Empirical Studies on Massive Data Operation for Route Choice and Travel Information

Hideki YAGINUMA *

Engineering Department,

Highway Planning Inc.,Tokyo, Japan

Tomoyuki ADACHI

Maintenance and Operation Department,

West Nippon Expressway Company Ltd., Osaka, Japan

Yasuo ASAKURA

Department of Civil Engineering,

Tokyo Institute of Technology, Tokyo, Japan

***** Email: yaginuma@doro.co.jp

Abstract

Recently, the inter-urban expressway in Japan tends to be networking, and multiple-routes become available in many origin-destination pairs. Adachi et al. [1,2] pointed out the networking effect of expressway that has improved the travel time and its reliability associated with decreasing travel distance in Shin-Meishin expressway between Nagoya and Osaka. In parallel with the improvement of network configuration, the road administrators provide traffic information to drivers such as travel time, jam length, occurrence of incident by using the roadside variable message sign (VMS). The information is expected to assist expressway user's route choice and to contribute to the efficient road operation. Thus, it is necessary to analyze user's route choice behavior in the multiple-route sections, and the sensitivity of their response to the provided information through VMS.

Some previous studies have discussed driver's behavior and response to VMS. Wardman et al. [3] and Chatterjee et al. [4] investigated the relationship between driver's route choice and VMS information by using stated preference (SP) survey. They developed the discrete choice model based on the traffic state variables and the individual socio-economic variables. Hato et al. [5] studied the route choice model considering information reference behavior under multiple information sources. These studies depend on the small size of data samples obtained by SP survey. On the other hand, Oguchi et al. [6] developed the driver's behavior

data generation system in the Tokyo metropolitan expressway by using multi-data sources such as automatic vehicle identification (AVI) data, traffic flow data by detectors, and VMS information log data. Nagai et al. [7] studied the route choice model using the OD data obtained with the electronic toll collection (ETC) system. The AVI data and ETC show the revealed preference of drivers, and the sample size is larger than the SP survey data. However, those data may fail to identify driver's route as the ETC data identify the pair of on-and-off toll gates.

This paper aims to analyze the relationship between expressway user's route choice behavior and provided travel information. We first intend to develop the data generation system of the individual user's route choice behavior by combining the multiple data sources such as ETC data, traffic detector data, and Free Flow Antenna (FFA) data. The proposed system is used to obtain a massive time series individual based route choice data. The generated route choice data are then linked with the VMS information log data. It becomes possible to analyze more detailed route choice behavior by using the massive behavioral data set matched with the provided travel information.

The proposed system consists of two modules. The first module is to estimate individual actual route choice by using ETC and FFA data. The ETC data contain individual user's ID, the type of car, and trip record such as the name of toll gates, day and time. FFA is the observation device installed in the expressway. The individual signal of an on-board ETC device is detected and the individual user's ID passing under the FFA is collected. The first module of the proposed system traces the same ID number recorded in ETC data source and FFA data source, and determines the route of the ID travelling in the expressway network. The second module is to match the route choice with the travel information. The actual travel information provided to a driver passing under the VMS at every moment is found in the VMS log data. The travel time from an exit gate to a roadside VMS is calculated by using traffic speed data obtained by detectors, and the corresponding travel information in the VMS log data is matched with the individual driver who has passed under the VMS at that time.

The system was applied to the actual expressway network. There are two routes available between Kusatsu and Toyota junctions. The length of Meishin route and Shin-Meishin route are 160 km and 126 km, respectively. The period of analysis is six months from December 2010 to May 2011 (182 day). The number of route choice data is 2.6 million. The east-bound

and west-bound directions are 1.34 million (7,385 samples/day) and 1.26 million (6,915 samples/day), respectively. This amount covers 99.9% of the total number of ETC users in the target area.

We have analyzed the share of each route segmented by car type, time/month period, and the contents of VMS information. Logit type route choice model was applied to the datasets to describe relationships between route choice and travel information. The estimated results of the random sampling with 10,000 observations, the ratio of likelihood was 0.557 and the t-value of each variable was 1% significant. In addition, we also estimated the segmented dataset such as days, months, and types of vehicles.

We also investigate about the massive data processing method. In particular, the optimal number of the sample from the massive datasets is discussed considering the trade-off between stable estimator and computational time. We also try to involve the data compression method such as data squashing in the context of data mining.

References

[1] Adachi, T., Fujikawa, K. and Asakura, Y., “Empirical Studies on Travel Time Reliability for Expressway with Alternative Routes”, 16th ITS World Congress, Stockholm, Sweden, 2009.

[2] Adachi, T., Fujikawa, K. and Asakura, Y., “Empirical Analysis for Improvement of Travel Time Reliability Due to Opening of New Expressway”, Proceedings of the 15th International Conference of Hong Kong Society for Transportation Studies, Hong Kong, China, 2010.

[3] Wardman, M., Bonsall, P.W., Shires, J.D., “Driver response to variable message signs: a stated preference investigation”, Transportation Research Part C, Vol.5, pp.389-405, 1997.

[4] Chatterjee, K.,Hounsella, N.B., Firminb, P. E. and Bonsall, P.W., “Driver response to variable message sign information in London”, Transportation Research Part C, Vol.10, pp.149-169, 2002.

[5] Hato, E., Taniguchi, M., Sugie, T., Kuwahara, M., and Morita, H., “Route Choice Model Considering Driver’s Information Acquisition and Reference Behavior under the Multiple Traffic Information Sources”, Journal of Infrastructure Planning and Management, Vol.597, IV-40, pp.99-111, 1998. (in Japanese)

[6] Oguchi, T., Hato, E., Taniguchi, M., Yoshii, M., Yoshii, T., Kuwahara, M., and Morita, H., “Survey on the Actual Phenomenon about Route Choice in Tokyo Metropolitan Expressway”, Journal of Infrastructure Planning and Management, Vol.590, IV-39, pp.87-85, 1998. (in Japanese)

[7] Nagai, M., Hibino, N., and Morichi, S., “A Study on Route Choice Behavior in Urban-Expressway Based on ETC-OD Data”, Proceedings of Infrastructure Planning, Vol.42 CD-ROM, 2010. (in Japanese)