Departure Time Decision of Railway Users under Travel Time Uncertainty

Kazuyuki TAKADA
Associate Professor
Graduate School of Science and Engineering
Tokyo Denki University
Ishizaka, Hatoyama-machi, Saitama
350-0394 Japan
Fax: +81-49-296-2911
E-mail: takada@g.dendai.ac.jp

Makoto FUJIU
Ph.D. Candidate
Graduate School of Interdisciplinary Information Studies
The University of Tokyo
7-3-1, Hongo, Bunkyo-ku, Tokyo
113-8656 Japan
Fax: +81-3-5458-6438
E-mail: fujiu@iis.u-tokyo.ac.jp

Shigeki YOKOYAMA
Consulting Engineer
Creative Research and Planning Co., Ltd.
1-5-4, Higashiyama, Meguro-ku, Tokyo
153-0043 Japan
Fax: +81-3-5776-0012
E-mail: s_yokoyama@crp.co.jp

Takanori SUZUKI
Graduate Student
Graduate School of Science and Engineering
Tokyo Denki University
Ishizaka, Hatoyama-machi, Saitama
350-0394 Japan
Fax: +81-49-296-2911
E-mail: 10rmk19@ms.dendai.ac.jp

Abstract: Recently evaluation of travel time reliability has become a great concern in transportation planning. Many studies have analyzed the travel time reliability so far and most of them focused on the reliability concerning road transports. Therefore, the reliability of railway service was analyzed in this study. It is thought that railway users recognize variability of travel time and decide their departure time in consideration of necessary spare time. Therefore, departure time decision of railway users in Tokyo metropolitan area was investigated in this study. Internet survey was conducted to prepare the data for analysis. Several factors which affect the length of spare time were extracted through the basic analysis. Regression model which can estimate the length of spare time was built by Markov Chain Monte Carlo simulation. Moreover, it is shown that consideration of user’s heterogeneity concerning distribution of the spare time is effective for the analysis.

Key Words: travel time reliability, departure time decision, MCMC simulation

1. INTRODUCTION

There are many incidents which cause delay of railway service. The delay has forced a large inconvenience for railway users. In Japan, Ministry of Land, Infrastructure, Transport and Tourism (MLIT) maintains the statistics concerning railway accidents and troubles with more than 30 minutes delay. This statistics can present when the accident happened, what the cause was, how large the influence was, and etc… According to this statistics, many railway accidents occur in Tokyo metropolitan area and the schedule of railway service is chronically delayed by several accidents every day. Railway users have received inconvenience such as actual time loss and loss of the opportunity cost. As the result, this influences the railway user’s behavior. Therefore, to clarify the travel behavior under travel time uncertainty is useful to consider railway planning.

Previous studies concerning travel time reliability mainly focused on road traffic and a few studies on railway transport. Therefore, this study investigates departure time decision of railway users under travel time uncertainty.
Over the past few decades, many studies analyzed the departure time choice. Most of them focused on road traffic. For instance, Gaver (1968) and Vickrey (1969) considered disutility of early arrival and Small (1982) considered disutility of delay arrival. Furthermore, De Palma et al. (1997), Senbil (2004), Park et al. (2005), Chen et al. (2010), Ramli et al. (2010) also analyzed departure time decision.

Meanwhile, there are many studies which examined travel time reliability and most of them focused on road traffic. Reitveld et al. (2001) defined uncertainty of arrival time as the reliability. The probability distribution of arrival time and departure time were specified for each transport mode. Bates (2001) arranged a concept of travel time reliability. Furthermore, Eliasson (2004), Karlstrom (2009) and Fosgerau (2010) analyzed travel time of single link of road traffic. Fosgerau (2009) and Fosgerau (2010) explained two approaches to analyze the travel time reliability. First approach is mean-variance approach and the other is scheduling approach. Susilawati et al. (2010) analyzed travel time reliability of ten corridors in the Adelaide Metropolitan road network. As the result, there were large differences in spare time among the corridors.

Previous studies concerning departure time choice and travel time reliability have been reviewed. However, there are no researches which estimated departure time choice model considering travel time reliability of each railway users under multiple OD link. Furthermore, factors of departure time choice considering travel time reliability are not still detected.

In this study, a utility function regarding departure time decision was considered by applying Fosgerau’s approach. First, departure time decision problem under travel time uncertainty was formulated as a minimization problem of expected negative utility. Then, departure time decision model considering probability distribution of the spare time was estimated. The MCMC simulation was utilized to estimate the model parameters. Moreover, heterogeneity of individuals towards distribution of the spare time was considered and the factors influencing on spare time were identified.

The contents of this study are as follows. In the next chapter, methodology to analyze the departure time decision under travel time uncertainty is explained. In chapter 3, general information of the data is explained, and estimation result of the departure time decision model is described. In chapter 4, outcome of this study is described.

### 2. DEPARTURE TIME DECISION UNDER TRAVEL TIME UNCERTAINTY

Before explaining the formulation proposed in this study, Fosgerau’s approach which can integrate the mean-variance approach and the scheduling approach is described shortly. Fosgerau (2009) and Fosgerau (2010) assumed a cost function as written in equation (1).

\[
U(D, T) = \alpha D + \omega T + \beta(T - D)^+ 
\]  

(1)

where \(D\) is departure time, \(T\) is travel time, and \((T - D)^+\) is schedule delay late. The first term is the cost of early starting which is opportunity cost of interrupting a prior activity. The second term is the cost of travel time \(T\). And, the third term is the cost of being late.

Meanwhile, travel time \(T\) is composed of two terms as seen in equation (2).
\[ T = \mu + \sigma X, \quad X \sim \Phi(0,1) \]  
\[ \text{where } \mu \text{ is travel time under free flow condition, } \sigma \text{ is a fluctuation of travel time, and } X \text{ is a standardized random variable. He assumed that traveler chooses his/her departure time so as to minimize an expected cost under uncertainty of travel time. A formulation of this mathematical problem is formulated as written in equation (3).} \]

\[ \text{E}(U(D)) = \min \left[ \alpha D + \omega \mu + \beta \int_{\frac{D}{\beta}}^{\infty} (\mu - \sigma x - D) \Phi(x) dx \right] \]  
\[ \text{The solution of this problem is written in equation (4).} \]

\[ D^* = \mu + \sigma \Phi^{-1} \left( 1 - \frac{\alpha}{\beta} \right) \]  
\[ \text{where } \frac{\alpha}{\beta} \text{ is defined optimal probability of being late by Bates(2001). By inserting the optimal head start } D^* \text{ into the cost function, optimal minimum expected cost is derived as seen in (5).} \]

\[ \text{E}(U(D^*)) = (\alpha + \omega) \mu + \beta \sigma \int_{1-\frac{\alpha}{\beta}}^{\infty} \Phi^{-1} (x) dx = (\alpha + \omega) \mu + \beta H\left( \frac{a}{b}, \Phi \right) \sigma \]  
\[ \text{(6)} \]
\[ (\alpha + \omega) \text{, the coefficient of } \mu, \text{ is the marginal cost of travel time, i.e. the value of time. Meanwhile, } \beta H\left( \frac{a}{b}, \Phi \right), \text{ the coefficient of } \sigma, \text{ is the marginal cost of fluctuation of travel time, i.e. the value of travel time reliability.} \]

We applied the Fosgerau’s approach to examine departure time decision of railway users who travel different Origin-Destination pairs. Disutility is defined as following equation in this study. In this study, travel time \( T \) is assumed to be composed of two terms as seen in equation (6).

\[ T = \mu + T_s \]  
\[ \text{First term, } \mu, \text{ is a minimum travel time based on timetable and second term, } T_s \text{, is spare time for not being late. The spare time is set by each traveler considering the occurrence probability of the delay of railway operation.} \]

\[ \text{Meanwhile, } T_s \text{ is a random variable whose probability density function is } f(T_s) \text{ and distribution function } F(T_s). \]

\[ \text{Figure-1 shows the relation between travel distance and length of spare time. Each color represents a category of travel distance. As shown in the figure, spare time of each category seems to follow exponential distribution. Therefore, in this study, it is assumed that the spare time follows exponential distribution. Moreover, it is thought that each traveler has different distribution. Therefore, exponential distribution with covariates is applied.} \]
Figure 1 Relation between travel distance and length of spare time

To estimate the length of spare time, survival analysis is applied. The survival analysis model is a model by which time until a certain event occurs or ends is made the object of the analysis at the time of a certain basis. In this research, spare time for each respondent was considered as explained variables of survival analysis. In this research, the analysis was done by applying proportional hazard model with covariate which increased or decreased its hazard.

The proportional hazard function is applied to the function of the distribution. The hazard function $h(t|x_i)$ is written in the following equation (7).

$$h(t|x_i) = \lambda \exp(\sum_{i=1}^{n} \theta_i x_i)$$  \hspace{1cm} (7)

where, $x_i$ is covariates, $\theta_i$ is parameter, and $\lambda$ is scale parameter.

Relationship between hazard function, $h(t|x_i)$, and survival function, $S(t)$, is shown in equation (8).

$$S(t) = \exp \left( - \int_{0}^{T} h(u) \, du \right)$$  \hspace{1cm} (8)

Hazard function is generally written in equation (9).

$$h(t|x_i)=\frac{f(t|x_i)}{S(t|x_i)}$$  \hspace{1cm} (9)

Then, probability density function with covariates derived as written in (10).

$$f(T_x) = \lambda \exp(\sum_{i=1}^{n} \theta_i x_i) \cdot \exp((-\lambda T_{x}) \exp(\sum_{i=1}^{n} \theta_i x_i))$$  \hspace{1cm} (10)

Thus, expected disutility, $E[U(D)]$, is derived and the mathematical problem of minimizing
the expected disutility is shown in the following equation (11).

\[ E[U(D)] = \min_{\alpha D + \omega u + \beta F_{D-\mu}(\mu - T_s - D)f(T_s)dT_s} \]  

Solution of this problem is the optimal departure time, \( D^* \), and it is derived as follows.

\[ D^* = \mu + F^{-1} \left( 1 - \frac{\alpha}{\beta} \right) \]  

Then, equation (12) can be rewritten in equation (13).

\[ D^* = \mu - \left( \frac{\ln(\frac{1}{1 - \frac{\alpha}{\beta} + 1})}{1 - \exp(\sum_{i=1}^{n} \ln q_i)} \right) \]  

It is assumed that \( \frac{\alpha}{\beta} \) is recognized optimum probability of being late.

3. ANALYSIS OF DEPARTURE TIME DECISION

3.1 Data
Nowadays, the internet survey has become a major survey method. Therefore, this study conducted an internet monitoring survey to collect data concerning travel behavior of railway users. The survey was consigned to the MACROMILL, Inc. Table 1 shows outline of this survey. The survey was conducted in July 11-12, 2009 and we obtained 1030 samples (515 samples a day). The number of available data for the analysis was 863.

Answerer attributes of this internet survey are as follows: First one is jobholder over 15-year-old (including part-time jobs); second one is person who lives in around Tokyo metropolitan area (Tokyo, Kanagawa, Chiba, and Saitama) and third one is person who uses train service daily. Question contents of questionnaire survey are as follows: railway usage behaviors (frequency of taking on trains, perceived travel time, delay times and lost times, etc…), personal attributes (sex age residential area, etc...), consumer satisfaction for railway service, and WTP for reducing perceived lost time. In addition to this internet survey, travel time, number of transits and distance of the OD pairs for each respondent were made using the timetable published by JTB Co., Nov. 2009.

3.2 Determining factors of spare time
This section briefly discusses the characteristics of collected data using figure 2 to 4.

Firstly, definition of three kinds of time is explained. “Travel time” is a travel time measured by timetable. “Perceived travel time” is travel time estimated by railway users considering the situation of daily schedule delay. “Spare time” is measured by the calculation that pulls "travel time" from "perceived travel time".

Figure 2 shows the relation between travel distance and average of spare time. It shows that travel distance and spare time are in proportion. It indicates that the railway users recognize the uncertainty of travel time and that they set their spare time in accordance with travel time / travel distance.
Table 1 Outline of survey

<table>
<thead>
<tr>
<th>Method</th>
<th>Internet survey</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conditions of respondents</td>
<td>- having income</td>
</tr>
<tr>
<td></td>
<td>- being over 15-year-old</td>
</tr>
<tr>
<td></td>
<td>- living in Tokyo / Kanagawa / Chiba / Saitama prefectures</td>
</tr>
<tr>
<td></td>
<td>- using railway service every day</td>
</tr>
<tr>
<td>Travel information</td>
<td>- Frequency of taking on trains in a week</td>
</tr>
<tr>
<td></td>
<td>- Travel route (Origin and Destination, transit point)</td>
</tr>
<tr>
<td></td>
<td>- Perceived travel time</td>
</tr>
<tr>
<td></td>
<td>- Lost time caused by service delay</td>
</tr>
<tr>
<td>Experience</td>
<td>- Experience of encountering the service delay</td>
</tr>
<tr>
<td>Willingness to pay (WTP)</td>
<td>- WTP for reducing perceived lost time</td>
</tr>
<tr>
<td>Satisfaction level</td>
<td>- Fare</td>
</tr>
<tr>
<td>for railway service</td>
<td>- Operation speed</td>
</tr>
<tr>
<td></td>
<td>- Equipment function for safety</td>
</tr>
<tr>
<td></td>
<td>- Operation time</td>
</tr>
<tr>
<td></td>
<td>- Congestion rate</td>
</tr>
<tr>
<td></td>
<td>- Easiness of transfer</td>
</tr>
<tr>
<td>Personal attributes</td>
<td>- Sex</td>
</tr>
<tr>
<td></td>
<td>- Age</td>
</tr>
<tr>
<td></td>
<td>- Marital status</td>
</tr>
<tr>
<td></td>
<td>- Presence of children</td>
</tr>
<tr>
<td></td>
<td>- Residential area</td>
</tr>
<tr>
<td></td>
<td>- Occupation</td>
</tr>
<tr>
<td></td>
<td>- Annual income level</td>
</tr>
</tbody>
</table>

Figure 3 shows the relation between number of transits and average of spare time. It shows that number of transits and spare time are in proportion. It indicates that the railway users set their spare time according to the number of transits.

Figure 4 shows the relation between frequency of taking on trains and average of spare time. It was shown that frequent user is likely to set more spare time.

As the results of analysis, several factors which influence on the length of spare time are specified. These factors are used as the explanatory variables of the departure time decision model in the following section.

### 3.3 Estimation of departure time decision model

Even though the parameters were estimated from the individual data considering covariance, the parameter only estimates the average value, common among all the samples. Same with the parameter $\lambda$ that describes the basic characteristics of spare time, which is also just an average.

This is equivalent to the fact that the distribution of the spare time among individuals should be the same under the same condition. However, it is reasonable that each individual set a different distribution according to his/her travel characteristics, experiences and thoughts even under the same conditions. Therefore in this study, we have estimated distribution parameter of the spare time with two cases, one as a common variable and another considering the heterogeneity of the distribution.
Figure 2 Relation between travel distance and spare time

Figure 3 Relation between number of transits and spare time

Figure 4 Relation between frequency of taking on trains and spare time
In this study, hierarchical Bayesian method is applied and the parameters in equation (13) were estimated by Markov Chain Monte Carlo simulation. There are two parameters in the expression, $\alpha$ and $\beta$, however these are unified to one parameter for the estimation. Respondents’ answer to the question about the expected travel time was considered as optimal departure time, $D^*$. 

Table 2 shows the result of parameter estimation. Model-1 is estimated without considering the heterogeneity concerning the distribution of the spare time. Meanwhile, Model-2 is estimated with considering the heterogeneity. Deviance Information Criterion (DIC) is used to evaluate the explanatory power of the model.

The covariates such as travel distance, number of transits, frequency of taking on trains, and optimum probability of being late are statistically significant as explanatory variables. Meanwhile, plus sign of the parameter indicates that the spare time increases in accordance with the increase of the value of concerned variable. For example, the longer travel distance become, the longer the spare time become. Sign's condition of all variables is reasonable.

The scale parameter, $\lambda$, of the Model-2 was estimated by each traveller to consider the heterogeneity of the characteristics of travellers and their trips. As seen in Figure 5, frequency distribution of $\lambda$ is distributed widely and not normally. This indicates the appropriateness of considering the heterogeneity.

Reproducibility of each model was verified. The results are shown in Figure 6, 7, 8, and 9. Figure 6 and Figure 7 shows the reproducibility concerning departure time and Figure 8 and Figure 9 shows the reproducibility concerning spare time. Coefficients of determination $R^2$, indicate the level of reproducibility. As shown in those figures, the model-2 which considers the heterogeneity can explain the departure time and spare time accurately.

Here, reproducibility of each model concerning departure time and spare time are examined to examine the result of considering heterogeneity. Figure 6 and Figure 8 respectively shows the departure time reproducibility of Model-1 and Model-2. Meanwhile, Figure 8 and Figure 9 also show the spare time reproducibility of Model-1 and Model-2.

Regarding the departure time reproducibility, the coefficients of determination of model-1 and model-2 are 0.9009 and 0.9973. Those values are so high that it indicates that both models have high reproducibility. However, the scatter situations of the plotted point in Figure 7 and Figure 8 are slightly different and obviously reproducibility of model 2 is higher than that of model-1.

Meanwhile, regarding the spare time reproducibility, the coefficients of determination of model-1 and model-2 are 0.0957 and 0.9586 respectively and there is large difference between the values. Moreover, the scatter situations of the plotted point in Figure 9 and Figure 10 are definitely different, and obviously reproducibility of model-2 is higher than that of model-1. The value of the model-2 is so high that it indicates that model-2 has high reproducibility.

As described above, model-2 considering heterogeneity in the scale parameter of spare time distribution has high reproducibility. In other words, the consideration of heterogeneity is effective to explain the departure time decision of railway users.
Table 2 Estimation result

<table>
<thead>
<tr>
<th></th>
<th>Model-1: without considering the traveler's heterogeneity in the distribution of $\lambda$</th>
<th>Model-2: with considering the traveler's heterogeneity in the distribution of $\lambda$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameter</td>
<td>Prior distribution</td>
<td>Parameter (t-value)</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>Parameter of exponential distribution</td>
<td>0.609 (1.400)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha / \beta$</td>
<td>Optimum probability of being late</td>
<td>0.264 (2.361)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta_1$</td>
<td>Logarithm of travel distance(km)</td>
<td>-0.165 (-4.125)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta_2$</td>
<td>Number of transits(times/trip)</td>
<td>-0.202 (-5.941)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta_3$</td>
<td>Frequency of taking on trains(times/week)</td>
<td>-0.039 (-2.438)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DIC</td>
<td>(Deviance Information Criterion)</td>
<td>5663.5</td>
</tr>
<tr>
<td>$\tau$</td>
<td>$\Gamma(\lambda=10, \nu=10)$</td>
<td></td>
</tr>
</tbody>
</table>

Figure 5 Frequency distribution of $\lambda$
4. CONCLUSIONS

Recently evaluation of travel time reliability has become great concern in transportation planning. In this study, the travel time reliability of railway service was investigated in terms of departure time decision of travelers.

In this study, internet survey was conducted to prepare the data for analysis. Several factors which affect the length of spare time were identified through the basic analysis. Moreover regression model which can estimate the departure time and the length of spare time was built. Furthermore, it was demonstrated that consideration of heterogeneity among individuals concerning distribution of the spare time is effective to estimate desirable model.

As the result of verification of reproducibility of the model, it became clear that the estimated model with considering heterogeneity has high performance to explain the departure time and spare time.

In this study, travel time reliability was examined from the view point of spare time. In the
future, we will continue to study on the travel time reliability and evaluate the user benefit caused by decrease travel time variability and reducing spare time.

REFERENCES


