Available online at www.sciencedirect.com
ScienceDirect
Transportation Research Part B 41 (2007) 823–840
TRANSPORTATION RESEARCH PART B
www.elsevier.com/locate/trb

A structural state space model for real-time traffic origin–destination demand estimation and prediction in a day-to-day learning framework

Xuesong Zhou [a], Hani S. Mahmassani [b,*]

[a] Department of Civil and Environmental Engineering, University of Utah, Salt Lake City, UT 84112, USA
[b] Department of Civil and Environmental Engineering, University of Maryland, College Park, MD 20742, United States

Received 28 November 2005; received in revised form 14 February 2007; accepted 15 February 2007

Abstract

Dynamic origin–destination (OD) estimation and prediction is an essential support function for real-time dynamic traffic assignment model systems for ITS applications. This paper presents a structural state space model to systematically incorporate regular demand pattern information, structural deviations and random fluctuations. By considering demand deviations from the a priori estimate of the regular pattern as a time-varying process with smooth trend, a polynomial trend filter is developed to capture possible structural deviations in real-time demand. Based on a Kalman filtering framework, an optimal adaptive procedure is further proposed to capture day-to-day demand evolution, and update the a priori regular demand pattern estimate using new real-time estimates and observations obtained every day. These models can be naturally integrated into a real-time dynamic traffic assignment system and provide an effective and efficient approach to utilize the real-time traffic data continuously in operational settings. A case study based on the Irvine test bed network is conducted to illustrate the proposed methodology.

Keywords: Real-time traffic estimation and prediction; Dynamic OD estimation and prediction; Kalman filter; Traffic system management

1. Introduction

The premise of Intelligent Transportation Systems (ITS) is the ability to sense prevailing conditions and rapidly devise actions to optimize system performance in real-time. Because the dynamics of traffic systems are complex, as they depend on the interaction of many independent agents (drivers) acting non-cooperatively in a spatially connected network, many situations call for strategies that anticipate unfolding conditions

* Corresponding author. Tel.: +1 301 405 0221; fax: +1 301 405 2585.
E-mail addresses: zhou@eng.utah.edu (X. Zhou), masmah@umd.edu (H.S. Mahmassani).

0191-2615/$ - see front matter
doi:10.1016/j.trb.2007.02.004

instead of adopting a purely reactive approach. Real-time simulation of the traffic network forms the basis of a state prediction capability that fuses historical data with sensor information, and uses a description of how traffic behaves in networks to predict future conditions, and accordingly develop control measures.

Dynamic origin–destination demand estimation and prediction is an important capability in its own right, and an essential support function for real-time dynamic traffic assignment model systems for ITS applications. The dynamic OD demand estimation and prediction problem seeks to estimate time-dependent OD trip demand patterns at the current stage, and predict demand volumes over the near and medium terms in a general network, given historical demand information and real-world traffic measurements from various surveillance devices (e.g. occupancy and volume observations from loop detectors on specific links).

Substantial research efforts have been devoted to dynamic demand estimation and prediction problems over the past 20 years. Existing models can be grouped into two classes: DTA-based vs. non-DTA-based, depending on whether a DTA component is incorporated into the estimation process (Kang, 1999; Chang and Tao, 1999; Peeta and Ziliaskopoulos, 2001). In this paper, existing models are categorized according to the underlying assumptions in representing dynamic demand processes. Assuming that the deviations of flow (demand) from historical averages define a stationary time series, the first group applies autoregressive (AR) models to the recursive estimation and prediction process. In the Kalman filtering formulation proposed by Okutani and Stephanedes (1984), the original data is first detrended from historical observations, then an AR model is used to estimate and forecast time-varying traffic flows on a single link. Along the same line, Ashok and Ben-Akiva (1993, 2000) formulated deviations of OD demand from historical averages as AR processes, and further developed a Kalman filter for real-time OD demand estimation and prediction, in which a 4th-order AR model is adopted based on several data sets. In general, an autoregressive model is suitable to describe a stationary random process with constant mean and variance. On the other hand, if the prevailing OD demand is structurally different from the regular demand pattern, demand deviations will not satisfy the fundamental stationarity assumption for AR processes, and such non-stationarity could seriously degrade the overall prediction performance. In addition, an AR type model with high-order terms requires extensive off-line calibration effort for the autocorrelation coefficients, and the corresponding augmented state space also dramatically increases the on-line computational burden, especially for large-scale network applications.

Alternatively, without requiring prior demand information, a simple random walk model can be relatively easily built for short-term demand prediction, corresponding to an AR(1) model with autocorrelation coefficient of 1. Cremer and Keller (1981, 1987), as well as Chang and Wu (1994) applied the random walk model to predict dynamic OD flow split parameters, by directly extending the latest estimates as the future forecasts. Although this model is effective for a slowly changing process, it might not be rich enough to capture non-linear trends in time-varying OD flows, especially for medium term prediction. In order to describe the non-linearity in dynamic OD demand, Kang (1999) and Mahmassani et al. (1998) proposed a polynomial trend filter to estimate time-dependent OD flows on a general network. This model used historical information, instead of on-line traffic measurements, to calibrate demand evolution processes. The filter was applied to the OD demand values directly rather than to deviations of these from a priori best-estimates, unlike the approach devised in the present paper.

In a closely related problem area, approaches for off-line time-varying OD demand estimation have also been proposed in the past decade, mostly for operational planning applications. Using a simplified assignment model, Cascetta et al. (1993) presented a generalized least-squares framework for estimating time-varying demand in a network. A bi-level DTA-based time-varying demand estimation formulation was introduced by Tavana (2001) and further extended by Zhou et al. (2003) to utilize multi-day link counts. In contrast, little attention has been given to procedures for effectively and systematically updating the historical demand information for on-line estimation and prediction purposes. Ashok (1996) suggested several heuristic approaches to update the historical demand estimate with recent estimates obtained in real-time, but no optimal updating formulation was given.

In general, regular OD trip desires can be viewed as a repeated process with similar within-day dynamic patterns. By utilizing knowledge from household interview surveys and off-line estimation results on multiple days, historical demand data represents the a priori estimate of the regular OD demand pattern. In particular, in the context of long-range demand prediction, reliable historical data can serve as an informative source under normal conditions. On the other hand, it is necessary to recognize the possible existence of structural