A suggestion of path enumeration algorithm based on massive behavioral space information

Y. Yamakawa & E. Hato
School of Urban Engineering, The University of Tokyo, Japan

Abstract

One of the difficulties of route choice modeling is to represent a generation process of choice set explicitly in a network including infinite paths. Since it is expensive and uses many man-hours to collect the data, the analysis of routes that a driver actually used has not been done. Therefore, in the existing route choice model, there is a problem that routes included in a choice set do not necessarily correspond to routes that a driver has actually used. However, with the development of Probe Person Technology that gives location data of the movable body, having a GPS phone, it is possible to collect the data of routes a driver has used without burdening the driver. Then, in this paper, the route data which is available on a private basis and obtained in Probe Person survey is analyzed. The choice set generation methods based on routes drivers actually used are suggested, and the case studies of route choice models are carried out by using it.
keywords: route choice modeling, choice set generation, GPS data.

1 Introduction

The route choice model is used for the traffic simulation and is vital to the analysis of road networks. But, the applicability of route choice models for traffic simulation is not sufficient, it is attributed to how to generate route choice set.

When we choose our routes in the road network, we must restrict the road network and generate the choice set in some stages (Bovy [1]). The following four processes are used to generate the choice set, Existing set $\rightarrow$ Known set Available set $\rightarrow$ Feasible set $\rightarrow$ Choice set. Though it is arguable that we choose our routes through these steps, actually, capturing processes to generate the choice set explicitly is essential for the route choice modeling. The Analysis of routes drivers actually used has not been made until now because it is expensive in direct costs and

in man-hours to collect data. In existing route choice models, Labeling Approach (Ben-Akiva *et al.* [3]), EBA Approach (Tversky [4]), and *K*-*th* Bounded Rationality Approach (Hato *et al.*, 1999, Maura *et al.*, 1999) are used to generate the choice set. But all of these generation methods are not based on driver's positional data, so the routes included choice set do not necessarily correspond to routes that the driver has actually used.

But, with the development of Probe Person Technology by which the location data of movable bodies are observed, it is possible to collect driver's routes data without any burden on the driver. It is also possible to compute route characteristics, for example path length, travel-time and so on, and to analyse the routes drivers actually used. In this paper the choice set, which is composed of the routes similar to routes actually used, generation algorithm is suggested by the data-oriented approach. With this choice set, route choice models parameters are estimated.

2 Route choice set generation

In this paper, two type of route choice set generation methods are suggested. Both of these methods use cognitive network. Cognitive network is composed of some (or one) routes that a driver used in consultation periods between an Origin-Destination pair (OD).

In the first method, the choice set is generated with cognitive networks, the path enumeration algorithm applying the shortest path algorithm, and EBA approach. With this method, a problem that *K*-*th* shortest path algorithm is computationally expensive is solved. The advantage of this method is to enumerate spatial paths. In the second method, the choice set is generated with the driver's route data obtained in Probe Person Data, that is cognitive network, directly.

A more realistic choice set based on the driver's route choice behavior is obtained with the second method, but with the second method the choice set can be generated only between OD pairs where a driver is used. On the other hand, with the first method, the choice set can be generated between any OD pairs. Two generation methods are defined in detail in the following way.

2.1 A method with some algorithms

2.1.1 Path enumeration algorithm
In generating the choice set, the alternative path enumeration algorithm is invented. A driver's route choice behaviors are influenced by the cognitive network. Then, we define the virtual cognitive network and enumerate paths with the algorithm using the shortest path algorithm. These are following 4 steps.

Step 1 As the virtual cognitive network, when driving between an OD pair, suppose a diamond shape composed of 4 straight lines. The angles between these lines and the line which connects an Origin and a Destination, is θ. See Step 1 in Figure 1.
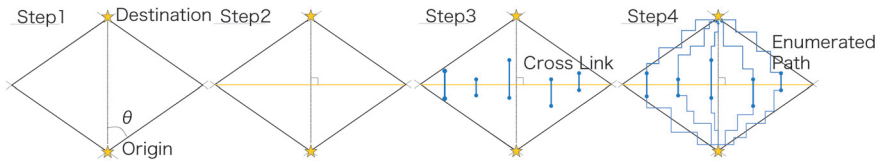
Figure 1: Images of path enumeration.

Step 2 Suppose a part of perpendicular bisector of the straight line which connects an OD pair, in this diamond shape. See Step 2 in Figure 1.

Step 3 Search the links which cross this perpendicular bisector. See Step 3 in Figure 1.

Step 4 With the shortest path algorithm, search the shortest path between Origin, and the up node of the cross path and between the down node of the cross path, and Destinations. Link these two shortest paths, and then the product is an alternative path. See Step 4 in Figure 1.

2.1.2 Choice set generation algorithm with EBA approach

Using alternative paths enumerated by the algorithm, the choice set is generated. Enumerated paths have some overlapping links with each other. Suppose that two (or more) paths are the same paths if their overlap rates exceed a threshold value. Overlap rate of route A to route B is defined as the rate of the sum of overlapping links of route A with route B to path length of route A. Also, behavioral characteristics in route choice vary among individuals such as the number of alternatives, and upper and lower limit of detour rate etc.(detour rate represents ratio of path length of choice route to the shortest path length.). Consequently, the choice set based on these characteristics is generated. The generation method is mainly composed of EBA approach, that paths which do not meet standards are eliminated. These standards are based on path characteristics. There are the following 6 steps.

Step 1 As parameters used in choice set generation algorithm, set upper and lower limit of detour rate, the number of alternatives and overlap rate.

Step 2 As to path length of alternatives, eliminate the paths whose detour rates do not lie between the upper limit and the lower limit.

Step 3 Add the path whose length is the shortest in alternatives to the choice set.

Step 4 Eliminate the path whose overlap rate to the path included in the choice set is higher than the pre-set one in alternatives.

Step 5 Repeat Step 3 and Step 4.

Step 6 If there are no alternatives left, or the number of paths included in choice set reaches the pre-set number of alternatives, add the route the driver actually used to the choice set and finish these processes.

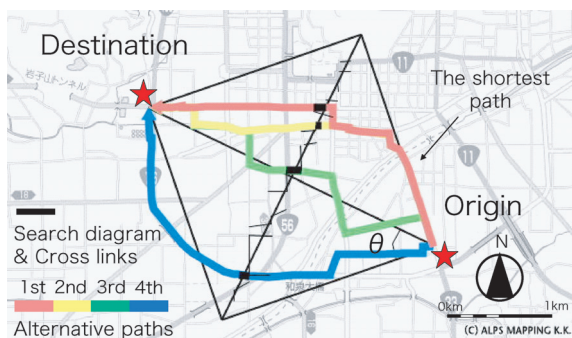Example of choice set generated with algorithms is shown in Figure 2.

Figure 2: Example of choice set with algorithms. (Search angle: 45°, Upperdetouring; 1.50, Lower 1.00, Alternatives: 4, Overlapping: 0.6)
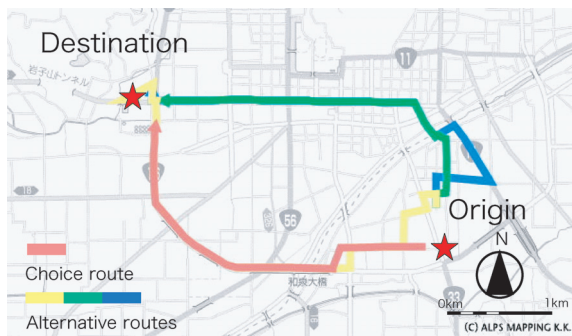

Figure 3: Example of choice set with routes a driver actually used.

2.2 A method with actually driving routes

As to an OD trip, paths which are used by a driver in the consultation period are thought as alternatives. Generate choice set including paths whose overlap rate to each other does not exceed 90%. Example is shown in Figure 3.

3 Route choice model with Probe Person Data

3.1 Data

Data used in this paper are from Matsuyama Probe Person Survey conducted in 2004 (MPP2004). An outline of this survey is presented below. The research zone is the Matsuyama urban area, composed of Matsuyama City, Iyo City, Tohon City, Tobe Town and Masaki Town. Consultation periods are 35 days from 26th January 2004 to 29th February 2004 in a 24-hour period. The number of subjects in this survey is 317 people.

3.2 Analysis of driver's choice routes

3.2.1 Basic characteristic

Set parameters in the path enumeration and the choice set generation algorithm by analyzing drivers' chosen routes. Analyze the detour rate to decide upper and lower limits of the detour rate. Relationships between the shortest path length and the detour rate of the chosen routes are shown in Figure 4. Explanatory note indicates ID of subjects. From this chart, in the traffic simulation, the assumption that drivers choose the shortest path is doubtful. The detour rate runs from 1.00 through 1.40. The longer the distance between Origin and Destination, and the length of the shortest path becomes, the fewer the shortest path is chosen.

Then, relationships between the number of OD pairs used by a driver in consultation periods and the number of alternative routes used between the OD pair are shown in Figure 5. In addition, the choice set is composed of paths whose overlap rate does not exceed 90% of each other. The numbers of OD pair varies among individuals, but most of them are under 4.
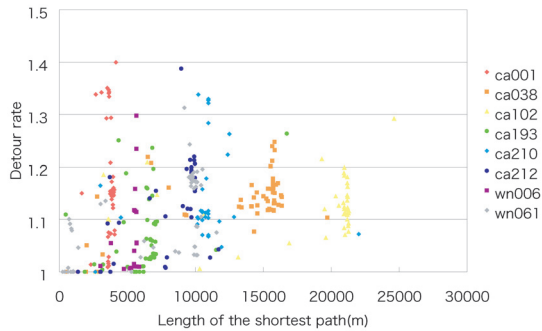

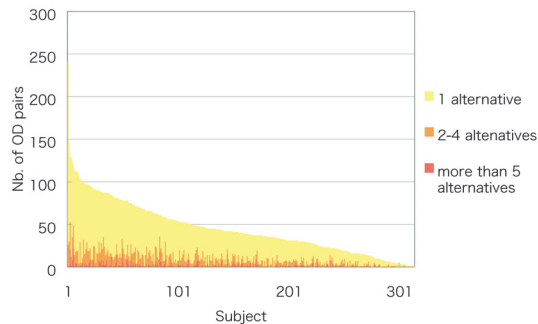Figure 4: Detour rate and the shortest path length.

Figure 5: Alternatives in choice set at an Origin-Destination trip,

3.2.2 Cognitive network

Search the angle parameter θ in the path enumeration algorithm captures spatiality of networks. The virtual cognitive network needs to be similar to the cognitive network which is composed of the actual driver's choice route in order that the choice set is a more realistic one. Here, the cognitive network is composed of the chosen routes and alternatives between an Origin and a Destination. Vary θ from 10° to 70° by 5°. Then, the overlap rate of the virtual cognitive network with the (realistic) cognitive network is defined as the cover ratio. When θ is increases and the cover ratio reaches 90% of maximal values of cover ratio at the OD trip, θ at that time is regarded as the search angle at the OD trip. Relationships between direct distances between OD pairs and search angles are shown in Figure 6. When direct distances between OD pairs are short, the search angle takes on a wide range of values. The longer distances between OD pairs are, the lower search angle becomes.

3.3 Framework of route choice model

In this paper, the MNL (Multinomial Logit) model is used in route choice modeling. The probability that a driver chooses route n is defined as P_n.

$$P_n = \frac{\exp(\mu V_n)}{\sum_{j \in C} \exp(\mu V_n)}$$

V_n is deterministic term of utility, C is choice set, and μ is scale parameter. In addition, suffixes which specifies individuals are omitted.

3.4 Estimation

As explained, variables, the path length, the travel-time, the travel-time variance, the arterial rate, and the number of turning are used for the parameter estimation. These variables can all be calculated using the data collected trough Probe Person technologies. The arterial ratio is the ratio of the sum of the link length which has
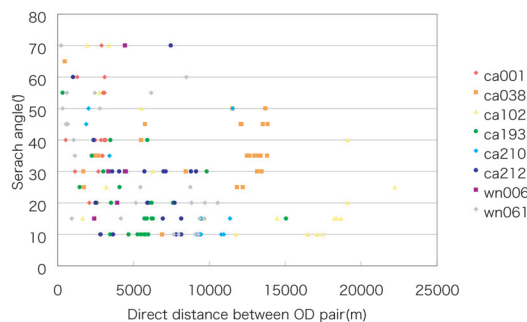

Figure 6: Direct distance between Origin and Destination, and search angle.

more than 2 lanes to the entire the path length. As to travel-time, drivers can not know the travel-time of chosen routes correctly, let alone that of alternatives at the route choice. The travel-time depends on the departure time, and varies from day-to-day, even though a driver departs at the same time each day. Probe Person Data with time and the location information with high accuracy will solve these problems. In this paper, hourly average transit velocities of a link are calculated, and time varying travel-times are calculated with the chart of these velocities. In the following estimation, one subject's data is used for the individual based on the route choice modeling.

3.4.1 Choice set with path enumeration algorithm

In Figure 7 it is shown that if the number of turning is used as an explaining variable, estimates are biased. So, as explaining variables, only the path length is used, and in each 4 models, different parameters are used in the path enumeration algorithm and in the choice set generation algorithm. Estimation results are shown in Table 1. In the model 1, the shortest path is involved in the choice set, since the lower limit of the detour rate is 1.00. Because path lengths of some paths in choice set are shorter than that of the chosen route, estimate of the path length is the positive number. This is against intuition. In model2-4, because the lower limit of the detour rate set 1.10, so the path lengths in the choice set are longer than these of in the model 1. Therefore, estimates of path length is a negative number, and this result correspond to intuition. Paths whose length took on a wide range of values will be enumerated to the choice set, by increasing the search angle, the upper limit of the detour and the overlap rate. So, the appropriateness of fit will be increasing.

3.4.2 Choice set with driver's running route

Some combinations of explaining variables are tested in each estimation. The models which fit well are shown in Table 2. In model5, 6, the travel-time is more explainable than the path length. This shows that the explaining variable made full
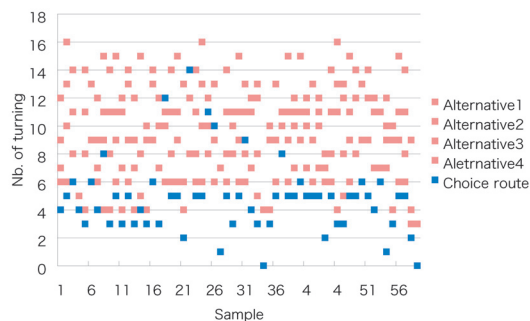

Figure 7: The number of turning in choice set with algorithms.

Table 1: Choice set with algorithms (t-statistics between parentheses).

	model1	model2	model3	model4
	Nb. of observation : 59			
Alternatives	4	4	4	4
Search angle (°)	30	30	40	45
Upper detouring	1.00	1.10	1.10	1.10
Lower detouring	1.30	1.40	1.40	1.50
Overlapping (%)	50	50	50	60
Path length (m)	1.611	−0.725	−0.586	−1.188
	(3.96**)	(−1.19)	(−0.94)	(−2.01*)
ρ^2	0.111	0.012	0.005	0.024
$\bar{\rho}^2$	0.098	−0.004	−0.006	0.013

Table 2: Choice set with driver's running route (t-statistics between parentheses).

	model5	model6	model7	model8	model9	model10
					Nb. of observation : 53	
Path length	−1.593		−1.346			−1.001
(m)	(−3.57**)		(−2.78**)			(−1.43)
Travel-time		−0.732		−0.484		1.667
(min)		(−3.70**)		(−2.19*)		(0.68)
Arterial				−4.291		−0.194
(%)				(−2.37*)		(−1.33)
Turn			−0.235	−0.402	−0.274	0.192
			(−3.57**)	(−3.03**)	(−4.46**)	(0.67)
Time variance					−0.565	−0.597
(min)					(−4.19**)	(−2.82**)
ρ^2	0.126	0.176	0.221	0.239	0.287	0.302
$\bar{\rho}^2$	0.114	0.164	0.197	0.203	0.263	0.242

use of the characteristics of Probe Person Data, such as the travel-time is valid. In model7, 8, it is shown that a combination of variables, such as the path length and the travel-time and other variables makes the model fit high. In all the combination of variables, model 9 which uses the number of turning and travel-time variance as explaining variables shows the highest model fit. Combined with the result of model 10, this implies that a driver prefers smallness of travel-time variance to shortness of travel-time.

4 Conclusion

In this paper, the route choice set generation methods based on a driver's positional information obtained from Probe Person Data are invented, and route choice models are established. Though the goodness of fit of the route choice model using the path enumeration algorithm is low, it is possible that the goodness of fit can be increased by the parameter tuning based on individual behavioral characteristics. Also, in this path generation method, the choosing route is added to the choice set at the final step. Then, it is important to enumerate the choosing route using the path enumeration algorithm for more reasonable choice set generation methods. Arterial rates of the most of enumerated paths are lower and the numbers of turning of these paths are more than that of the chosen route. So, if these explaining variables are used, estimates will be biased. To solve this problem, it is assumed that the threshold value of EBA approach employs these variables, or paths enumerated by labeling approach should be introduced to the choice set. On the other hand, the close match of fit with the driver's chosen routes is comparatively high, and many models with some combination of explaining variables are established. Especially, the fact that the travel-time variance influences the route choice behavior must be taken into account. If individual ultra long-term travel-activity data is available using Probe Person, the choice set will be obtained between more OD pairs, and the route choice model which can applied to traffic simulation will be established.

References

[1] Bovy, S. and Stern, E., *Route Choice : Wayfinding in Transport Networks*, Kluwer Academic Publishers, 1990
[2] Frejinger, E. and Bierlaire, M., Capturing correlation with subnetworks in route choice models. *Transportation Research Part B* 41, pp. 363–378, 2007
[3] Ben-Akiva, M., Bergman, J., Daly, J., Ramaswamy, R., Modeling inter urban route choice behaviour. Proceeding of 9th International Symposium on Transportation and Traffic Theory, pp. 299–330, 1984
[4] Tversky, A., Elimination by aspects: A theory of choice. Psychological Review, Vol. 79, Issue 4, pp. 281–299, 1972