ROUTE CHOICE BEHAVIOR IN NETWORK

by Liu MingWei** and Eiji Hato***

[1]1. Introduction

The precondition of the route choice problem is "human being" that means the choice is acted on the personal recognition and decision-making process. Ben-Akiva, et al (1999) described that there are 3 characteristics in the problem, these are: 1) universal choice set is very large; 2) a driver can't consider all alternatives; 3) alternatives are always relates to each other because of common links. In order to solve the problems mentioned above, this study describes the questions of how awareness of alternative routes affects travelers' choices and how the chosen route can be predicted while considering the traveler's awareness with a three-stage decision making route choice behavior model. In this model, the first stage is network recognition, second is the path enumeration, and the third is the path choice.

For this study, the data is acquired from Matuyama GPS pro-person investigation. The paper examines the route choice behaviors in the recognition network and Matuyama network and moreover a 3-level CNL model is constructed considering the paths' link type and network constitution.

2 Behavior Modeling

The behavior modeling is divided into 3 stage, network recognition, path generation and path choice. The model's structure is shown as Figure 1.

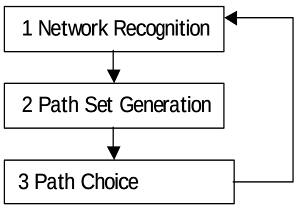
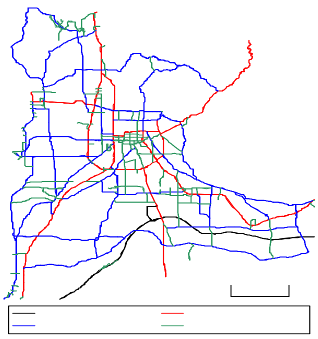


Figure 1 Model structure

(1) Network recognition



by

Figure 2: Matuyama network

The set of all existing routes is called the universal set or full network. The universal set we used is Matuyama network as shown in Figure 1, which is consisted by 1785 links and 746 nodes. For a driver, links in the network are divided into 3 parts: well familiar, not well familiar and unfamiliar. The person's recognition network is made up by the links belong to the first two parts. In this study, a person's travel behavior was observed for 31 days. The information of passed links can be attained from the data. The network made by these reported links is the recognition network. There are two examples, person A and B. A is a worker, man, 37 years old, B is unemployed, woman, 30 years old. A's recognition network is made up by 744 reported links and B's is 242 links.

The goodness of network is embodied by two indexes, matching rate and choosing the shortest path rate. Matching rate is the rate of including choosing route in the enumerated route. Choosing the

＊Keywords: recognition network, discrete choice model, 3-level CNL

**Graduate Student, Department of Civil Engineering, Kyushu University, (Hakozaki 6-10-1,Higashi-ku, Fukuoka, Japan TEL092-642-3275, FAX092-642-3275)

***Member of JSCE, Transport Studies Unit, Ehime University,, (Bunkyo-cho 3, Matuyama 790-8577 Japan, TEL089-927-9676, FAX089-927-9676

shortest path rate is defined that the rate of the chosen one is the shortest one. As shown in Table 1, the comparison results of the two rates rise in the recognition network.

Table 1: Comparing full network and recognition network

	Full Network	Recognition Network
Matching Rate	0.78	0.94
Choosing Shortest Path Rate	0.64	0.78

(2) Path enumeration process

In this process, K-th shortest path algorithm's Screening method is generalizations or repeated applications of shortest path algorithm that generate a collection of paths. The paths with long distance from the origination to the destination were deleted beforehand, and the ones with the length within 2 times of the shortest one made up of the choice set.

(3) Path choice process

The analysis of travel behavior is always discrete, meaning that the models represent the choice behavior of individual travelers. Because of the common link, an important issue in route choice modeling is whether the interactions between the alternatives can be good embodied. In this process, three kinds of models: MNL (can not consider the constitution of the network), CNL (link-based structure), and 3-level CNL (link-based adding considering the link type), are used for estimation.

a) MNL

One limitation of MNL model is the assumption that error terms are independent and identically (i.i.d) Gumbel distribute. This reduces the model can't consider the overlapping problem. The probability of the MNL model is:

$$P(k \mid C_n) = \frac{e^{V_{kn}}}{\sum_{j \in C_n} e^{V_{jn}}} \quad (1)$$

where V_{kn} is path *k*'s impedance for person n.

b) CNL

CNL allows alternatives to belong to more than one nest with different 'degree' of membership. In the context of route choice, CNL employs a link-based nesting structure. The structure of CNL is as Figure 3.

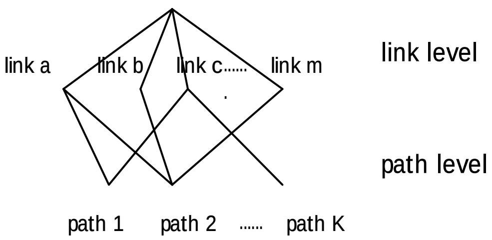


Figure 3 CNL structure

The probability of choosing path *k* is:

$$P(k) = \sum_{m=1}^{M} P(k \mid m) P(m) \quad (2)$$

where $P(m)$, the probability of choosing link *m* is given by:

$$P(m) = \frac{e^{u_1 I_m}}{\sum_{m'} e^{u_1 I_{m'}}} \quad (3)$$

and $P(k|m)$, the probability of choosing path *k* from link *m* is given by:

$$P(k \mid m) = \frac{(a_{mk} e^{V_k})^{1/\mu_1}}{\sum_{j \in C_m} (a_{mj} e^{V_j})^{1/u_1}} = \frac{(a_{mk} e^{V_k})^{1/u_1}}{I_m} \tag{4}$$

where u_1 is log-sum parameter with $0 \le u_1 \le 1$, α_{mk} is the inclusion weight given by:

$$a_{mk} = \frac{l_m}{L_k} \delta_{mk} \tag{5}$$

where l_m is the length of link *m*, L_k is the length of route *k*, δ_{mk} is the link-route incidence dummy, that is, $\delta_{mk} = 1$ when route *k* travels link *m*, and 0 otherwise, and I_m, the inclusive value is given by:

$$I_m = \ln \sum_{j \in C_m} (a_{mj} e^{V_j})^{1/u_1} \tag{6}$$

c) 3-level CNL

There have 4 types of the links in the network: express high way, national road, prefectural road and local street. In this model, the links are divided into 4 parts as mentioned above. The three levels are link type (highest), link (middle) and path (lowest). The structure of 3-level CNL is as Figure 4:

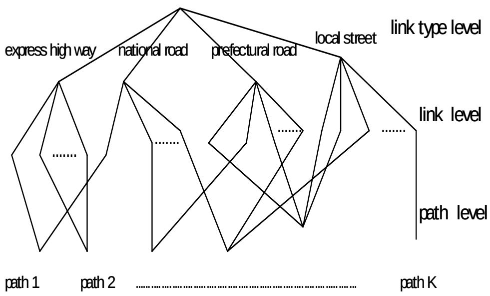


Figure 4: 3-level CNL structure

The probability of choosing path *k* is:

$$p(k) = \sum p(k \mid m, l) * p(m \mid l) * p(l) \tag{7}$$

where $P(k \mid m, l)$, the probability of choosing path *k* in link *m* and type *l* is given by:

$$p(k \mid m, l) = \frac{\exp(V_k)}{\sum_{k' \mid m, l} \exp(V_{k'})} = \frac{\exp(V_k)}{\exp(J_{m|l})} \tag{8}$$

the conditional probability of choosing a particular link *m* in link type *l* is given by:

$$p(m \mid l) = \frac{\exp[u_1 (J_{m|l})]}{\sum_{m'|l} \exp[u_1 (J_{m'|l})]} = \frac{\exp[u_1 (J_{m|l})]}{\exp(I_l)} \tag{9}$$

and the probability of choosing level *l* is given by:

$$p(l) = \frac{\exp[u_2 (I_l)]}{\sum_{l'} \exp[u_2 [I_{l'})]} = \frac{\exp(u_2 (I_l))}{\exp(H)} \tag{10}$$

where u_1 is log-sum parameter with $0 \le u_1 \le 1$, u_2 is the log-sum parameter with $0 \le u_2 \le 1$, $J_{m|l}$ the inclusive value for link *m* in link type *l* is given by:

$$J_{m|l} = \log \sum_{k'|m,l} \exp(V_{k'}) \quad (11)$$

I_l is the inclusive value for type *l* given by:

$$I_l = \log \sum_{m'|l} \exp[u_1(J_{m'|l})] \quad (12)$$

3 Estimation Research

(1) Data

The data used in this study is from "Matuyama GPS prob-person travel behavior investigation", holding during 29/1/2003 to 28/2/2003. Investigated persons living in Matuyama who were older than 18 years and having a car license. 172 samples are used.

(2) Utility function

The utility function is given by:

$$U_k = aDis\tan ce_k + \varepsilon_k \quad (13)$$

where U_k is the utility of the path *k*, a is the parameters to be estimated, $Dis\tan ce_k$ is the length of the path *k*,. ε_k is the unobserved part of the utility *k*. "Distance" should have a negative effect on the choice of the paths because drivers like to choose the shorter one.

(3) Result

From the result of Table 1, the parameters are not only not in the area they should be but also not significant, meaning that there have some biases in the estimation on full network. While from the result of Table 2, except the value of μ_2 in the 3-level CNL, the parameters are all in the area they should be and are significant at 1 percent level. 3-level CNL has the smallest loglikelihood value, meaning that the type of link will be considered when a driver chooses the route. Link-based structure models', which can better reflects the network constitution and overlapping problem, have better result than MNL. The reason why μ_2 is out of the area is that in the function of choosing probability of highest level, link type level, there has no other variables except the inclusive value.

From the result, we can see the result is improved largely using the recognition network comparing to the full network and 3-level CNL has the smallest loglikelihood value.

Table 2 Full network

parameters	MNL	CNL	3-Level CNL
a	0.02(0.36)	-9.72(-0.44)	
u1		-8.64(-0.46)	
u2			
loglikelihood	-364.64	-309.6	

Table 3 Recognition network

parameters	MNL	CNL	3-Level CNL
a	-0.75(-8.52**)	-0.62(-12.224**)	-0.65(-6.82**)
u1		0.07(7.12**)	0.001(0.084)
u2			2.15(4.51**)
loglikelihood	-228.82	-193.32	-190.68

Significance: * at 5%, ** at 1%

4 Conclusions

This paper describes a three-stage decision-making behavior route choice model. The structure of the model is simple and can embody the driver's decision procedure. In the recognition level, recognition network, which is made by recognized links, is made. In the path choice level, a 3-level CNL discrete choice model, which can consider paths' link type and can better describe the network's constitution, is constructed.

For next study, the person's learning procedure of the network will be adapted to the network recognition procedure.

References

1) Moshe E. Ben-Akiwa and Michel Bierlaire "Discrete Choice Methods and Their Applications to Short Term Travel Decisions." In Handbook of Transportation Sciences, Randolph W.Hall, ed, 1999

2) Michael Scott Ramming " Network Knowledge and Route Choice." *MIT Ph.D*, 2001

3) R.D.Shier "On algorithms for Finding k Shortest Paths in a Network." *Networks*, 9,195-214, 1979

4) Hato, E. and Asakura, Y. Incorporating Bounded Rationality Concept into Route Choice Model for Transport Network Analysis. European Transport Conference 2000 in Cambridge, Behavioural Modelling, Vol.441, pp.1-12. 2000.