A study on evaluation method of advanced transport information system based on discrete choice analysis

A STUDY ON EVALUATION METHOD OF ADVANCED TRANSPORT INFORMATION SYSTEM BASED ON DISCRETE CHOICE ANALYSIS

Tetsuro Hyodo
Tokyo University of Mercantile Marine
Shigeru Morichi
Kei-ichi Ogawa
Tokyo Institute of Technology

1. INTRODUCTION

In later years, to analyze that how information affects when people decide which transportation is used is particularly needed. Because our society is increasingly information-oriented, the transportation is not exception, we easily use varieties of information by a route guidance system, a parking guidance system, a bus-location system and a timetable. Especially, the recent advanced information systems of transport, such as Advanced Traveler Information System or Parking Guidance Information System, have controlled various kinds of transport behavior, however the methodology which can describe the interaction between the transport information system and individual transport behavior has not been established. The simulation study or experimental study (c.f., Mahmassani et al.(1990), Iida et al.(1994), Fujii et al.(1994)) have been mainly reported, whereas, travel behavioral analysis which describes the changes of individual choice by the information systems has not been fully considered in the simulation systems. If a method of travel behavioral analysis which is analyzed the effect on travel information is defined, using behavioral models to an analysis on several information systems is possible.

This study supposes a new modeling method which can measure the effects of the information system on transport behavior. The feature of this model is the combination of the decision theory under uncertainty and the logit model based on the probabilistic utility theory. The basic formulation of this model is as follows.
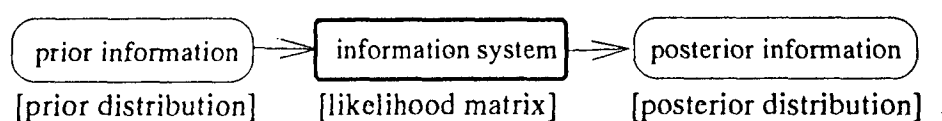

Fig. -1 Concept of information system and individual behavior

The above figure shows that one has prior information of trips before going and receiving travel information changes his/her prior information to posterior information after. We can derive the definition of the value of information system from the above formulation. The purpose of this paper is the following:

1) to model the relation between the accuracy of the information system and the probability of individual discrete choices

2) to measure the value of the information system

3) to estimate the parameters of individual behavior and the information system as a whole system.

2. EVALUATION METHOD OF VALUE OF INFORMATION SYSTEM BASED ON LOGIT MODEL

2.1 General formulation of the evaluation method

First, we define an expected profit of alternative k as $E[x_k]$, and assume that one chooses the alternative which has the maximum expected profit among the set of alternatives. The expected maximum profit before reception of the information is presented as

$$EU(\eta_0) = \max_k (E[x_k]) \qquad . \tag{1}$$

η_0 means 'null-information'. And we introduce likelihood matrix which indicates the joint probability between the prior state (θ_i) and the signal from the information system (η_j) as $f(\theta_i, \eta_j)$. For example, the subjective weather forecast (e.g. fine or rain) before reception of the information (e.g. weather forecast on TV etc.) is one of the prior state and the actual posterior information by the information system (e.g. 'forecasts fine' or 'rain') is regarded as a signal from the system. The marginal probabilities of each state and signal are shown as follows.

$$\sum_i f(\theta_i, \eta_j) = q(\eta_j) \tag{2}$$

$$\sum_j f(\theta_i, \eta_j) = p(\theta_i) \tag{3}$$

The expected maximum profit when a signal η_1 was provided is written as

$$EU(\eta|\eta_1) = \max_k (E[x_k|\eta_1]) \qquad . \tag{4}$$

$$\text{where } E[x_k|\eta_1] = \left(\sum_i f(\theta_i, \eta_1) x_{ik}\right)\Big/ q(\eta_1)$$

x_{ik} : profit of k-th alternative when the state is i.

Considering all set of signals, we can define the following expected maximum profit.

$$EU(\eta) = q(\eta_1)EU(\eta|\eta_1) + q(\eta_2)EU(\eta|\eta_2) + \cdots \tag{5}$$

The value of the information system equals to the difference between the prior and posterior expected maximum profit. Here, we define VI as the value of the information system. And VI can be calculated from the following.

$$VI = EU(\eta) - EU(\eta_0) \tag{6}$$

The value of VI is non-negative, if the profit function is convex function.

2.2 The method of measuring value of information system using Logit model

2.2.1 Basic formulation

If the utility maximization is assumed instead of profit maximization as the decision rule of travel behavior and the expectation of the maximum utility is defined instead of the expectation of the maximum profit, the methodology mentioned in 2.1 can be applied to Logit model (Asakura et al.(1993), Kobayashi et al.(1993)).

For example, we suppose route choice model with two routes: the route-1's average travel time is $\bar{\theta}$ and travel information is given to drivers on route-1, on the other

hand, the travel time of route-2 is fixed to X_2 and there is no information on route-2. The expectation of the maximum utility EU is defined as

$$EU = \log\left(\sum_i \exp[V_i]\right)$$

(7)

where $V_i = \beta x_i$, β :parameter, x_i: travel time of i-th route

Then the value of EU before the reception of the information is represented as

$$EU(\eta_0) = \log\left(\exp[\beta\bar{\theta}] + \exp[\beta X_2]\right)$$

(8)

And after the reception of signal η_j, the following indicates the posterior value of EU.

$$EU(\eta_0|\eta_j) = \log\left(\exp\left[\beta E[\theta|\eta_j]\right] + \exp[\beta X_2]\right)$$

(9)

Consequently, the expectation of posterior maximum utility is derived as follows.

$$EU(\eta) = \int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} f(\theta,\eta_j)d\theta\right) \times EU(\eta|\eta_j)d\,\eta_j$$

(10)

The value of the information system (VI) can be calculated by equation (6), and the value is non-negative because equation (7) is convex function. Moreover, if value of time (VT) is measured, we can evaluate VI in monetary term as follows.

$$VI \times VT / |\beta|$$

(11)

We can define several formulations of likelihood matrix. In this paper, we introduce the following bi-variate normal distribution as the likelihood matrix function to simplify the modeling.

$$f(\theta,\eta) = \frac{1}{2\pi|\Sigma|^{1/2}}\exp\left[-\frac{(\theta-\bar{\theta},\eta-\bar{\eta})\Sigma^{-1}(\theta-\bar{\theta},\eta-\bar{\eta})'}{2}\right]$$

(12)

where, $\Sigma = \begin{pmatrix} \sigma_\theta^2 & \rho\sigma_\theta\sigma_\eta \\ \rho\sigma_\theta\sigma_\eta & \sigma_\eta^2 \end{pmatrix}$, ρ : correlation coefficient between θ & η

In the above equation, θ means subjective prior travel time and η means provided travel time information and the correlation coefficient (ρ) indicates the subjective index of the accuracy of the information system. Attention should be paid to that the prior travel time and the accuracy of the information system are subjective variables, therefore, they differ from observed travel time or correlation. Apparently, we can suppose that the subjective and observed distribution or correlation are identical as a special case. The equation $\rho = 0$ represents 'null-information' as well as $\rho = 1$ represents 'perfect information' as shown in Fig.-2. And the following formulations are derived from bi-variate normal distribution.

$$E[\theta|\eta] = \bar{\theta} + \frac{\sigma_\theta}{\sigma_\eta}\rho(\eta-\bar{\eta})$$

(13)

$$V[\theta|\eta] = \sigma_\theta^2(1-\rho^2)$$

<div align="right">(14)</div>

$$V[\cdot] : \text{variance}$$

The marginal distribution is calculated as

$$\int_{-\infty}^{\infty} f(\theta,\eta)d\theta = \frac{1}{\sqrt{2\pi}\,\sigma_\eta}\exp\left[-\frac{(\eta-\bar{\eta})^2}{2\sigma_\eta^2}\right].$$

<div align="right">(15)</div>

Then equation (10) is rewritten as follows.

$$EU(\eta) = \int_{-\infty}^{\infty}\left(\frac{1}{\sqrt{2\pi}\,\sigma_\eta}\exp\left[-\frac{(\eta-\bar{\eta})^2}{2\sigma_\eta^2}\right]\right)\times\log\left(\exp\left[\beta\left(\bar{\theta}+\frac{\sigma_\theta}{\sigma_\eta}\rho(\eta-\bar{\eta})\right)\right]+\exp[\beta X_2]\right)d\eta$$

<div align="right">(16)</div>

The discussed information system in this·paper is evaluated subjectively by trip makers, so that it seems that the prior and provided travel time have same distribution. From this, the further analyses in this paper assume the following equalities; $\bar{\theta} = \bar{\eta}$, $\sigma_\theta = \sigma_\eta$.
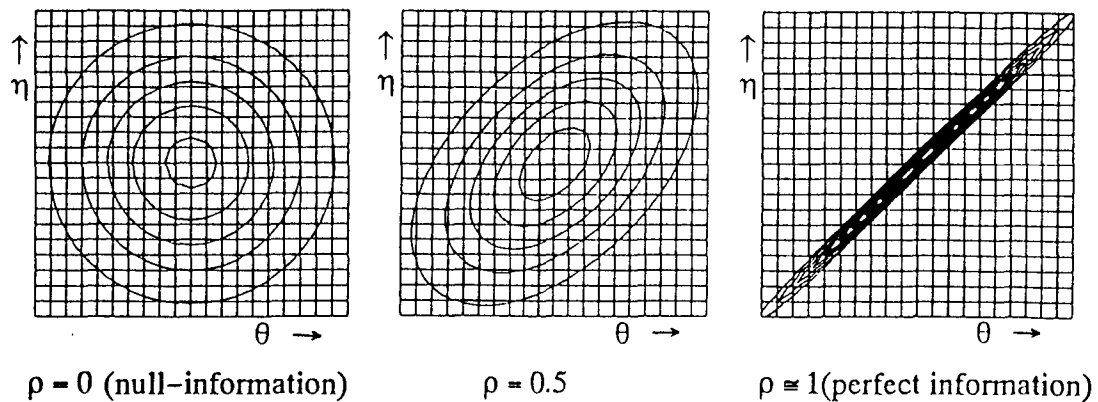

$\rho = 0$ (null-information) $\rho = 0.5$ $\rho \cong 1$(perfect information)

Fig. -2   Figure of likelihood matrix with several values of correlation coefficient

2.2.2   The relationship between VI and other variables

We examine the relationship between the value of the information system (VI) and other variables such as the accuracy of the information system and prior variance (σ_θ^2) and travel time of the another alternative (X_2). The following basic situations are supposed.

● The information of travel time is provided on route-1.

● Travel time of route-1 is normally distributed with mean 50[min.] and variance 20^2 [min.^{2}].

● Travel time of route-2 is fixed to 50[min.].

● Parameter of travel time (β) equals to -0.04.

The value of the information system can be calculated from the difference between

equation (16) and (8). The bi-variate normal distribution mentioned in 2.2.1 is defined as a likelihood matrix of the information system.

The calculation result of the relationship between VI, ρ and σ_θ is shown in Fig.-3. It represents that the accuracy of the information system (ρ) increases the value of the information system (VI). The maximum value of VI under the condition $\sigma_\theta = 20$ is 0.075, which is the case of 'perfect information', and if the value of time is 10 [US$/hour], VI is measured in monetary term as follows.

$$0.075 \times (10/60)/|-0.04| \cong 0.31[\$]$$

(17)

Fig.-3 describes that the variance of prior information (σ_θ^2) is also in proportion to VI. This reason is that the information system is more valuable for one who does not have much prior information. For example, travelers who visit a city for the first time would feel the travel information system as an important tool.

Fig.-4 shows the relationship between VI ,ρ and travel time of route-2 (X_2). The result indicates that the maximum VI is obtained on the points which are $X_2 = 50$ in each ρ. This is because that the value of the information system is maximized when the alternatives have identical value of utility. In other words, when there are large difference between the alternatives' utilities, the value of the information system is recognized to be small, even if the accuracy of the information system is high.
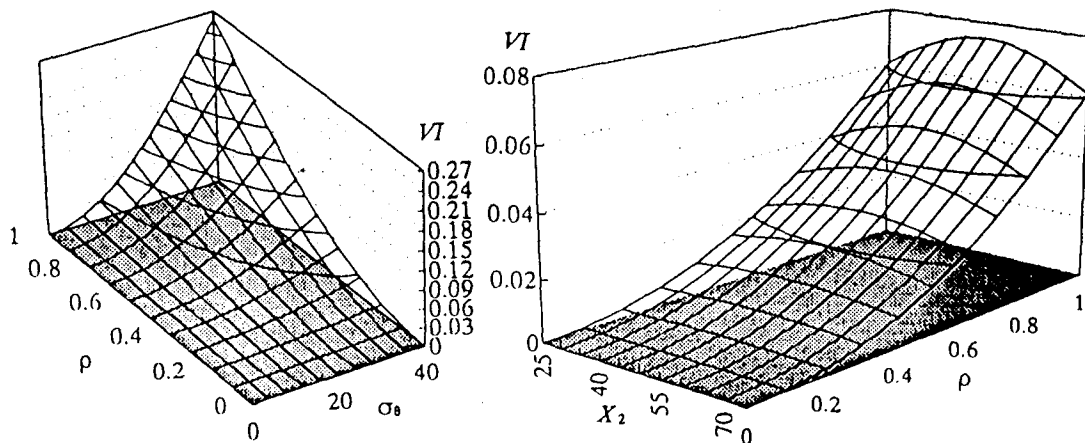

Fig.-3　Relationship between VI, ρ & σ_θ　Fig.-4　Relationship between VI, ρ & X_2

3. CONSIDERATION IN APPLICABILITY OF THE EVALUATION METHOD OF INFORMATION SYSTEM

3.1 Estimation method of information system accuracy

The functions and the parameters which are related to the evaluation method of the information system are summarized in Table-1.

Table-1   Principal functions and parameters of the evaluation method

FUNCTION	PARAMETER
prior distribution $p(\theta)$	average and variance etc.
likelihood matrix $f(\theta,\eta)$	accuracy ρ
utility function	parameter of utility function β

The parameters of utility function are estimated from conventional data of travel behavior such as Revealed Preference data or Stated Preference data. However, the parameters of prior distribution and likelihood matrix could not be derived from the conventional data.   The specific surveys (e.g., in-depth interview) could make possible to identify the parameters of prior distribution, but it is difficult to know the parameter of likelihood matrix.   The likelihood matrix parameter could be estimated by the maximization of log-likelihood function which involves the conventional utility function and the likelihood matrix of the information system.

For example, we assume the bi-variate normal distribution as the likelihood matrix. The perceived travel time before and after the reception of travel information is represented as follows.

- before the reception (prior) : $\overline{\theta}$
- after the reception (posterior) : $\overline{\theta} + (\sigma_\theta / \sigma_\eta)\rho(\eta - \overline{\eta})$

The average and variance is supposed to be identical between the prior and posterior reception of information (c.f., 2.2.1), therefore, $\sigma_\theta = \sigma_\eta$ and $\overline{\theta} = \overline{\eta}$ can be substituted to the above equation.   Then the posterior perceived travel time is rewritten as

$$(1-\rho)\overline{\theta} + \rho\eta .\tag{18}$$

The expectation of posterior perceived travel time is described as a dividing point between prior travel time and provided travel time with weight parameter ρ. Therefore, the accuracy parameter (ρ) can be estimated together with other unknown parameters (β).

3.2 Application to utility function considering variation of travel time

Variation of travel time is one of the important factors for route or modal choice

behavior. It is considered that the travel information decreases the variation, so that the variable of the variation should be introduced to the utility function for our purpose. Several researches on the formulation of the utility function have been developed (c.f., Hall(1983)), in this paper we deal with the utility function including standard deviation of travel time so as to simplify the model building.

First, we assume that the utility function which is explained by average travel time and difference between the average and actual time. The difference indicates the uncertainty factor derived from the variation of travel time. The additional assumption is that the utility increases when one arrives the destination early, on the other hand, the utility decreases when one arrives late. The weight of the difference would not be same between early arrival and late arrival, therefore, we suppose the utility function which consists of the following parts.

- utility of average time :　　　$-\alpha_0 \bar{x}$
- utility of early arrival :　　　$\alpha_1(\bar{x} - x)$ 　　　　　　　　　　　(19)
- utility of late arrival :　　　$-\alpha_2(x - \bar{x})$ 　　　　$(\alpha_0, \alpha_1, \alpha_2 \geq 0)$

And the density function of travel time $(g(\cdot))$ is introduced, then the expectation of the utility is calculated as follows.

$$E[V] = \int_{-\infty}^{\bar{x}} (\alpha_1(\bar{x} - x) - \alpha_0 x) g(x) dx + \int_{\bar{x}}^{\infty} (-\alpha_2(x - \bar{x}) - \alpha_0 x) g(x) dx \qquad (20)$$

If normal distribution $(N(\bar{x}, \sigma^2))$ is supposed as $g(\cdot)$, the following equation is derived.

$$E[V] = -2\alpha_0 \bar{x} + \frac{\alpha_1 - \alpha_2}{\sqrt{2\pi}} \sigma = \beta_1 \bar{x} + \beta_2 \sigma \qquad (21)$$

Consequently, the utility function can be explained by the average and the standard deviation. Equation (21) also shows that parameter β_2 is negative when $\alpha_1 < \alpha_2$. This case represents that one who has tendency to avoid risk has negative β_2 parameter.

The utility function considering both variation of travel time (equation (21)) and likelihood matrix of information system (equation (13),(14)) is derived as follows.

$$V^- = \beta_1 x^- + \beta_2 \sigma_\theta + h(\cdot) \qquad (22)$$

$$V^+ = \beta_1((1 - \rho)x^- + \rho x') + \beta_2 \sigma_\theta \sqrt{1 - \rho^2} + h(\cdot) \qquad (23)$$

In the above equations, superscript '-' means 'prior', '+' means 'posterior' and 'T' means 'provided information', and $h(\cdot)$ means other term except variables of travel time. Using the above utility functions, we can estimate unknown parameters (i.e., β_1, β_2, ρ) by the maximization of log-likelihood function.

4. EXPERIMENTAL ANALYSIS ON ESTIMATION OF INFORMATION SYSTEM'S ACCURACY USING SP DATA

4.1 Abstract of experiment data

The purpose of this chapter is to test the validity of the proposed model by the interview survey data. The home based survey was conducted in a residential area in Yokohama-city where is 20-25 Km from the Central Business District in Tokyo. When people drive to the C.B.D. from home, they have three routes. The first route is via expressway (route-1) and the second is via national road and expressway (route-2) and the third is via only national road (route-3). The questionnaires of the Stated Preference (SP) were as follows.

● How long do you perceive it takes from your home to the C.B.D. ?
  - answer: average, minimum and maximum time by each route
● If you go to the C.B.D. on business, which route do you choose ?

Then, a random hypothetical information of travel time was informed to each examinee, and the following question was asked again as posterior choice result.

● If you know this information before your departure, which route do you choose ?

The number of the examinees was 185, and personal attributes (c.f., age, income etc.) were also asked.
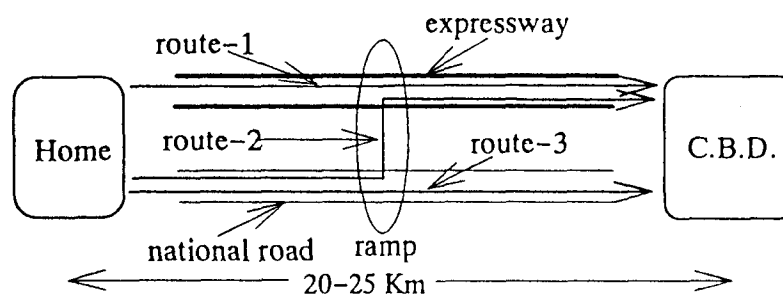

Fig.- 5   The three routes in the interview survey

4.2 The parameter estimation of route choice model and evaluation of information system

At first, we estimate parameters of conventional route choice model by the prior choice result. The model formulation is logit model as follows.

$$P_i = \frac{e^{V_i}}{\sum_{j=1}^{3} e^{V_j}} \tag{24}$$

$$\text{where, } V_i = \sum_k \beta_k x_{ik}$$

The estimation result is shown in Table-2. 'Model-2' includes 'width of travel time' variable, which is difference between the maximum and minimum perceived travel time. And we assume that the width of travel time (Δx) is in proportion to standard deviation with parameter v as follows.

$$\sigma = v \cdot \Delta x \tag{25}$$

Then the utility function of 'model-2' is written as

$$V = \beta_1 \bar{x} + \beta_2 v \Delta x + h(\cdot) \tag{26}$$

β_2 and v can not be estimated separately, so that the estimated parameter of 'width of travel time' in Table-2 means parameter of β'_2 which equals to $\beta_2 v$ ($= \beta'_2$). The estimation result of β'_2 is negative, therefore, it is clear that the examinees have a tendency to avoid risk.

Using the utility function of model-1 and the posterior choice result, we estimate the accuracy parameter (ρ) of the imaginary information system by equation (18). The prior and posterior utility function is described as follows.

-prior : $\quad V^- = \beta_1 x^- + h(\cdot)$

-posterior : $V^+ = \beta_1 \big((1-\rho)x^- + \rho x'\big) + h(\cdot) \tag{27}$

The parameters of both utility functions are estimated by the following joint log-likelihood function.

$$L^-(\beta_1) + L^+(\beta_1, \rho) \rightarrow \max_{\beta_1, \rho} \tag{28}$$

(function $h(\cdot)$ is excluded to simplify the description)

And the utility function based on model-2 is derived as follows from equation (22),(23),(26).

-prior : $\quad V^- = \beta_1 x^- + \beta'_2 \Delta x' + h(\cdot)$

-posterior : $V^+ = \beta_1 \big((1-\rho)x^- + \rho x'\big) + \beta'_2 \Delta x \sqrt{1-\rho^2} + h(\cdot) \tag{29}$

In the same way, unknown parameters (β_1, β'_2, ρ) can be estimated to maximize the joint log-likelihood function.

The results are shown in Table-3. The accuracy parameters of the both models are estimated 0.65-0.66 and the t-values indicate that they are statistically significant. It represents that there is not large difference of estimated ρ between model-1 and model-2 and the accuracy can be measured stably.

Table-2   Estimation result of base model   (t-value in parentheses)

	model-1	model-2
travel time* [generic]	-0.08464 (-6.15)	-0.08086 (-5.67)
width of travel time* [generic]		-0.009728(-1.30)
dummy of income**[route-1]	1.128 (3.25)	1.105 (3.10)
dummy of location***[route-1]	-0.8540 (-2.01)	-0.8943 (-2.11)
constant [route-1]	-1.697 (-4.56)	-1.645 (-4.35)
constant [route-2]	-2.635 (-7.78)	-2.632 (-7.09)
$\bar{\rho}^2$	0.1981	0.2009
hit ratio [%]	67.57	68.65

*: all 'travel time' means 'perceived travel time and the unit is minute
**: more than 10 million yen/year =1, otherwise = 0
***: distance from home to the nearest ramp is more than 6 Km=1, otherwise =0

Table-3   Estimation result of information system's accuracy (t-value in parentheses)

	model-1	model-2
travel time [generic]	-0.06998 (-7.92)	-0.06706 (-7.51)
width of travel time [generic]		-0.01193 (-2.37)
dummy of income [route-1]	0.8018 (3.70)	0.7602 (3.48)
dummy of location [route-1]	-0.4761 (-1.73)	-0.5472 (-1.98)
info. system accuracy (ρ) [generic]	0.6546 (5.47)	0.6627 (6.50)
constant [route-1: prior]	-1.320 (-4.85)	-1.276 (-4.72)
constant [route-2: prior]	-2.490 (-7.56)	-2.502 (-7.69)
constant [route-1: posterior]	-0.8427 (-3.23)	-0.8488 (-3.45)
constant [route-2: posterior]	-1.758 (-6.25)	-1.822 (-6.41)
$\bar{\rho}^2$	0.1589	0.1656
hit ratio [%]	66.95	68.07

The sensitive analysis using parameters of Table-3 is represented in Fig.-6. The value of utility is calculated only by variables of travel time and information accuracy ('travel time', 'width of travel time' and 'information system accuracy'). The figure shows the sensitivity of utility to the information accuracy with each model. The prior travel time and the width of travel time are set to 50 minutes and 40 minutes, respectively, and we calculate three kinds of utility that are informed time is 45 (X^I=45), 50 and 55 minutes. It is obvious that the utility curve of model-1 is linear because of its formulation of utility function. On the other hand, the utility by using model-2 shows non-linear curve. It represents that the accuracy can decrease the uncertainty of travel time and increase the utility of each alternative.

Our study's another interest is to evaluate the difference of perceived accuracy parameter between several types of trip maker. Because it is necessary to clarify market segment which can be more affected by the information system. In this study, the same sets of variables in Table-3 used and the parameters are estimated

for three segments;

1) persons who are using information tools of travel time daily or not,

2) persons who are more than 45 years old or not,

3) persons who took the route more than 3 times or not.

Only the estimated results of ρ in each segment are displayed in Table-4. It suggests that the experience of the travel is related to the evaluation of the information system accuracy, in other words, the persons who do not have much experience depend on the information system relatively.
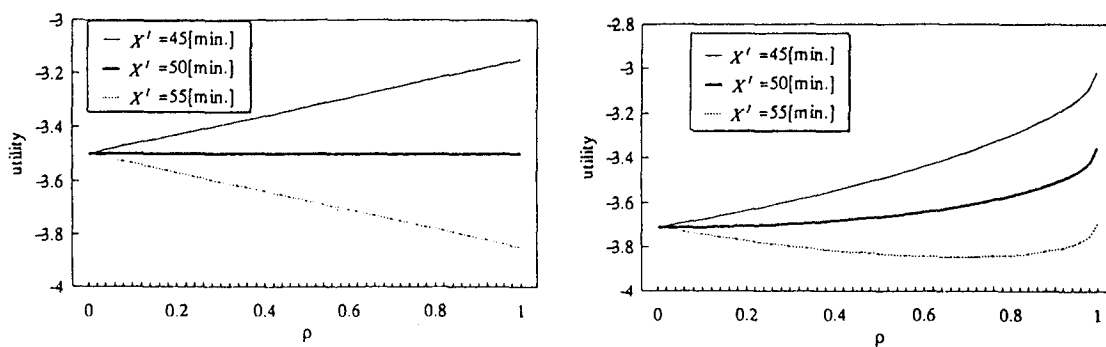

Fig. -6 Sensitivity analysis of utility to accuracy parameter (ρ)

Table-4 Estimation result of ρ in each segment (t-value in parentheses)

	ρ [by model-1]	ρ [by model-2]	# of samples
using travel info. daily*	0.6307 (4.03)	0.6329 (5.69)	146
using travel info. not daily	0.6973 (3.96)	0.7225 (5.41)	39
more than 45 years old	0.6534 (4.25)	0.6866 (5.98)	127
under 45 years old	0.6575 (3.44)	0.6193 (3.97)	58
more than 3 times experiences*	0.5642 (3.64)	0.5963 (5.14)	139
less than 3 times experiences	0.8116 (4.21)	0.7555 (5.38)	46

*: persons who are using travel information daily
**: persons who took the route more than 3 times

5. CONCLUSION

The modeling method for evaluation of transport information systems based on individual behavior was proposed in this paper, and it was represented that the method is able to measure the value of the information system. The experimental study shows that the parameters of the models could be estimated from interview survey. It is considered that the proposed method has general structure, and it can

be applied to various examples. The evaluation of the information system in this paper is limited to the subjective one, therefore, the modeling method which can deal with observed distribution of travel time should be developed in the further analysis. The other topics to be studied are as follows.

1) The results in this paper show that the prior information plays a major role as well as the provided information. It is important to measure the individual prior information and to clarify the decision process of it.

2) It was indicated that the experience of the travel would affect the travel behavior after the reception of the travel information, therefore, the proposed method should be expanded into dynamic analysis.

3) The simulation system based on the method in this paper should be developed. The evaluation of the optimum arrangement or accuracy of the travel information system could be examined by the system.

REFERENCES

Asakura, Y.; Morikawa, T. (1993). Evaluation of parking information systems using behavioral choice models. *Proceedings of 26th ISATA*,   Conference on ATT/IVHS, pp.221-228

Fujii, S.; Iida, Y. & Uchida, T. (1994). Dynamic traffic simulation to evaluate vehicle navigation systems. Vehicle Navigation & Information Systems Conference Proceedings, pp.239-244

Hall, R.W. (1983). Travel outcome and performance: The effect of uncertainty on accessibility. *Transportation Research-B*. Vol.17B, No.4, pp.275-290

Iida, Y.; Uno, N. & Yamada, T. (1994). Experimental analysis approach to analyze dynamic route choice behavior of driver with travel time information.   Vehicle Navigation & Information Systems Conference Proceedings, pp.377-382

Kobayashi, K.; Ikawa, O. (1993). Providing route guidance information for automobile navigation. *Journal of infrastructure planning and management*, JSCE, No.470, pp.185-194 (in Japanese)

Mahmassani, H.; Jayakrishnan, R & Herman, R. (1990). Network traffic flow theory: microscope simulation experiments on supercomputers, *Transportation Research A*, Vol.24A, No.2, pp.149-162