Transportation Research Part C 36 (2013) 245–260

Contents lists available at ScienceDirect

Transportation Research Part C

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

A microstate spatial-inference model for network-traffic estimation

Adeyemi J. Fowe [a], Yupo Chan [b,*]

[a] Nokia (HERE Maps), 425 Randolph Street, Chicago, IL 60606, USA
[b] University of Arkansas at Little Rock, 2801 S. University Avenue, Little Rock, AR 72204, USA

ARTICLE INFO

Article history:
Received 21 July 2012
Received in revised form 24 August 2013
Accepted 24 August 2013

Keywords:
Spatial inference
Traffic estimation
Entropy maximization
Golden-section search

ABSTRACT

In an Advanced Traveler Information System (ATIS), sensors are often used to monitor and obtain traffic information on a real-time basis. Knowing that traffic sensors cover only a fraction of the road network, we investigate how to estimate traffic volumes on arcs that are not covered by sensors. By exploiting the spatial properties and the topology of a network, we derive a microstate model that can be used to estimate these traffic volumes. Based on entropy maximization, we present a microstate surrogate for competing techniques such as traffic assignment, and algebraic method or topological approach in estimating traffic flow. Being an entropy model, it also has advantage over these competing techniques in terms of the prerequisite information required to enable the model. Being a microstate rather than a steady-state model, it takes into account the fluctuation of traffic and it executes fast enough to allow real-time estimation of traffic flow. By covering the entire network flow this way with only a limited number of sensors, it will help in better driver routing decisions and traffic management tactics while being cognizant of today's budgetary constraints facing operating agencies. The algorithm has been tested successfully in Little Rock, Arkansas and in a controlled experiment with a randomly generated 100-node/522-arc grid network.

1. Introduction

Due to the extensive instrumentation requirement, complete coverage of traffic information in a road network requires high infrastructural and operational cost. This leaves a complete coverage of the network practically unachievable in agencies operating under a tight budget. Lack of sensor coverage for real-time traffic information, however, has impeded the deployment of many ATIS applications. This in turn has inspired research on optimal placement of sensors by Liu and Danczyk (2008), Ban et al. (2009), Kianfar and Edara, (2010), Gentili and Mirchandani, (2011), and Barcelo et al. (2013)just to name a few. Most recently, Ng (2012) has proposed a node-based model for the network sensor-location problem. He showed that a minimum subset of the network arcs and their locations existed for the estimation of the unobserved arc flows. Our paper goes beyond sensor placement. Starting with a limited number of sensors already in place, we present a way to estimate traffic volumes on arcs where traffic is unknown. Should the level of accuracy be acceptable, we are able to cover the entire network, possibly without increasing the number of sensors. This is achieved by exploiting the microstate relationship between vehicle movements, based on the topology of the network. In contrast to macrostate, a microstate description of traffic would identify each trip individually about where it is coming from and where it is heading, instead of just accounting for the trip ends (Golan, 1998). We will show that this is a viable approach for an ATIS, rather than the competing methodologies that estimate an aggregate, steady-state traffic volume for the network.

* Corresponding author.
E-mail addresses: james.fowe@nokia.com (A.J. Fowe), ychan@alum.MIT.edu (Y. Chan).

0968-090X/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.trc.2013.08.011

A classic way of estimating traffic is through traffic assignment. In spite of significant advances in the field, Zhang (2011) pointed out that the behavioral foundations of the models, as well as the practical implications of various behavioral assumptions, were rarely discussed. Zhang found out that there were significant discrepancies between arc-flow estimates arising from different behavioral foundations. Aside from this, traffic assignment is often inappropriate for real-time applications, when accurate origin–destination (O–D) demands are not available. Traffic assignment is also computational intensive, limiting its use when traffic has to be estimated "on the fly." Since the network configuration is given and can be represented by its link-path incidence matrix, some argue that the problem of network observability can be well solved by either the *algebraic method* (Hu et al., 2009) or *topological approach* (Castillo et al., 2008). A study of network observability provides useful information on which selected O–D-pair flows or arc flows are informative on other O–D-pairs and arc flows. Based on this, a practical application is to guide the placement of sensors to collect the most relevant traffic counts to infer a network-wide traffic pattern. However, the topological approach requires prior knowledge of O–D or path flows, or route-choice probabilities given by some traffic-assignment rules, and to formulate the flow-conservation equations. This constrains the applicability of the approach in real-world situations, particularly over real time, when such information is an estimate at best. On the other hand, the algebraic method, in relating the unobserved arc flows to the arc-path incidence matrix, may give rise to multiple solutions in terms of the set of *basis links* for sampling. In addition, there is a need for complete path enumeration—a computationally intensive undertaking. This discounts greatly its potential for real-time applications. All these considerations mean that we need to go beyond traffic assignments, algebraic and topological methods to estimate network traffic volumes.

Going beyond traffic estimation for a particular time slice, we wish to endow traffic operators with a look-ahead capability regarding what the traffic pattern is like a few minutes from now. Kamarianakis and Prastacos (2005), Min et al. (2011), and Zhang et al. (2012) are among the first to apply space-time models for traffic forecasting, focusing exclusively on arcs covered by sensors. Their work distinguishes from others by embracing the fundamental principle that there is a closer traffic relationship between nearby arcs than further-away arcs (Tobler, 1965)—a principle that will be exploited in full in this paper as well. Starting from a microstate level, we extend their work by estimating traffic volumes on unknown arcs. Furthermore, the current approach is based on more robust information. Instead of just projecting forward the arc volumes recorded by sensors, we estimate unknown arc traffic using *information theory* (or *entropy maximization*). The Principle of Maximum Entropy states that, subject to known or testable information, the probability distribution which best represents the current state of knowledge is the one with largest entropy (Jaynes, 1982). Most recently, Giffin and Caticha (2007) stated that Bayes' Rule and the Principle of Maximum Entropy were completely compatible, lending further credence to the entropy maximization paradigm. To facilitate a look-ahead capability, we estimate traffic one time-period at a time, including a future time-period. When performed repeatedly, the proposed approach has the potential for real-time ATIS applications.

Synthesizing the above literature, we wish to take the best features of all existing techniques while avoiding their worst features. In this regard, we are happy to report here an approach that distills the flow-continuity feature of traffic assignment, the network geometry of the topological approach, and the node-arc incidence information of the algebraic method into a single model. In this model, we consider the upstream and downstream neighbors as the *region of influence* of an arc in traffic estimation. Our model attempts to determine the most appropriate traffic estimates based on neighboring observations in this region. In an eventual aggregate statement of a travel pattern it is useful to be robust enough to accommodate as many detailed, microstate patterns as possible, where each microstate represents a stochastic instance of traffic flow. The most likely aggregate flow-pattern is assumed to be one with the greatest number of possible microstates. This embraces the tenet of entropy maximization, which attempts to capture all possible microstate patterns. In this context, entropy is interpreted in the spatial context as a measure of the frequency with which a traffic event occurs between neighboring arcs. We will estimate the unknown arc traffic by Shannon's seminal work in information theory (Shannon, 1948), and show that it is not only accurate, but computationally feasible for real-time applications. While travel demand matrices have been routinely estimated by entropy maximization, this is the first attempt to apply entropy maximization to arc traffic estimation—to the best of our knowledge.

2. Model formulation

Let $\overline{Q}^l$ represent a one-dimensional column vector of observed traffic flows $\bar{q}_l$ in the lth order for every arc i, $i = 1, 2, \ldots, m$ in a network. Let $\widetilde{W}^l$ be an $m \times m$ matrix of the lth-order neighbors with weight entries w_{ij}^l denoting the relationship between neighboring arcs i and j. When $l = 1$, only first-order (or immediate) neighbors j of subject arc i are considered. This says that only the set of arcs directly connected to arc i exert influence on the subject arc, dismissing arcs more than one "hop" away. Arcs with non-zero weight values in each row of $\widetilde{W}^l$ represent upstream arcs, while the arcs with non-zero weight values in a column represent downstream arcs. The upstream-arc j for a subject-arc i suggests that there may exist some arc-j traffic-flow into the arc i, while the downstream-arc j for arc i implies possible arc-i traffic-flow into arc j.

2.1. Key parameters

Let the set $\Gamma^l(i)$ stand for the *region of influence* for arc i, with elements k representing both the upstream and downstream arcs among the set of lth-order neighbors for arc i, $\{l\}$. i.e.,