An Efficient Estimation Method for Traffic States Using Incomplete Floating Car Data in a Large Network

School of Engineering, Hokkaido University	○Student Member	Kaifeng	Duan
Faculty of Engineering, Hokkaido University	Member	Ryuichi	Tani
Faculty of Engineering, Hokkaido University	Member	Kenetsu	Uchida
DOCON CO., LTD.	Nonmember	Keita	Zushii
DOCON CO., LTD.	Member	Osamu	Nagaoka

1. Introduction

The field of traffic planning has witnessed a significant impact due to the advancements in information and communication technology. To cope with a road network where some of the link data is unavailable, the traffic states estimation is formulated as a full information maximum likelihood estimation (FIMLE) problem with UE constraints in this study. An efficient algorithm for solving the FIMLE problem based on the sensitivity analysis for UE is also presented.

2. Methods

2.1. Full information maximum likelihood estimation (FIMLE)

We assume that an hourly generated traffic demand Q follows a lognormal distribution, $Q \sim LN(\mu, \sigma^2)$ where μ and σ^2 are parameters of the distribution. The traffic states of all road network can be estimated by maximizing the likelihood function of the multivariate normal distribution as

$$\max \quad L\left(\mu, \sigma^2 \middle| \widehat{\boldsymbol{d}}_g \forall g \in \{1, \ldots, G\}\right) = \prod_{g \in \{1,\ldots,G\}} \frac{\exp\left(-\frac{1}{2}\left(\widehat{\boldsymbol{d}}_g - \boldsymbol{\mu}_g(\mu)\right)^T \boldsymbol{\Sigma}_g^{-1}(\sigma^2)\left(\widehat{\boldsymbol{d}}_g - \boldsymbol{\mu}_g(\mu)\right)\right)}{2\pi^{\frac{n(g)}{2}}\left|\boldsymbol{\Sigma}_g(\sigma^2)\right|^{\frac{1}{2}}} \tag{1}$$

s.t.

$$\sum_{i \in I} \sum_{j \in J_i} c_{ij}\left(f_{ij} - f_{ij}{}^*\right) \geq 0 \tag{2}$$

$$\sum_{i \in I} \sum_{j \in J_i} f_{ij} = q_i \tag{3}$$

$$f_{ij} \geq 0 \ \forall i \in I, \forall j \in J_i \tag{4}$$

Equations (2)-(4) are the user equilibrium constraints, where $\boldsymbol{\Sigma_g}(\sigma^2) = \boldsymbol{M}_g \boldsymbol{\Sigma}_{T'_a}(\sigma^2)\boldsymbol{M}_g^T$, $\boldsymbol{\mu}_g(\mu) = \boldsymbol{M}_g \boldsymbol{\mu}_{T'_a}(\mu)$ and $\widehat{\boldsymbol{d}}_g = \boldsymbol{M}_g \widehat{\boldsymbol{d}}_g^T \ \forall g \in \{1, \ldots, G\}$. $n(g)$ is the number of observed traffic data in gth day. Note that $\mathbf{M}_g$ is a matrix used for extracting the corresponding parts of $\boldsymbol{\Sigma}_{T'_a}(\sigma^2), \boldsymbol{\mu}_{T'_a}(\mu)$ or $\widehat{\boldsymbol{d}}_g^T$ to the observed traffic states.

2.2. Sensitivity analysis

It is complex to solve the likelihood function with equilibrium constraints directly and which requires a large amount of computation time. However, the method of sensitivity analysis converts the equilibrium constraints to linear constraints, which makes the objective function finding the optimal solution easier.

The perturbation parameter for OD demands will be denoted by $\boldsymbol{\varepsilon}$. The mean vector of OD demands considering perturbation parameter $\boldsymbol{q}(\boldsymbol{\varepsilon})$ is represented by $\mathbf{q}(\boldsymbol{\varepsilon}) = \boldsymbol{q} + \boldsymbol{\varepsilon}$. The path flow calculated at $\mathbf{q}(\boldsymbol{\varepsilon})$ can be approximated by $\tilde{\boldsymbol{f}}(\boldsymbol{q}(\boldsymbol{\varepsilon})) = \boldsymbol{f}^*(\boldsymbol{q}) + \boldsymbol{\nabla}_{\boldsymbol{\varepsilon}}\boldsymbol{f} \cdot \boldsymbol{\varepsilon}$, where $\boldsymbol{f}^*$ is vector of path flows at user equilibrium, $\boldsymbol{\nabla}_{\boldsymbol{\varepsilon}}\boldsymbol{f}$ is calculated by the method of sensitivity analysis. Thus, $\tilde{\boldsymbol{f}}(\boldsymbol{q}(\boldsymbol{\varepsilon}))$ is the vector of path flows calculated at $\boldsymbol{q}(\boldsymbol{\varepsilon})$. The link flows can be defined as $\boldsymbol{v}(\boldsymbol{\varepsilon}) = \tilde{\boldsymbol{f}}(\boldsymbol{q}(\boldsymbol{\varepsilon})) \cdot \Delta$, The ratio of link flows to mean OD demands in is calculated as $\hat{p}_a(\boldsymbol{\varepsilon}) = v_a(\boldsymbol{\varepsilon})/q(\boldsymbol{\varepsilon})$.

Keywords Full information maximum likelihood estimation, ETC2.0 data, Sensitivity analysis
Contact 5 Chome Kita 8 Jonishi, Kita Ward, Sapporo, Hokkaido 060-0808, Japan
Graduate School of Engineering., Hokkaido University

The maximum likelihood estimation problem can be written as follows.

$$\max L\left(\mu, \sigma^2, \boldsymbol{\varepsilon} \middle| \widehat{\boldsymbol{d}}_g \forall g \in \{1, \dots, G\}\right) = \prod_{g \in \{1,\dots,G\}} \frac{\exp\left(-\frac{1}{2}\left(\widehat{\boldsymbol{d}}_g - \boldsymbol{\mu}_g(\boldsymbol{\varepsilon}, \mu)\right)^T \boldsymbol{\Sigma}_g^{-1}(\sigma^2)\left(\widehat{\boldsymbol{d}}_g - \boldsymbol{\mu}_g(\boldsymbol{\varepsilon}, \mu)\right)\right)}{2\pi^{\frac{n(g)}{2}} \left|\boldsymbol{\Sigma}_g(\sigma^2)\right|^{\frac{1}{2}}} \tag{5}$$

s.t.

$$\sum_{i \in I} q_i(\boldsymbol{\varepsilon}) = \exp\left(\mu + \frac{\sigma^2}{2}\right) \tag{6}$$

3. Numerical Calculations

The road network in Asahikawa, Japan, is used for the numerical calculation. The network has 340 directed links, 135 nodes, and 182 OD pairs.

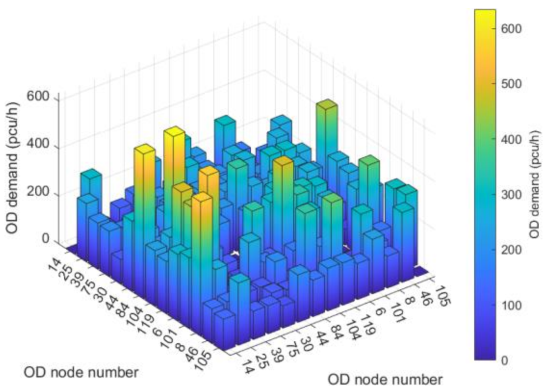

Figure 1. Demand for each OD pair for the Asahikawa network.

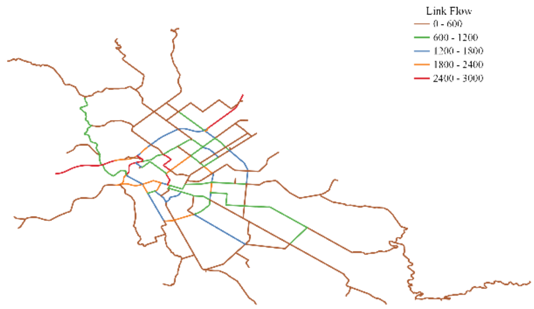

Figure 2. Estimated average link flow.

Figure 1 shows the estimated mean demand for each OD pair (there are a total of 182 OD pairs). The total OD demand estimated is 3.2406×10^4[pcu/hour]. The estimated OD demands for OD pairs 44-14, 44-6, 104-6, and 84-14 are significantly larger than those for other OD pairs, given that the centroid node of 84 is located near a primary school and the centroid nodes of 44, 104 and 14 are located near residential areas. In the vicinity of centroid node 6, there is the expressway from the outskirts to the center of Asahikawa City.

Figure 2 shows the estimated average link flow for the Asahikawa network. In this study, a road section in the network is comprised of opposing two directed links, and since the flows of opposing two directed links are essentially the same. The link flow shown in Figure 4 is the average of mean link flows for the two opposing directed links. The average link flow in the central part is around 600-1,200 [pcu/hour], and the links in red have flows between 2,400-3,000 [pcu/hour]. The reason for this is that the links are located near the centroid nodes.

4. Conclusions

This study developed a method to estimate the traffic states by the FIMLE problem with UE constraints. In addition, a solution algorithm for the FIMLE problem based on the sensitivity analysis for UE is also presented.

Since the market penetration rate of the ETC2.0 system is not high, therefore the fully observed data cannot be obtained. The proposed method in this study can address such incomplete traffic data.

Numerical calculations are carried out in the real Asahikawa network using real observed link delays. Especially, the method proposed in this study estimated relatively reasonable link delays in the Asahikawa network, and the issue of traffic condition estimation for multiple time periods in a large network will be addressed in the future study.

References

Tani, R., Uchida, K. : An Estimation Method for Spatio-temporal Traffic States Based on Incomplete Traffic Observation. The 8th International Symposium on Dynamic Traffic Assignment, 2021.

