed that optimizing seating plans can be expected to improve both passenger experience and railway operator revenues.

Keywords: inter-city express trains, seating allocation, potential demand estimation, suboptimal choice behavior, revenue management

1. Introduction

Most ordinary coaches on Japanese inter-city express trains have two seat types: (1) reserved seats, which guarantee passengers a fixed seat for the duration of their trip in exchange for an additional fee, usually about 500 yen (equivalent to about 5 U.S. dollars) and (2) non-reserved seats, where passengers are free to sit in any available seats. The drawback of not reserving a seat is that a passenger may have to stand if no free seats are available. A summary of the features for each type of seat is shown in Table 1, and suggests that non-reserved seats offer some advantages over reserved seats, because the passenger is free to board any service and find a seat if one is available without having to pay an additional fare for reservation. However, Japanese inter-city express trains are sometimes very crowded, so there is still demand for reserved seats.

In most cases, the number of cars on a train and allocation of number of seat types are usually fixed yearly on the basis of empirical data. However, fluctuating demand makes it difficult to satisfy passenger needs on these types of trains, indicating that reserved/non-reserved seat allocation on inter-city express trains does not necessarily correspond to passenger demand [1]. The consequence of this

<table>
<thead>
<tr>
<th>Table 1 Features of two types of seats: reserved/non-reserved seats</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reserved seats</td>
</tr>
<tr>
<td>Seat</td>
</tr>
<tr>
<td>Service to board</td>
</tr>
<tr>
<td>Where to seat</td>
</tr>
<tr>
<td>Fee</td>
</tr>
<tr>
<td>Preferred by</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
mismatch in demand and supply of appropriate seating is poorer customer experience and falling revenue for railway operators. As such, a method to optimize seating plans is expected to be an effective solution, as described in [2, 3, 4].

This paper describes the seating plan optimization system which has been developed and the models involved in the system, which make it possible to estimate potential demand and passenger behavior when they meet with congestion. A case study which was conducted demonstrated that optimized seating plans can improve both passenger experience and railway operator revenue.

1.1 Abbreviations and terminology

In this article, we use some abbreviations and terminologies described as below.
- RS (s) ... reserved seat (s).
- NRS (s) ... non-reserved seat (s).
- ICT (s) ... inter-city express train (s).
- LF ... load factor.
- OD ... passenger “origin” and “destination” stations.

Potential demand ... number of people who wish to take ICT. It consists of actual demand and those who finally give up using ICT due to reasons including congestion.

Actual demand ... number of passengers actually used ICT.

2. Model for establishing potential passenger demand

2.1 Existing actual demand data

To estimate the potential demand, it is necessary to know the actual demand of the target ICTs. Three kinds of actual demand data can be obtained, as described in Table 2.

This data is employed to calibrate the potential passenger demand estimation model.

(1) Sales results accumulated in the seat reservation system
All RSs are sold through a seat reservation system. Sales results for RSs for all days and trains can be extracted from the system. This data also gives passenger ODs. Nevertheless, it is important to take into account that sales results do not necessarily correspond to actual demand, because tickets may not necessarily be used in accordance with their validity. It is also necessary to take into account that the data does not provide sales results for NRSs.

(2) OD survey of ICTs
For some ICTs, surveys are conducted to obtain OD data. These surveys are mostly carried during weekdays once a year, by collecting used tickets. The surveys provide OD data with relatively higher accuracy not only for RS passengers but also for NRSs. They do not however give insight into any short-term change (e.g. seasonal change) in demand, because they are conducted only once a year. Furthermore, passengers using a commuter pass are not included in the survey results.

(3) Visual count of ridership
The train conductors of ICTs have to report the ridership in predetermined target sections. They walk through the train and generally count the passengers in each seat type. In cases where the target section consists of some station-to-station sections and some passengers get on board or drop off within the target section, the record should be the maximum number of the ridership through the target section. The most important feature of this data is that it is recorded every day, on all trains and for all seat types. Analysis of this set of data, can therefore reveal changes in seasonal, daily or weekly demand, etc. The drawback with this set of data however is that it is relatively less accurate.

<table>
<thead>
<tr>
<th>Data collection method</th>
<th>Type of data</th>
<th>Frequency</th>
<th>Seat type</th>
<th>Features</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>From the seat reservation system</td>
<td>OD</td>
<td>Everyday</td>
<td>RS and first class seats</td>
<td>OD data of RS is available</td>
<td>Does not always correspond with actual demand</td>
</tr>
<tr>
<td>Collecting used tickets</td>
<td>OD</td>
<td>Mostly Once a year</td>
<td>All types</td>
<td>OD data of NRS is available</td>
<td>Cannot grasp short-term changes in demand</td>
</tr>
<tr>
<td>Visual count by the train conductors</td>
<td>Ridership in target sections</td>
<td>Everyday</td>
<td>All types</td>
<td>Demand data on all days and of all seat types are available</td>
<td>Less accurate than other data</td>
</tr>
</tbody>
</table>

Table 2 Existing real demand data

QR of RTRI, Vol. 58, No. 2, May 2017
than other data and does not indicate OD.

2.2 An online survey

To establish a potential passenger demand estimation model for NRS, an online survey was conducted. In this survey, passengers were asked about what they did if the train they were boarding which had coaches offering NRSs was crowded. The outline of the survey is described in Table 3. Passengers were also asked how they selected their seat type and how much they were influenced by the price difference between RS and NRS [5].

The results are shown in Table 4. Most passengers indicated they would board a coach with NRS, even though only few were available so long as they could get a seat. If no free seat was available (i.e. they have to stand in the carriages or in aisles), 35% said they would avoid boarding.

2.3 The potential passenger demands estimation model

A model was designed to estimate potential passenger demand from the data described above. The outline of the model is shown in Fig. 1. This model is intended to esti-

mate the potential OD data of some ICTs on a particular date from actual demand data for ICTs. Because there are few actual OD data for the NRS, an attempt was made to estimate the potential OD data indirectly: i.e. using this model, ridership was estimated for some sections and converted to OD data.

(1) Calculate the average ridership count reported on the same date over past years, taking into account the sequence of days. This was treated as actual ridership for the section on the target date and train.

(2) From the estimated ridership calculated in (1), potential ridership can be calculated using an equation

<table>
<thead>
<tr>
<th>Table 3 Outline of the online survey</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target audience of the survey</td>
</tr>
<tr>
<td>- High-school student age to 69-year-old adults</td>
</tr>
<tr>
<td>- Live in Japan (excluding Okinawa Pref.)</td>
</tr>
<tr>
<td>- Have experience of getting on ICT (not including Shinkansen super express) operated by JR (Japan Railways) in the last year</td>
</tr>
<tr>
<td>Period of the survey</td>
</tr>
<tr>
<td>From Jan. 30th to Feb. 13th, 2014</td>
</tr>
<tr>
<td>Number of respondents</td>
</tr>
<tr>
<td>Target: 8130 Valid responses: 6036 (response rate: 74.2%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 4 Answers for “What do you do when you want to board a car with NRS and are faced with passenger congestion?”</th>
</tr>
</thead>
<tbody>
<tr>
<td>Do not mind Mind a little but do not feel uncomfortable Board the ICT although it feels uncomfortable Feels uncomfortable and do not board the NRS car (*)</td>
</tr>
<tr>
<td>(1) 3357 1854 628 197</td>
</tr>
<tr>
<td>(2) 3134 1989 743 170</td>
</tr>
<tr>
<td>(3) 832 714 2388 2102</td>
</tr>
<tr>
<td>(4) 501 350 1989 3196</td>
</tr>
</tbody>
</table>

(*) For example: either book a RS, or wait for next service, etc.

(1) There are some window seats available but the aisle seats are fully occupied (adjacent, front and rear seats surrounding each of the empty seats are occupied)

(2) There are some aisle seats available but the window seats are fully occupied (ditto)

(3) Cannot be seated due to full occupancy (but no one is standing)

(4) Cannot be seated due to full occupancy and the deck is crowded
describing the relationship between potential demand and actual demand.

(3) Convert the potential ridership calculated in (2) into potential OD data for the target date and train, by applying the results of the OD survey for the target ICT as OD pattern data.

In detail each method works as follows.

2.3.1 Relationship between the potential and actual demand (for NRS)

An attempt was made to formulate a relationship between actual ridership (actual demand) and the number of passengers who avoided boarding ICTs due to passenger congestion (lost demand), by utilizing the result of the online survey described in 2.2. Firstly, it was assumed that the LF was 75% if either window or aisle were still available. This is because usually half the seats on ICTs are windows seats and the other half are aisle seats. Then, it was assumed that when the LF was between 75% and 100%, the behavior of passengers intending to get on was represented by the average of the second and third rows in Table 4. Under this assumption, the ratio between passengers corresponding to the actual demand (i.e. who take the ICT) and passengers corresponding to the lost demand (i.e. who decide not to board the ICT) can be calculated as below:

\[
\frac{2102}{3934} = \frac{11705}{367}
\]

Hence, when the LF is between 75% and 100%, the lost demand is estimated to be 367/11705 of the actual demand which exceeds the amount corresponding to a LF of 75%.

In the same way, when the LF is over 100%, the ratio is:

\[
\frac{832}{2102} = \frac{3934}{2102}
\]

Therefore, when the LF is over 100%, lost demand can be estimated from the sum of 2102/3934 for actual demand exceeding that corresponding to a LF of 100% and the lost demand generated when the LF is between 75% and 100%.

This estimation makes it possible to formulate a relationship between the potential LF and actual LF of NRS. The result is shown in Fig. 2. Note that it is considered that lost demand is not generated when the LF is less than 75%.

2.3.2 Relationship between potential and actual demand (for RS)

Estimating potential demand for RS is more difficult than for NRS, for the following reasons:

(1) The LF for RS, i.e. the number of seats reserved divided by the number of RSs, never exceeds 100%. For this reason, it is not possible to conclude with absolute certainty that “an ICT with a high actual LF has high potential demand.”

(2) Sometimes a seat reservation request is rejected despite the LF being less than 100%. This is because the request is accepted only when there is a seat available for the entire OD route. Therefore, it is not possible to know whether any seat reservation requests were rejected or not, from the LF.

(3) Most seat reservation request rejections are not recorded in the seat reservation system due to the system’s specifications. In addition, the records of the rejections do not necessarily represent the passenger’s initial choice. For these reasons, it is not possible to observe how many seat reservation request rejections actually occurred through the seat reservation system.

Some thought was given to designing a method for estimating the potential demand of the RS from actual demand data in consideration of these problems. First, while there is the problem (1) shown above, it was decided to formulate the relationship between the potential and actual LF of the RS, which is similar to that of the NRS, because there was no other data from which to estimate the volume of potential demand. Next, it was assumed that seat reservation rejection occurred only when the LF was over 85%, based on the results of interviews with railway operators. Furthermore, sales results for ICTs recorded in the seat reservation system were analyzed and it was found that most of the RSs were reserved on the day travel and reservation rates on the days running up to the service rarely exceeded 85%. This insight indicated that the following assumptions could be applied:

(a) No reservation requests were rejected when the actual LF of the RS was under 85%. When the LF was over 85%, rejections occurred at a constant rate as the LF exceeded 85%. In other words, it was possible to satisfy the equations described below, where \( \nu \) is the potential LF and \( x \) is the actual LF:

\[
y = x \quad (\text{if } x \leq 0.85) \quad (3)
\]

\[
y = x + p(x - 0.85) \quad (\text{if } x > 0.85; p \text{ is a coefficient}) \quad (4)
\]

(b) Reservation rejection occurs only on the day of service.

(c) There is a linear relationship between the rate of reservation on the previous day of its service and the potential LF.

With these assumptions, a linear regression analysis was conducted between the rate of reservation on the previous day of service and the actual LF (from assumption (a), this is equal to the potential LF) for every ICT with a LF under 85%. Coefficient \( p \) was then calculated to make it fit the regression line for ICTs with a LF of over 85%. Using the obtained value of \( p \), a relational expression between the potential LF and actual LF has been calculated.
2.3.3 Method for obtaining OD data from ridership count in each section

As described in 2.1, there is little actual measured OD data for NRS. As such, a method was designed to find potential OD data by adjusting past real OD data to correspond with estimated potential ridership, as follows.

(1) Past OD survey results (or results from another past survey, if the target date was a holiday) for target ICTs are taken as the basic OD data. The number of passengers taking the ICT from station $i$ to station $j$ in this basic OD data is described as $t_{ij}$.

(2) For every station found in the basic OD data, a coefficient $a_i$ is defined. It is assumed that the number of passengers from station $i$ to station $j$ in the OD data is calculated by $t_{ij} \times a_i$ (show Fig. 3.)

(3) The OD data assumed in (2) is converted into the ridership for every section. Then the coefficients $a_i$ are calculated, minimizing the square error between the converted ridership and the potential ridership. Potential OD data is calculated using the values found for $a_i$. For example, when the OD data assumed in (2) correspond to the data shown in the table in Fig. 3 below, and the potential ridership in the section between station 3 and station 4 is 200, under this method, $a_i$ values are calculated so that the sum of 100$a_1$, 70$a_2$ and 60$a_3$ approximates 200 as much as possible.

In other words, this method is built on the assumption that the distribution of passengers who get on at a station among the stations at which they get off is constant regardless of the day of service. Applying this method based on this assumption, the potential OD data can be calculated by adjusting the basic OD data to suit to the estimated potential ridership which is input into the method.

3. Establishing second-best choice behavior model

To utilize the potential demand to determine how the transport service should be, it is necessary to consider how customers think when they meet with congestion; in other words, understand how potential demand changes in relation to actual demand. To measure this, a questionnaire survey was conducted among some ICT passengers. The results gave insight into their choice preference when they could not get a seat / could not get a reservation. For RS passengers, the question was “How would you have behaved if you had not been able to buy a RS ticket due to full occupancy?” For NRS passengers, the question was “How would you have behaved if you had met with congestion in the car with NRS when boarding?” The outline of the survey is described in Table 5.

The result is shown in Table 6. About half the RS passengers would have tried to buy a RS on another train. About 40% would have decided to board the car with NRS. Only 4% answered that they would have avoided boarding the ICT.

63% of NRS passengers would have chosen to board the car with NRS, despite the prospect of having to stand during the journey. In other words, only 37% of them would have avoided boarding the NRS. This is corresponds with the result of the online survey described in 2.2. 25% of passengers would wait for the next train: 8% would try to buy a RS on another train. The result is shown in Table 6. About half the RS passengers would have tried to buy a RS on another train. About 40% would have decided to board the car with NRS. Only 4% answered that they would have avoided boarding the ICT.

Chi-squared tests were performed on the results, to check whether the attributions of passengers and trains significantly affected preferred choice or not. The results of these tests imply that choice preference varies significantly with certain factors, namely age and number of accompanying persons for RS passengers, and age and OD for NRS.

### Table 5 Outline of the onboard survey

<table>
<thead>
<tr>
<th>Target audience of the survey</th>
<th>Those who:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Period of the survey</td>
<td>Jan. 17th and 18th, 2015</td>
</tr>
<tr>
<td>Target ICT</td>
<td>10 trains on each day</td>
</tr>
<tr>
<td>Number of respondents</td>
<td>Distributed: 4194</td>
</tr>
<tr>
<td></td>
<td>Valid responses: 3905 (response rate: 93.1%)</td>
</tr>
</tbody>
</table>

Fig. 3 An example of applying the conversion method
Table 6   Passenger behavior had they been unable to have the seat they were occupying

<table>
<thead>
<tr>
<th></th>
<th>Passengers of the RS</th>
<th>Passengers of the NRS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Book a RS on an earlier service</td>
<td>261</td>
<td>-</td>
</tr>
<tr>
<td>Book a RS on a latter service</td>
<td>291</td>
<td>209</td>
</tr>
<tr>
<td>Board a car with NRS on an earlier service</td>
<td>6</td>
<td>-</td>
</tr>
<tr>
<td>Board a car with NRS on this train</td>
<td>447</td>
<td>1748</td>
</tr>
<tr>
<td>Board a car with NRS on a latter service</td>
<td>15</td>
<td>470</td>
</tr>
<tr>
<td>Abandon the journey on the ICT</td>
<td>41</td>
<td>93</td>
</tr>
<tr>
<td>Others / invalid answers</td>
<td>83</td>
<td>241</td>
</tr>
</tbody>
</table>

From these tests, a decision was made to establish various second-best choice behavior models based on these factors. Figure 4 is an example showing a model for the NRS passengers travelling over a long distance and who are in their 30s. NRS passengers have four choice alternatives as a second-best choice shown in Fig. 4. RS passengers have six choice alternatives; including "switch to a RS on an earlier service," "switch to a NRS on an earlier service."

Fig. 4 An example of a second-best choice behavior model

4. The system developed and the optimization algorithm

A seating plan optimization system was developed using the models described above. The outline of the algorithm is shown in Fig. 5. This system is based on a genetic algorithm, whose solutions are the combinations of seating plans of each target ICT and those in each section on the target date. In this system, to evaluate the solution, the relationship between potential demand and the solution has to be considered, especially the effects on the demand caused by a certain seating plan. Therefore, this system involves a simulation of the selling process of RS before departure and the seating behavior on the NRS when a certain ICT arrives, which is referred to as "S&S simulation" hereafter.

The calculation flow of the system is described below.

1. Estimate the potential OD on the target date and trains from past real demand data by applying the potential passenger demand estimation models described in Chapter 2.
2. Produce various seating plans, as initial solutions of the genetic algorithm.
3. For each of the solutions produced in (2), conduct S&S simulation using the potential OD estimated in (1).
4. Evaluate the solution in consideration of the OD data from S&S simulation derived in procedure (3) and calculate the evaluation indices of each solution as an objective function. Based on the genetic algorithm, the solutions are generated and selected, using a certain
index which is set in advance. The objective function can be chosen among four types of evaluation index; the value of the passenger-kilometer of those who takes an ICT but cannot be seated (minimize), the difference between the LF of the RS and that of the NRS (minimize), the LF (maximize) and the total revenue (maximize).

(5) Iterate (3) and (4) repeatedly, and when the terminal condition defined in advance (e.g. the limit of calculation time) is satisfied, terminate the continuous calculations under this system and output the best solution.

The optimized solution is displayed in the form of a seat map, as shown in Fig. 6.

---

5. Result of model validation

To validate the accuracy of the re-productivity of the models and the S&S simulation, OD data was estimated for a series of ICTs operated over a period of two weeks by using the system and this was compared to real demand data.

The estimated OD data was converted to ridership for each section, and compared to real visual ridership counts. Although there were some outliers, the coefficient of the correlation R was 0.787 for RS and 0.800 for NRS. These values imply that the models and S&S simulation are sufficiently accurate for re-productivity and could be suitable for practical use.

6. Case study results

To evaluate the effect of this kind of flexible seating plans, the seating plans optimization system was applied to a series of ICTs operating along a given line, as a case study. The results are shown in Table 7. Two types of evaluation index were applied, and calculations were performed for a two-day period (a weekday and a holiday.) The evaluation indices of the current seating plan were also calculated and compared to the indices of the optimized seating plans. The results suggest that optimization appears to make it possible to decrease the passenger-kilometers of passengers boarding an ICT without being able to find a seat and to increase revenues.

---

Table 7  The effect of optimizing seating plans

<table>
<thead>
<tr>
<th>Date</th>
<th>Minimize P-K</th>
<th>Maximize total revenue</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>P-K</td>
<td>Total revenue</td>
</tr>
<tr>
<td>A weekday</td>
<td>-74.1%</td>
<td>+0.1%</td>
</tr>
<tr>
<td>A holiday</td>
<td>-60.1%</td>
<td>+0.1%</td>
</tr>
</tbody>
</table>

* P-K means “passenger-kilometers of passengers boarding an ICT and unable to find a seat.”
** The percentages in Table 7 show the results of the comparison with the evaluation indices of the current seating plan.

7. Conclusion

Models were developed to estimate potential passenger demand and second-best choice behavior, and a seating plan optimization system was designed for ICTs. This system helps to generate solutions which improve passenger experience and increase railway operator revenue.

The system is an effective tool for railway operators to review their ICT seating plans. While further studies are needed to adapt the system to practical train operation, e.g. adapting the system so that passengers can see whether a seat is RS or NRS, it is hoped that this effective flexible seating plan system for ICT services can be improved in cooperation with ICT operators.

---

References


Authors

**Shingo NAKAGAWA**
Assistant Senior Researcher, Transport Planning and Marketing Laboratory, Signalling and Transport Information Technology Division
Research Areas: Transport Behavior Analysis, Revenue Management

**Munenori SHIBATA**, Dr. Eng.
Senior Researcher, Transport Planning and Marketing Laboratory, Signalling and Transport Information Technology Division (Former)
Research Areas: Transport Planning, Travel Behavior Modeling, Revenue Management

**Noriko FUKASAWA**
Senior Chief Researcher, Laboratory Head, Transport Planning and Marketing Laboratory, Signalling and Transport Information Technology Division
Research Areas: Transport Behavior Analysis, Travel Data Utilization