Transportation Research Part C 81 (2017) 57–82

Contents lists available at ScienceDirect

Transportation Research Part C

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

On the variance of recurrent traffic flow for statistical traffic assignment

Wei Ma, Zhen (Sean) Qian *

Department of Civil and Environmental Engineering and Heinz College, Carnegie Mellon University, Pittsburgh, PA 15213, United States

ARTICLE INFO

Article history:
Received 3 June 2016
Received in revised form 13 May 2017
Accepted 14 May 2017

Keywords:
Statistical traffic assignment
Probability distribution
Demand variance
Route choice variance
Variance decomposition
Data driven

ABSTRACT

This paper generalizes and extends classical traffic assignment models to characterize the statistical features of Origin-Destination (O-D) demands, link/path flow and link/path costs, all of which vary from day to day. The generalized statistical traffic assignment (GESTA) model has a clear multi-level variance structure. Flow variance is analytically decomposed into three sources, O-D demands, route choices and measurement errors. Consequently, optimal decisions on roadway design, maintenance, operations and planning can be made using estimated probability distributions of link/path flow and system performance. The statistical equilibrium in GESTA is mathematically defined. Its multi-level statistical structure well fits large-scale data mining techniques. The embedded route choice model is consistent with the settings of O-D demands considering link costs that vary from day to day. We propose a Method of Successive Averages (MSA) based solution algorithm to solve for GESTA. Its convergence and computational complexity are analyzed. Three example networks including a large-scale network are solved to provide insights for decision making and to demonstrate computational efficiency.

1. Introduction

For decades, static traffic assignment has been playing a pivotal role in roadway planning and operations. Conceptually, the traffic assignment problem maps Origin-Destination (O-D) demands $q \in \mathbb{R}_+^{R\times S}$ to link flow $x \in \mathbb{R}_+^N$ or path flow $f \in \mathbb{R}_+^K$,

$$\phi : q \mapsto x, f \tag{1}$$

where R, S, N, K is the cardinality of the set of origin nodes, destination nodes, links and paths. ϕ captures individual travelers' route choice bahvior.

The classical traffic assignment problem overlooks the variation of demand and flow, a critical feature of recurrent traffic. The variation of recurrent traffic differs from the variation of non-recurrent traffic by nature. Under the non-recurrent flow, O-D demands and travel behavior are abnormally disrupted, subject to possibly significant change in infrastructure capacity or demand characteristics. Consequently, the supply and/or demand follow a completely different pattern from a regular day when traffic conditions are stabilized from day to day. The variation of recurrent traffic, on the other hand, is due to the random nature of demand and supply without substantial external interventions. The term "reliability" in the literature oftentimes measures how different is a non-recurrent pattern is away from a recurrent pattern. However, this paper studies the variation of recurrent patterns to understand the network reliability under repetitive traffic patterns.

* Corresponding author.
E-mail addresses: weima@cmu.edu (W. Ma), seanqian@cmu.edu (Zhen (Sean) Qian).

http://dx.doi.org/10.1016/j.trc.2017.05.009

The increasing quantity and quality of traffic data enable modeling more complicated statistical features of demand than the "mean". Instead of sampling traffic counts or travel times on sparse locations and several days of a year, prevailing technologies collect various kinds of data all years along in high spatio-temporal resolutions. For example, we acquire 5-min traffic counts data of all Tuesdays/Wednesdays/Thursdays in May 2015 from PeMS (http://pems.dot.ca.gov/) for two fixed sites on California SR-41. The daily time-varying counts on each day and their daily average are plot in Fig. 1. Without incidents, the counts of those two locations clearly exhibit recurrent patterns. The 5-min counts at the same time of day vary by up to 50 veh/5 min, but the overall trend seems repetitive from day to day. If one would make decisions on design, planning or operations of this road segment, using the mean is hardly sufficient. Knowing the variance of flow allows a better understanding of network performance and reliability. In fact, one would expect similar results from examining other performance metrics, such as travel time.

To understand the statistical features of recurrent flow patterns, we therefore model demands, link and path flows by multivariate random variables, Q, X, F, respectively. This paper generalizes and extends most existing statistical assignment models, namely ϕ, to map the probability distribution of demands to that of link and path flow.

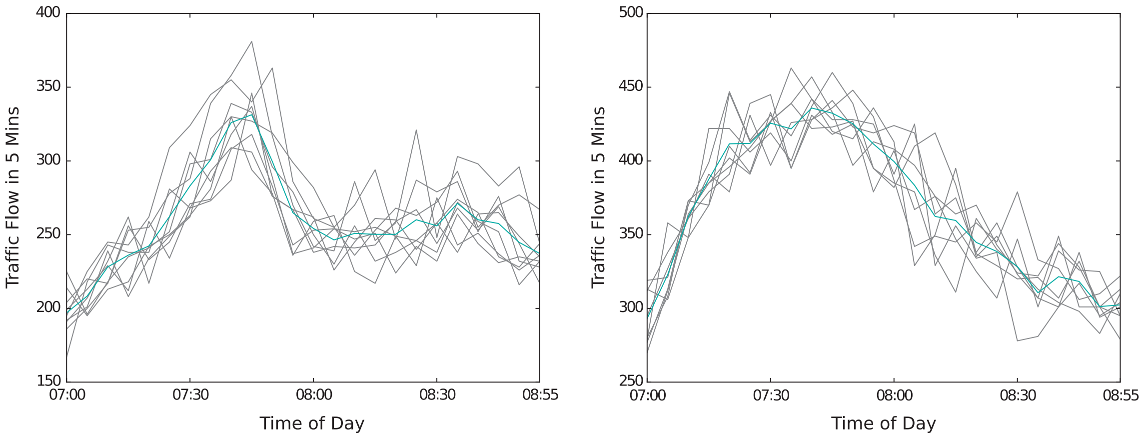
$$\phi : Q \mapsto X, F \tag{2}$$

where ϕ captures the stochastic route choice behavior. In the generalized statistical traffic assignment (GESTA), the variation of observed link and path flow stems from three sources, demand variation, route choice variation and measurement errors. Route choice variation is induced by both individual random choices and individual perception variation on travel costs. Demand or perception variation alone, sometimes referred to as "uncertainty", has been modeled and discussed in the literature to some extent. GESTA presents a statistical model incorporating all three variations coherently. The model is mathematically proven to be consistent with the underlying demand characteristics. It is a generalization of the classical deterministic traffic assignment models and other statistical assignment models embedding only a particular type of variation. The probability distributions of traffic flows and travel times, the outputs of GESTA, can be used for reliability-based network design (Yao et al., 2014) and robust routing (Wu and Nie, 2011). In addition, the covariance matrix of network flows and travel times is central to spatio-temporal travel time estimation and prediction (Min and Wynter, 2011; Nantes et al., 2016).

Another important feature of GESTA is the data-driven framework. The probability distributions and travel behavior models can be calibrated and validated using emerging field data collected years along. In the real world, we measure some elements of the flow vectors X, F on the daily basis. We can therefore use statistical inference and data mining models to estimate the probability distribution of demands Q and identify the best ϕ mapping to fit those data.

In addition to the traditional traffic counts data, data from various emerging sources can also be added to the GESTA framework to tune the model, thanks to its clear multi-level variance structure. For instance, speed measurements on individual links that are acquired from GPS years long can help build the probability distribution of travel costs (Ng et al., 2011; Rahmani et al., 2015; Coifman et al., 2016; Qin et al., 2017) and hence estimate the variance of route choices and O-D demands; On each day, the data we collect can be used to update the model following a general Bayesian Inference framework; The sampled vehicular trajectory data from GPS, Bluetooth (Bhaskar and Chung, 2013) and automated vehicle identification (AVI) (Mei et al., 2015) technology can help select route choice model and tune other parameters. In this paper, we define the assignment model ϕ, discuss its features and properties, while leaving the estimation process as future research.

Existing literature has examined each of the three types of variations that lead to flow variation, namely O-D demands, route choice and flow measurements. Each type of variation was modeled separately at different levels as shown in Fig. 2.



(a) One site on CA SR41 NB (b) One site on CA SR41 SB

Fig. 1. Daily time-varying traffic counts. Blue curve is the daily average. Each grey curve represents one day in May 2015. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)