ESTIMATION OF A TIME DEPENDENT OD MATRIX FROM TRAFFIC COUNTS USING DYNAMIC TRAFFIC SIMULATION

TOSHIO YOSHII
Institute of Industrial Science, University of Tokyo
7-22-1 Roppongi, Minato-ku TOKYO 106-8558, JAPAN

MASAO KUWAHARA
Institute of Industrial Science, University of Tokyo
7-22-1 Roppongi, Minato-ku TOKYO 106-8558, JAPAN

Abstract

This study proposes a method of estimating a time dependent OD (Origin-Destination) matrix from traffic counts using dynamic traffic simulation. Then, the accuracy of the estimated OD is examined in relation to the accuracy of simulation parameters, the number of observed counts and additional information.

INTRODUCTION

This study proposes a method of estimating a time dependent OD(Origin-Destination) matrix from traffic counts using dynamic traffic simulation and investigates how accurately the OD matrix can be estimated in relation to the accuracy of simulation parameters, the number of observed counts and additional information.

A time-dependent OD matrix in a general network is required to analyze the optimum traffic control and planning. Many dynamic traffic simulation models have been developed in order to reproduce traffic condition and evaluate policies of traffic control, such as signal control, one-way traffic control and so on. Dynamic models need a time-dependent OD as the input data, especially one composed of small OD zones. On the other hand, as it is hard to directly observe a dynamic OD matrix, some estimating methods from traffic counts have been proposed. Yang *et al.* (1990) estimated OD matrix so as to minimize the integrated squared error between observed and predicted link traffic flows with an efficient solution method developed Fourier transformation. However, this method does not include time-dependent driver's route choice behavior. Ashok and Ben-Akiva (1993) proposed the On-Line estimator using Kalman filtering. This method, however, requires a ratio of traffic flow on each route in advance and a prior OD flows. Cascetta *et al.* (1993) and Oneyama *et al.* (1996) proposed estimation methods with route choice activities using the Maximum Likelihood Method or the Entropy Maximizing Method which are extended from ones for the static OD matrix estimation.

However, using an OD matrix estimated by these conventional methods, simulation can not reproduce observed traffic counts because relationships between OD volumes and traffic counts used in these estimation methods are not the same as those in the simulation. Therefore, we intend to use the relationship observed in the simulation also for the OD estimation method. Namely, we propose an OD estimating method using the dynamic traffic simulation, so that the traffic simulation can reproduce observed traffic counts more accurately. After development of this estimating method, we apply the method to a real network with our dynamic traffic simulation model, SOUND (Kuwahara *et al.*, 1996), and examine its validity.

As a result, the proposed method successfully estimates the OD matrix consistent with the observed counts, although the OD cannot be uniquely determined. We understand the relationship between the reproducibility of estimated OD matrix and affective factors such as simulation parameters, the number of observed counts and additional information.

THE METHOD OF ESTIMATING A TIME-DEPENDENT OD MATRIX

We have already developed a method (Oneyama *et al.*, 1996) to estimate time-dependent OD volumes from traffic counts in a general network with driver's route choice activities. The method consists of two parts; (a) construction of the relationship between the time dependent OD volumes and traffic counts at each link and (b) estimation of a unique time-dependent OD matrix subject to the relationship obtained in (a). In the first part, we define a three-dimensional network to relate OD matrix to traffic flow on links. We then propose a method of estimating route choice probabilities. In the second part, we employ the Entropy Maximizing Method for the static OD matrix estimation and extend it to time-dependent method.

In this study, we propose a new algorithm replacing the first part with one using the simulation. The algorithm consists of two parts; (a) simulating part and (b) estimating part. In the simulating part, a dynamic traffic simulation model reproduces traffic condition, and the relationship between OD volumes and traffic counts is obtained. On the other hand, in the estimating part, a dynamic OD matrix is estimated subject to the relationship calculated in the simulating part. These two parts are repeatedly implemented to obtain OD volumes. An outline of the algorithm is summarized as follows:

Step 0-1 :Prepare a three-dimensional network.
Step 0-2 :Define driver's route choice model.
Step 0-3 :Set parameters of traffic capacity and the route choice model in the simulation.
Step 0-4 :Initialize the OD volumes.
Step 1 :Implement the traffic simulation and identify on which links each OD volume passes.
Step 2 :Update a time-dependent OD matrix by applying the entropy maximizing method under constraints of the relationship between the OD volumes and observed link flows.
Step 3 : Replace current OD volumes with updated OD volumes.
Step 4 : If converge, finish. Otherwise, go to Step 1.

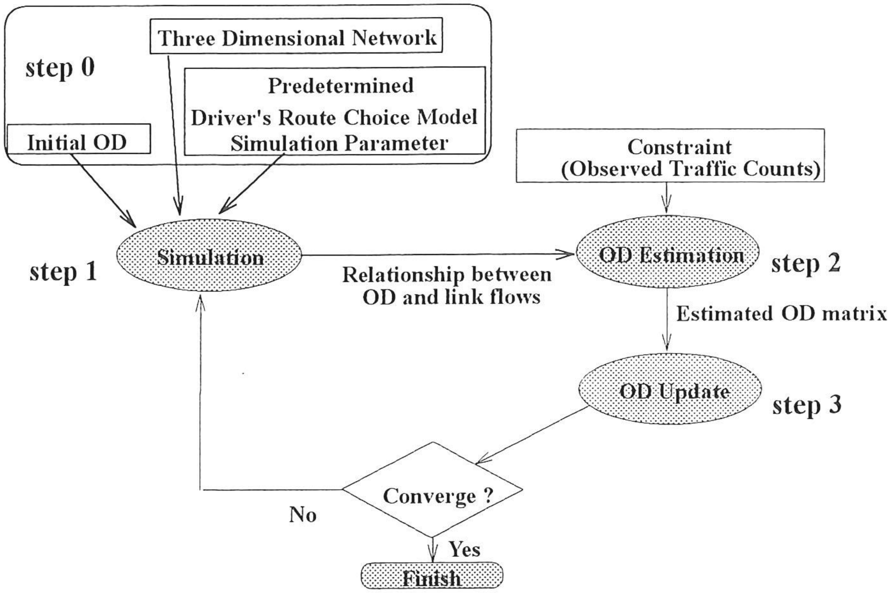

Figure 1 - A Proposed Estimation Algorithm

Three Dimensional Network

First, we define a three dimensional network. The time axis has been divided into time- intervals of equal length as shown in Fig.2 and Fig.3 in which a vehicle trajectory can be drown. For example, thick lines in both figures show the trajectory of a vehicle departing from origin at a time-interval h_r to destination along the route (link ***a-f***). In this case, the vehicle passes links ***a*** and ***b*** at time-interval h_r, links ***c***, ***d*** and ***e*** at time-interval h_r+dt and link ***f*** at time-interval h_r+2dt.

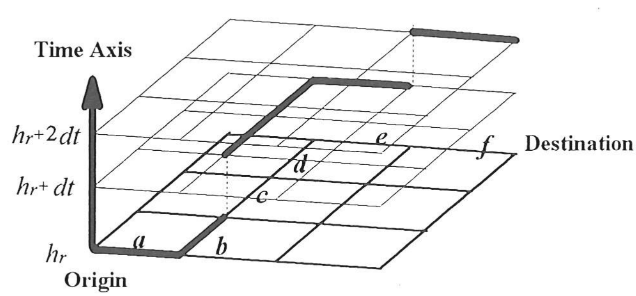

Figure 2 - A Three Dimensional Network

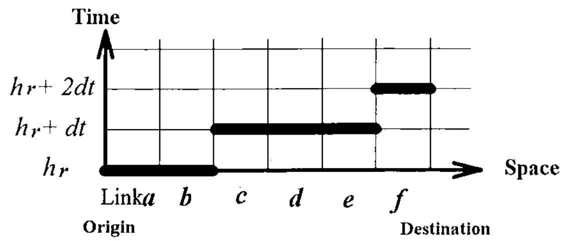

Figure 3 - A Vehicle Trajectory in a Time-Space Plane

Driver's Route Choice Behavior

Let us assume user's route choice probability such that

$$P_{kw}(h_r) = \text{Prob}\left[C_{kw}(h_r) + \varepsilon_{kw}(h_r) \leq C_{mw}(h_r) + \varepsilon_{mw}(h_r)\right] \qquad \forall m, \qquad (1)$$

where

$P_{kw}(h_r)$:route choice probability of path k of a vehicle departing from origin r of OD pair w at time-interval h_r

$C_{kw}(h_r)$:cost of path k of a vehicle departing from origin r of OD pair w at time-interval h_r

$\varepsilon_{kw}(h_r)$:an error term of $C_{kw}(h_r)$

Assuming that error term $\varepsilon_{kw}(h_r)$ has the Wible distribution, we obtain the following well known Logit model with parameter θ :

$$P_{kw}(h_r) = \frac{\exp(-\theta \cdot C_{kw}(h_r))}{\sum_m \exp(-\theta \cdot C_{mw}(h_r))} \qquad (2)$$

Dynamic Traffic Simulation

In Step 1, we use a dynamic traffic simulation model, SOUND (Kuwahara *et al.*, 1996), which reproduces dynamic traffic condition incorporating drivers' route choice behavior. SOUND requires time-dependent OD volumes and parameter values such as link traffic capacities, θ in the route choice model and others.

OD Estimation and Update

In Step 2, the OD volumes are estimated subject to constraints defined in Step 1 by applying the Entropy Maximizing Method (Oneyama *et al.*, 1996).

In Step 3, replace current OD volumes with calculated one in Step2 using Method of Successive Average (Sheffi, 1985) as shown in eqn. (3).

$$Q_{wh}^{n+1} = \frac{n \cdot Q_{wh}^{n} + R_{wh}}{n+1} \qquad (3)$$

where

Q^{n}_{wh} : OD volume on the n th iteration departing from origin r of OD pair w at time-interval h

R_{wh} : OD volume calculated at Step 2 departing from origin r of OD pair w at time-interval h

n : iteration number

APPLICATION

Application Network

We apply our estimating method to Toyota City area, whose network is shown in Fig.4. There are 18 origin nodes, 31 destination nodes and 540 OD pairs, and we observed traffic counts at 21 intersections and 13 parking facilities from 6:30 a.m. to 9:30 a.m. As traffic data have been obtained every 15 minutes, the time-interval is set as 15 minutes. Therefore, there are 12 time-interval and 6,480(=540 OD pairs × 12 time-interval) OD cells in this application. Since every directional flow in each observation point was counted, the total number of traffic observation data is 1,998.

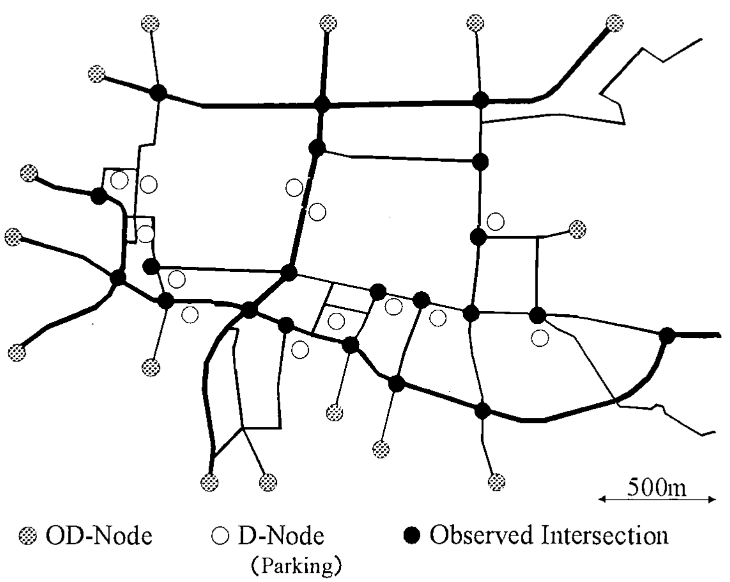

Figure 4 - Application Network (Toyota City Area)

Initialization

We set every cell of the initial OD matrix as volume 1[veh/15min.]. Traffic capacity at each link for the SOUND model has been determined by the geometric design and signal parameters, and parameter θ in the route choice model (that has been set as 0.01[sec^{-1}]).

Application Results

Application results are shown in figure 5. It shows the difference between observed traffic counts and reproducing flows in simulation. The vertical axis shows the averaged errors E^{avg}_{flow} determined in eqn(4). Average of traffic counts is about 60[veh/15min.].

$$E_{flow}^{avg} = \frac{\Sigma_a \Sigma_t \left| Q_{at}^{sim} - Q_{at}^{obs} \right|}{N^{obs}} \quad (4)$$

where

Q_{at}^{sim} : reproduced traffic flow at link a at time-interval t

Q_{at}^{obs} : observed traffic flow at link a at time-interval t

N^{obs} : number of traffic observation data

From this figure, it is surely understood that this iterating method converges well within several iteration. Figures 6 and 7 show the comparison between observed traffic counts and reproducing flows in simulation in result of using OD patterns estimated by new estimation method and previous estimation method (Oneyama *et al.*, 1996) respectively. From these figures, we confirm that simulation can reproduce observed traffic counts more accurately by using OD pattern estimated by our new method (Fig.6).

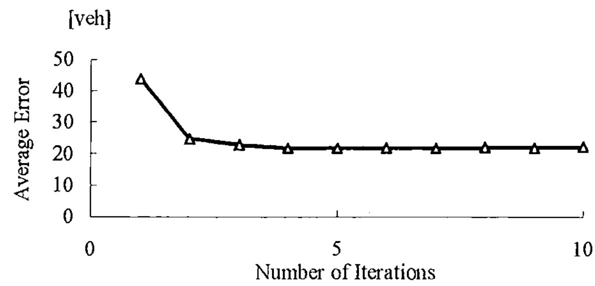

Figure 5 - Difference between Observed and Reproduced Traffic Flow

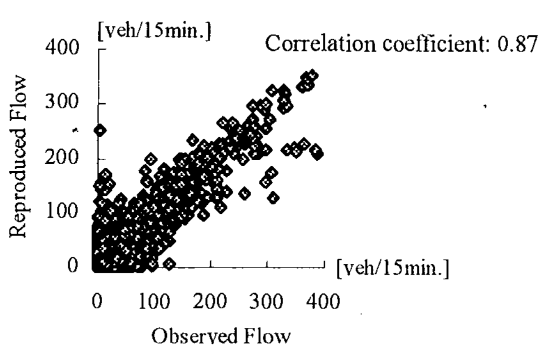

Figure 6 - Comparison between Observed and Reproduced Traffic Flow(*New* Method)

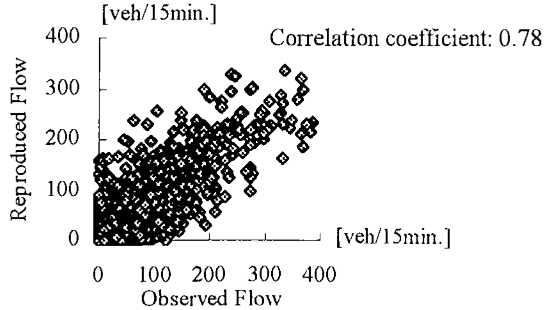

Figure 7 - Comparison between Observed and Reproduced Traffic Flow(*Previous* Method)

REPRODUCIBILITY OF OBTAINED OD MATRIX

The Reproducibility of Estimated OD Matrix

In order to investigate accuracy(reproducibility) of the estimated OD matrix, we implemented the estimating method under the evaluation framework using assumed counts. The evaluation framework is shown in Fig.8. At first, we regard estimated OD matrix in actual case as assumed OD. Then, we have produced link traffic counts by implementing SOUND under the assumed OD to take traffic counts. After that, using the same way as actual case, our estimating method estimates the OD matrix. In this way, it has been compared to assumed one in order to evaluate the reproducibility of the estimated OD matrix.

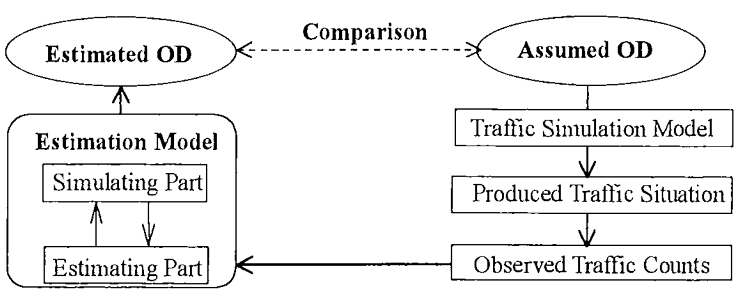

Figure 8 - Evaluation Framework (Assumed Case)

Fig.9 shows the difference between assumed OD volumes and estimated one. Average of OD volume is about 6[veh/15min.]. Vertical axis shows the averaged errors E_{od}^{avg} determined in eqn(5).

$$E_{od}^{avg} = \frac{\Sigma_r \Sigma_s \Sigma_t \left| q_{rst}^{est} - q_{rst}^{asm} \right|}{N^{od}} \qquad (5)$$

where

q_{rst}^{est} : estimated OD volume departed at origin r to destination s at time-interval t

q_{rst}^{asm} : assumed OD volume departed at origin r to destination s at time-interval t

N^{od} : number of OD pairs

A time-interval is set as 15 min. It seems that this estimation converges at several iterations.

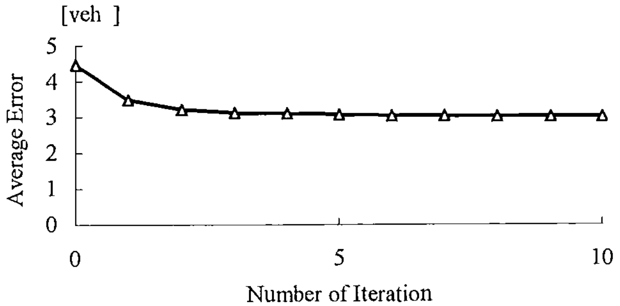

Figure 9 - Difference between Assumed and Estimated OD

Fig.10 compares assumed link flows and reproduced ones. Correlation coefficient is 0.92. Comparing to the actual case using observed counts(shown in Fig.6), the reproducibility becomes better. This reason for the difference is that many kind of error items, such as the observation error, simulation accuracy and others, are included in actual case, but not in this assumed case.

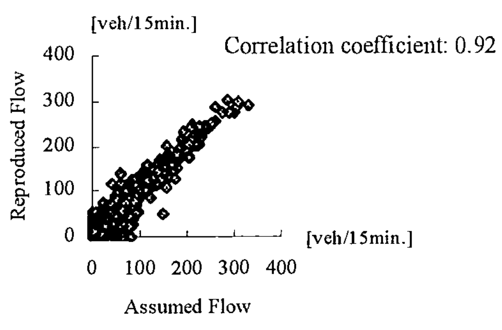

Figure 10 - Assumed and Reproduced Flows

Fig.11 compares assumed OD volumes and estimated ones. Correlation coefficient is 0.42. From these figures, we remark that reproduced link flows are reasonably agree with assumed ones but estimated OD volumes are not well fit to assumed ones. This result would be caused by the estimation method, which tries to reproduce link flows. As we have to estimate 6,480 OD volumes subject to 1,998 constraints (observed counts) in this application, there are many OD patterns that satisfy the constraints. This method picks up one of these OD patterns so as to maximize its entropy. Therefore, in order to estimate OD volumes more accurately, we should have to get more information about actual OD volumes and include it to the estimation method.

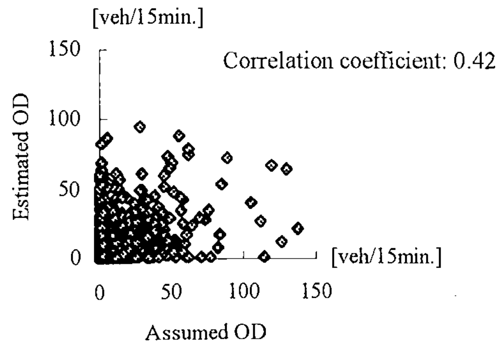

Figure 11 - Assumed and Estimated OD Volumes (time-interval = 15- min.)

We implement this estimating method with 1-hour time-interval instead of 15-min. Fig.12 compares assumed OD volumes with estimated ones. Comparing to the result with 15 minutes time-interval (shown in Fig.11), estimated OD is more consistent with assumed one. In generally, OD volumes at an OD pair much fluctuate within short time, and in case of much fluctuation, the entropy of the OD pattern becomes smaller. Whereas this method determines estimated OD so as to maximize its entropy. Therefore, estimated OD pattern is less consistent when OD volumes have more fluctuation. It seems that 1-hour time-interval is enough to cover fluctuation but 15 minutes is not enough. In addition to this reason, relationship between travel time from origin to destination and the time-interval is quite important. In case 1-hour time-interval, most of trips can reach his destination within same interval, on the other hand, in case 15 minutes time-interval, many trips cannot reach. We understand if time-interval is longer, more consistently we can get estimated OD matrix because of these reason. However, it should be careful that the OD matrix in longest time-interval is equal to static one. Dynamic estimation must be required to realize the optimum traffic control and planning.

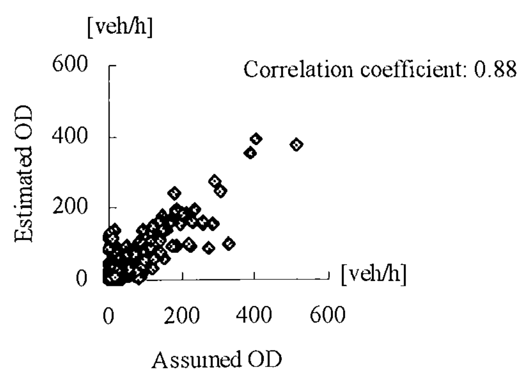

Figure 12 - Assumed and Estimated OD Volumes (time-interval = 1- hour)

Sensitivity of the Estimated OD Matrix to Simulation Parameters

Using this framework with 15 minutes time-interval, we investigate how accurately the OD matrix can be estimated in relation to the accuracy of simulation parameters, that is traffic capacity at each link and parameter θ in the route choice model. In both cases, the evaluation has been done after

10 iterations, because 10 iterations are enough to converge by the result shown in Fig.9.

Traffic Capacity

We investigate to what extent the accuracy of traffic capacity at each link affects the estimated OD. The estimation method has been implemented 10 cases. In each case, less than ±1% or ±3% errors are included in traffic capacities at each link using the random sequence. Fig.13 shows the average of difference in 10 cases between the estimated OD and assumed one. From this figure, we confirm that more exactly simulation parameters (traffic capacity) have been set, in other words, more accurately simulation model reproduces actual traffic condition, more accurately we can obtain the estimated OD matrix.

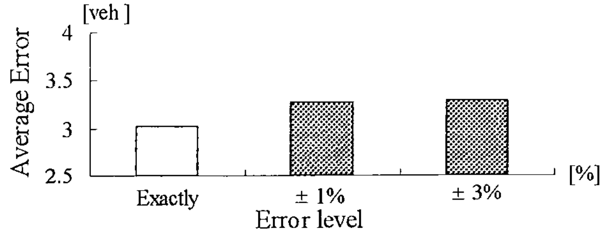

Figure 13 - Effect of the Accuracy of Traffic Capacity

Parameter θ in Route Choice Model

For another investigation, we investigate to what extent θ in Route Choice Model affects the estimated OD. Fig.14 shows the difference between the estimated OD and assumed one. From this figure, it would be confirmed that more exactly route choice behavior has been described, more accurately we can obtain the estimated OD matrix.

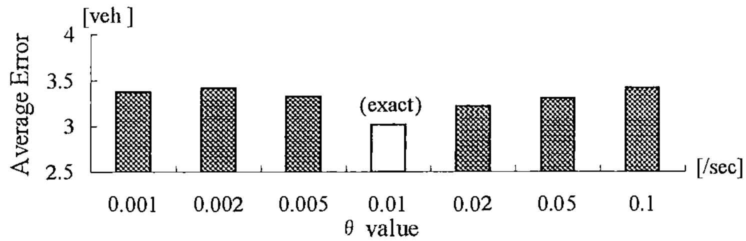

Figure 14 - Effect of the Accuracy of Route Choice Model

Sensitivity of the Estimated OD Matrix to the Number of Observed Counts

In this section, we investigate how accurately the OD matrix can be estimated in relation to the number of observed traffic counts. The estimation method has been implemented under different number of observed traffic counts, 1,998, 1,092 and 642. Number of observed point is restricted and generally less than number of OD pairs, and 1,998 counts is almost the maximum number in this case. Fig.15 shows the difference between the estimated OD and assumed one. From this figure, it can be confirmed that more number of observed traffic counts, more accurately we can obtain the estimated OD matrix. Furthermore, it remarks that the number of traffic counts strongly affects the reproducibility of estimated OD matrix because difference between the results of each case is very large.

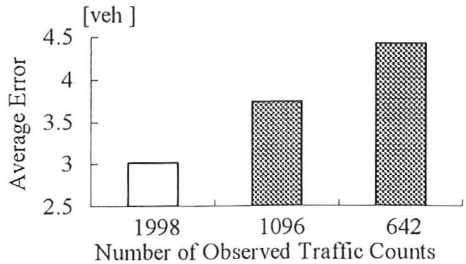

Figure 15 - Effect of the Number of Observed Traffic Counts

Sensitivity of the Estimated OD Matrix to the Additional OD Information

Using additional information about OD volumes, the estimated OD matrix is expected to be more accurate. Therefore, we focus the information about exact trips of several OD pairs, and investigate how accurately the OD matrix can be estimated in relation to the number of OD information. After exact volumes of 12 and 90 OD pairs are informed, the estimation method has been implemented. The number of informed OD pairs are 2% and 16% of total number of OD pairs, and the number of informed trips are 7% and 25% of total trips respectively. Fig.16 shows the difference between estimated OD and assumed one in both indexes average of estimated OD pairs (except informed OD pairs) and every OD pairs. From this figure, though the estimated OD surely becomes accurate with more information, difference between the results of each case is not so large. Average error of estimated OD pairs are almost the same. Then it remarks that OD information does not strongly affect the reproducibility of estimated OD matrix because of the followings. Even if exact volumes of 90 OD pairs are informed, there are still many OD patterns which satisfy constraints because the number of remaining OD volumes (5,400 OD) to be estimated is still greater than the number of observed counts (1,998 counts). It is required to get exact OD pattern that the number of OD volumes is equal to (or less than) the number of constraints.

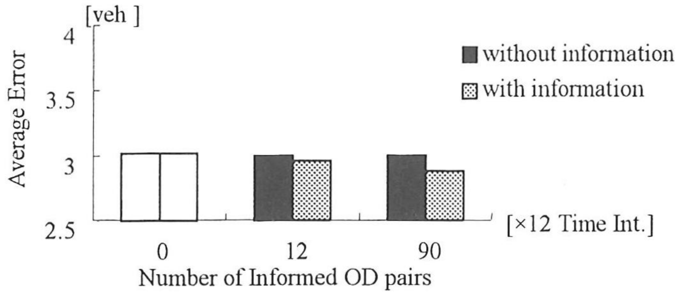

Figure 16 - Effect of the Number of Informed OD Pairs

SUMMARY AND FUTURE SCOPE

We propose a method of estimating a time dependent OD matrix from traffic counts using dynamic traffic simulation. Then, we investigate how accurately the OD matrix can be estimated in relation to the accuracy of simulation parameters. As a result, though this method can not estimate an exact OD matrix, we understand to what extent the estimated OD matrix differs from exact one and the reproducibility of the estimated OD is

a) affected by the accuracy of simulation parameters.
b) strongly affected by the number of observed counts.

c) not strongly affected by additional OD information which cannot equalize the number of OD volumes to the number of observed counts.

Since these results are derived from an application, these may charge depending on OD patterns and application networks.

Some of future research topics would be:

a) We should understand how accurately OD volumes can be estimated by other alternative method.
b) We should understand to what extent traffic condition reproduced by the simulation includes errors in relation to the estimated OD and simulation parameters.
c) An actual traffic data set is required to validate the method.
d) Although the route choice behavior of a driver is included in simulation models, there are still enough rooms to be studied on the human factor.

REFERENCES

Ashok, K. and Ben-Akiva, M.E.(1993) Dynamic Origin-Destination Matrix Estimation and Prediction for Real-Time Traffic Management Systems. **Transportation and Traffic Theory**, pp465-484

Cascetta, E., Inaudi, D. and Marquis, G. (1993) Dynamic Estimators of Origin-Destination Matrices Using Traffic Counts. **Transportation Science 27(4)**, pp363-373

Kuwahara, M., Yoshii, T., Morita, H. and Okamura, H. (1996) A Development of a Dynamic Traffic Simulation Model for Urban Road Networks: SOUND. **Research Report at Institute of Industrial Science, Univ. of Tokyo 48(10)**, pp49-52 (In Japanese)

Oneyama, H., Kuwahara, M and Yoshii, T.(1996) Estimation of a Time Dependent OD Matrix from Traffic Counts. **The Third Annual World Congress on Intelligent Transport Systems '96 Orlando**, CD-ROM

Sheffi, Y.(1985) **Urban Transportation Networks : Equilibrium Analysis with Mathematical Programming Methods.** Prentice Hall

Yang, H., Iida, Y. and Sasaki, T.(1990) Dynamic Estimation of Real Time OD Trip Matrix from Traffic Counts. **Proceedings of Infrastructure Planning 13**, pp599-606 (In Japanese)