Transportation Research Part C 40 (2014) 93–110

Contents lists available at ScienceDirect

Transportation Research Part C

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

A generalized sensor location model for the estimation of network origin–destination matrices

Shou-Ren Hu [*], Han-Tsung Liou [1]

Department of Transportation and Communication Management Science, National Cheng Kung University, Tainan City 70101, Taiwan

A R T I C L E I N F O

Article history:
Received 8 December 2013
Received in revised form 6 January 2014
Accepted 6 January 2014

Keywords:
Network sensor location problem
Origin–destination (O–D) demand estimation
Heterogeneous traffic information
Link-flow estimator

A B S T R A C T

This study examined the network sensor location problem by using heterogeneous sensor information to estimate link-based network origin–destination (O–D) demands. The proposed generalized sensor location model enables different sensors' traffic monitoring capabilities to be used efficiently and the optimal number and deployment locations of both passive- and active-type sensors to be determined simultaneously without path enumeration. The proposed sensor location model was applied to solve the network O–D demand estimation problem. One unique aspect of the proposed model and solution algorithms is that they provide satisfactory network O–D demand estimates without requiring unreasonable assumptions of known prior information on O–D demands, turning proportions, or route choice probabilities. Therefore, the proposed model and solution algorithms can be practically used in numerous offline transportation planning and online traffic operation applications.

1. Introduction

Trip origin–destination (O–D) demand matrices are critical components in transportation network modeling, and provide essential information on trip distributions and corresponding spatiotemporal traffic patterns in traffic zones in vehicular networks. Trip O–D matrices also reflect traffic loadings and flow intensities in transportation networks, and are crucial inputs in determining short-term traffic control schemes and long-term transportation improvement programs, as well as offline transportation planning and online traffic management. Trip O–D demand matrices have traditionally been estimated by conducting household surveys or roadside interviews; however, this is unfeasible because of the high cost and data recording errors involved. Inferring network O–D demand matrices using corresponding link flows is an effective alternative approach, because link flows, which are a set of traffic flows associated with the vehicular trip distributions of different O–D pairs, are easily obtained. Past studies on the estimation of network O–D demand matrices using link flow information have generally assumed that link flows are fully observable. However, in practice, highway agencies face budget constraints in implementing comprehensive sensor deployment plans, and assuming the full observability of link flows is unreasonable. Because of the rapid development of information and communication technologies (ICTs), applications of advanced sensor technologies to traffic management and operation have become widespread and essential. As a result, determining the strategic deployment of traffic sensors to obtain necessary traffic information for network O–D demand estimation has become crucial in transportation network research.

---

[*] Corresponding author. Tel.: +886 6 2757575x53203; fax: +886 6 2753882.
E-mail addresses: shouren@mail.ncku.edu.tw (S.-R. Hu), iroya.liou@gmail.com (H.-T. Liou).

[1] Tel.: +886 6 2757575x53271; fax: +886 6 2753882.

0968-090X/$ - see front matter © 2014 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.trc.2014.01.004

1.1. The network O–D demand estimation problem

Traditional means for collecting O–D demand data by using random sampling techniques (Qrtuzar and Willumsen, 2001), such as roadside interviews, postcard surveys, and license plate recognition (LPR), are costly, time consuming, and labor intensive. Moreover, such methods are prone to sampling and data recording errors, which influence the confidence intervals of the surveyed O–D data, resulting in perturbed data that does not allow true O–D demand values to be accurately approximated. Beginning in the early 1980s, studies began to investigate using vehicle detectors (VDs) to estimate network O–D demands based on link flow information because of the rapid development of ICTs and intelligent transportation systems (ITS). These approaches can be divided into two categories:

(a) Statistical methods, such as the generalized least squares (GLS) (Casetta, 1984), maximum likelihood (ML) (Spiess, 1987), entropy maximizing (EM) (Teodorovic et al., 2002), Kalman filter (KF) (Nihan and Davis, 1987), and Bayesian inference methods (Maher, 1983).
(b) Mathematical programming methods, such as the linear programming (Sherali et al., 1994), nonlinear programming (Yang and Zhou, 1998), and bi-level programming methods (Yang et al., 1992).

In addition to link flow information, other types of traffic information have been incorporated into the estimation process, such as infrared beacons (Horiguchi et al., 2001), automatic vehicle identification (AVI) (Sherali et al., 1994), probe data (Eisenman and List, 2004), cellular data (Sohn and Kim, 2008), and LPR (Castillo et al., 2008d). These methods involve collecting data that link flows do not address, including vehicular trajectories, turning proportions, and partial path flows. These advanced approaches can potentially improve the accuracy of O–D demand estimates. In addition, studies have considered integrating heterogeneous traffic information to estimate network O–D demands (Dixon and Rilett, 2002; Eisenman and List, 2004; Castillo et al., 2008d; Hu and Liou, 2012). However, such an approach requires a link–path or an O–D/path/link incidence matrix to facilitate the path, or an O–D flow estimation process; link–path and O–D/path/link incidence matrices are difficult to obtain in practice because of the path enumeration problem. On the other hand, inferring network O–D demands simply using link flows could result in multiple solutions in conventional network O–D demand estimation models. To address the multiple solutions problem, certain studies have used path flow as a decision variable to estimate a set of O–D demands. This approach is called path flow estimation (PFE) (Sherali et al., 1994; Bell et al., 1997; Bell and Grosso, 1998; Chen et al., 2005); PFE is effective for obtaining an O–D demand estimate, and offers ease of calculation and rapid convergence. Despite these merits, most O–D demand estimation models based on link flow inferences or PFE typically involve the assumption that prior O–D information, O–D/path/link incidence matrices, turning proportions, or route choice probabilities are known in advance. These assumptions contradict reality, because some of these assumptions result from traffic assignment procedures or travelers' route choice behaviors given a set of unknown O–D demands. A robust network O–D demand estimation model that does not involve unreasonable assumptions is necessary for use in practical applications.

1.2. The network sensor location problem

In the literature, most developed network O–D demand estimation models using link flow information or PFE have either assumed that full link flow information is available (Casetta, 1984; Spiess, 1987; van Zuyen and Williumsen, 1980; Nihan and Davis, 1987; Maher, 1983) or that partial information on link flow measurements is available (Sherali et al., 2003). Theoretically, the quality of O–D demand estimates is directly related to the number of observed link flow measurements. Determining the appropriate number and installation locations of traffic sensors under certain budget constraints provides necessary link flow information to enable the satisfactory estimation of a set of unknown network O–D demands. This is referred to as the network sensor location problem (NSLP) (Bianco et al., 2001; Hu et al., 2009). Specific problem domains for the NSLP have been proposed. According to the network observability problem (Castillo et al., 2007), traffic flow states can be fully or partially inferred based on observed partial link flow measurements. Gentili and Mirchandani (2012) classified the NSLP into two sub problems: the sensor location flow-observability problem (e.g. traffic flow inference) and the sensor location flow-estimation problem (e.g., O–D demand estimation problem, travel time estimation). In the literature, three approaches to solving the NSLP have been proposed:

(a) The mathematical optimization approach, in which an objective is designed to maximize flow coverings (Yang and Zhou, 1998; Yim and Lam, 1998; Teodorovic et al., 2002; Ehlert et al., 2006; Hu and Liou, 2012) or minimize specific requirements (Chen et al., 2004; Minguez et al., 2010), under certain flow conservation constraints, budget constraints, or set-covering rules, to determine a set of deployed links.
(b) The algebraic approach, in which one-step or iteration procedures involving matrix algebra based on specific incidence matrices, such as the link–path incidence (Castillo et al., 2007, 2008a,b,c, 2013; Hu et al., 2009) or the link–node incidence matrices (Ng, 2012), were used to determine a set of deployed links.
(c) Graph theory, in which the spatial relationships between links and nodes in a network structure are decomposed to search for a set of critical links (Gentili and Mirchandani, 2005; Bianco et al., 2001; He, 2013).