Transportation Research Part C 36 (2013) 446–459

Contents lists available at ScienceDirect

Transportation Research Part C

journal homepage: www.elsevier.com/locate/trc

A Bayesian approach to traffic estimation in stochastic user equilibrium networks ☆

Chong Wei [a,b,*], Yasuo Asakura [b]

[a] MOE Key Laboratory for Urban Transportation Complex Systems Theory and Technology, Beijing Jiaotong University, Beijing 100044, China
[b] Department of Civil and Environmental Engineering, Tokyo Institute of Technology, 2-12-1-M1-20, Ookayama, Meguro, Tokyo 152-8552, Japan

ARTICLE INFO

Article history:
Received 25 May 2013
Accepted 30 June 2013

Keywords:
Path flow estimation
Bayes' theorem
Likelihood
Stochastic user equilibrium
Markov chain Monte Carlo

ABSTRACT

This study proposes a statistical model to estimate route traffic flows in congested networks. In the study, it is assumed that route traffic flows conform to the stochastic user equilibrium (SUE) principle while being treated as random variables in order to exploit the stochastic nature of traffic. The proposed model formulates the distribution of these random variables as the conditional distribution of route flows given the observed link flows and the SUE principle. Here, the SUE principle is accounted for through the likelihood of user behaviours rather than by using a bi-level formulation. In this study, the Bayesian theorem is applied to derive the probability density function (PDF) of the conditional distribution. Based on the PDF, characteristics such as the means and variances of route/link traffic flows are estimated using a blocked Metropolis–Hastings (M–H) algorithm. To facilitate the use of prior knowledge, a hierarchical form is designed to provide a straightforward way to integrate prior knowledge into the traffic estimation model. The performance of the proposed method is tested on the Sioux–Falls network through a series of numerical examples.

1. Introduction

Knowing the traffic state of each link across a network is a fundamental requirement of telematics systems. However, contemporary traffic sensors such as inductive loops and Global Positioning System (GPS) devices can only cover parts of a road network, and the traffic volumes on unobserved links are not obtained. In order to improve the availability of traffic data, we need to develop a method that can estimate the traffic flows on unobserved links.

This kind of problem, which is usually referred to as a path flow estimation problem in the literature (see Bell, 1991), is related to the Origin–Destination (OD) matrix estimation problem that has been widely discussed over the past few decades. Sasaki (1967) pointed out the usefulness of an entropy model for estimating travel demands. Van Zuylen and Willumsen (1980) and Van Zuylen (1981) presented methods to find the most likely O–D matrix using maximum entropy and minimum information estimators. A number of previous studies also indicated that the traffic assignment model should be integrated into the O–D matrix estimators in order to take into account the dependency between route travel times and route flows in congested networks. For example, Yang et al. (1992) developed a bi-level model in which route flows are constrained by the

☆ This paper was presented at the 20th International Symposium on Transportation & Traffic Theory. It therefore also appears in the complete proceedings of the 20th ISTTT in [Procedia – Social and Behavioral Sciences, vol. 80C (2013), pp. 591–607].

* Corresponding author at: MOE Key Laboratory for Urban Transportation Complex Systems Theory and Technology, Beijing Jiaotong University, Beijing 100044, China. Tel.: +86 10 5168 7154.

E-mail addresses: weichongcn@gmail.com, chwei@bjtu.edu.cn (C. Wei).

0968-090X/$ - see front matter
http://dx.doi.org/10.1016/j.trc.2013.06.013

User Equilibrium (UE) principle; Bell et al. (1997) proposed a method to estimate route flows in Stochastic UE (SUE) networks.

On the other hand, researchers have noted that the stochastic nature of traffic flows should not be ignored. Bell (1991) and Cascetta (1984) looked at the problem from a statistical perspective and suggested that a generalised least-squares estimator be used to account for the random terms of O–D demands and route traffic flows. More recently, Xie et al. (2011) showed the benefit of combining a maximum entropy model and a least-square estimator for traffic estimation problems. Unlike its use of a generalised least-squares estimator, the approach of Spiess (1987) employs a class of maximum likelihood estimators to solve the O–D estimation problem, a method that shows the potential effect of the likelihood principle. Watling (1994), Lo et al. (1996), Hazelton (2000), and Parry and Hazelton (2012) also constructed different likelihood-based approaches to the problem. Maher (1983), Castillo et al. (2008), Hazelton (2008), Li (2009), Yamamoto et al. (2009) and Perrakis et al. (2012) formulated likelihood-based models from Bayesian perspectives. Nevertheless, almost all of these likelihood-based approaches were originally developed for uncongested networks, so there is still a need to explore the application of the likelihood principle to capture the stochastic nature of traffic flows in a congested network.

This study utilizes a Bayesian approach to solve the problem. Unlike previous studies, however, the proposed method takes into congestion effects through the likelihood of route choices on congested networks, rather than using a bi-level formulation.

We assume that traffic flow patterns conform to the SUE principle while treating route flows as random variables and seeking to estimate the characteristics of the random route flow variables. To do this, we need to obtain the probability distribution of the route flow variables. Since we consider that the route traffic flow patterns are constrained by both link traffic counts and the SUE principle, this probability distribution must be precisely represented as the conditional distribution of route traffic flows for given link traffic counts, in conformity to the SUE principle.

To exploit the conditional distribution, Bayes' theorem is adopted in this study in order to decompose the conditional distribution into a likelihood function and prior probability distributions. We first address the likelihood function and the prior probability distributions, and then combine the results to yield a formulation of the conditional distribution.

To obtain the characteristics of link flow variables, we develop a blocked Metropolis–Hastings algorithm to sample the conditional distribution of route flows, and then aggregate the samples to produce the characteristics of link flows or O–D flows.

The highlights of the proposed method are as follows:

(1) The proposed model is a likelihood-based statistical estimation model that can take into account users' contemporaneous interactions on a congested network without using bi-level formulations and can guarantee the estimate is unique; the method does not find the equilibrium solution in each iteration and does not impose specific requirements on the application of user behaviour models.
(2) The model can handle the inconsistencies among observed link traffic counts.
(3) The model can work with a hierarchical form to flexibly integrate prior knowledge.

The remainder of this article is organised as follows. Section 2 formulates the conditional probability distribution of route traffic flow variables. Section 3 outlines the blocked Metropolis–Hastings algorithm that is used to estimate the characteristics of link traffic flow variables. Section 4 provides some numerical examples that demonstrate the effectiveness of the proposed method. Section 5 offers a conclusion to the study.

2. Methodology

Let R be the set of routes; N, the set of O–D pairs; R_n, the set of routes that connect the O–D pair n; L, the set of links across the network; and L^*, the set of observable links. I_n, denotes the set of users who make trip between O–D pair n. $\mathbf{q} = [q_1, \ldots, q_{|N|}]$ is the vector of O–D demands; $\mathbf{c}_n = \{c_i|\forall i \in I_n\}$ is the set of the route choice result variables corresponding to O–D pair n; and $\mathbf{y} = [y_1, \ldots, y_{|R|}]$ is the vector of the route flow variables (note that the traffic flow on route r, y_r, is also the number of users who choose the route); and $\mathbf{x}^* = \{x_l^*|\forall l \in L^*\}$ denotes the link traffic counts during a given time period. Other notations will be defined when they are first introduced.

2.1. Estimate traffic with pre-specified O–D demand

We begin by solving a primary estimation problem in which we aim to estimate $\mathbf{y}$ based on the following factors:

(1) The O–D demand vector $\mathbf{q} = [q_1, \ldots, q_{|N|}]$ is given as a pre-condition that will not be estimated along with the other variables.
(2) A part of the links can be observed, i.e. $\mathbf{x}^*$ is available and is used to estimate $\mathbf{y}$.
(3) We presume the same user behaviour principle as in Bell et al. (1997) in considering that the route flows in a network conform to the SUE principle.