Transportation Research Part B 91 (2016) 229–249

Contents lists available at ScienceDirect

Transportation Research Part B

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

Node modeling for congested urban road networks

Saif Eddin Jabari*

Division of Engineering, New York University Abu Dhabi, Saadiyat Island, P.O. Box 129188, Abu Dhabi, U.A.E.

ARTICLE INFO

Article history:
Received 9 July 2015
Revised 31 May 2016
Accepted 1 June 2016
Available online 3 June 2016

Keywords:
Urban networks
Kinematic waves
Node models
Traffic holding
Flow maximization
Invariance principle
Simultaneous movements and conflicts
Signalized intersections

ABSTRACT

First-order network flow models are coupled systems of differential equations which describe the build-up and dissipation of congestion along network road segments, known as link models. Models describing flows across network junctions, referred to as node models, play the role of the coupling between the link models and are responsible for capturing the propagation of traffic dynamics through the network. Node models are typically stated as optimization problems, so that the coupling between the link dynamics is not known explicitly. This renders network flow models analytically intractable. This paper examines the properties of node models for urban networks. Solutions to node models that are free of traffic holding, referred to as holding-free solutions, are formally defined and it is shown that flow maximization is only a sufficient condition for holding-free solutions. A simple greedy algorithm is shown to produce holding-free solutions while also respecting the invariance principle. Staging movements through nodes in a manner that prevents conflicting flows from proceeding through the nodes simultaneously is shown to simplify the node models considerably and promote unique solutions. The staging also models intersection capacities in a more realistic way by preventing unrealistically large flows when there is ample supply in the downstream and preventing artificial blocking when some of the downstream supplies are restricted.

1. Introduction

In congested urban networks, we now know that combating a traffic problem at one location might best be done by adjusting the controls at another location. But how to do so requires an understanding of how congestion propagates in a network. Over the past six decades, numerous modeling approaches have been developed that can successfully describe the spatio-temporal evolution of congestion along individual road segments (Aw and Rascle, 2000; Kerner and Konhäuser, 1994; Lighthill and Whitham, 1955; Paveri-Fontana, 1975; Prigogine and Andrews, 1960; Richards, 1956; Treiber et al., 1999; Zhang, 1998). Most of these theories were developed for roads of infinite length (or closed circuits) and draw their inspiration from models that describe fluid flow or gas dynamics.

It was not until the mid 1990's that traffic flow theories of road segments *with* boundaries began to emerge (Daganzo, 1995; Holden and Risebro, 1995; Lebacque, 1996). Today, we have good theories for describing/computing flows through simple junctions that connect one incoming road segment to two (or more) outgoing road segments, *diverge nodes*, and

* Corresponding author. Tel.: +97126284247.
E-mail address: sej7@nyu.edu
URL: http://engineering.nyu.edu/people/saif-eddin-jabari

http://dx.doi.org/10.1016/j.trb.2016.06.001

those that connect one outgoing segment to two (or more) incoming segments, *merge nodes* (Daganzo, 1995; Jin, 2010; 2015; Jin and Zhang, 2013; Ni and Leonard, 2005). These node models, combined with suitable models of flow in individual links, referred to as *link models*, can be used to effectively model/compute traffic dynamics in expressway networks. In such networks, the link models play a major role in representing traffic conditions since the network nodes tend to be located at fairly long distances from one another. The situation is quite different for urban/city networks, where the road segments are short and the nodes can have multiple inbound links and multiple outbound links. In such networks, the nodes play a more critical role in describing the dynamics in the network.

When modeling flow rates across link boundaries, certain conditions need to be satisfied to ensure that the boundary flows make sense physically. For example, flows into (out of) the upstream (downstream) boundary of a link should not exceed the supply (demand) at the boundary. Treatments in which boundary flows are appended as source terms to continuity equations without imposing these conditions can lead to ill-posed problems. In general, the conditions on the boundary flows are imposed in such a way as to ensure realism in addition to well posedness of the mathematical model. Some of the literature on general node models have focused on the propagation of kinematic waves through nodes (Coclite et al., 2005; Garavello and Piccoli, 2006; Holden and Risebro, 1995; Jin, 2012). Due to the finite speeds at which traffic information propagates through network links, it is reasonable to investigate the node models independently of the link dynamics as in Flötteröd and Rohde (2011); Lebacque (2005); Smits et al. (2015); Tampère et al. (2011). In general, all of these models satisfy supply and demand restrictions at link boundaries and mass-balance requirements on flows across the nodes. Most of them also consider restrictions on the distribution of demands to different outgoing links so as to reflect path/destination desires of drivers.

In many papers, flow maximization is imposed as an entropy condition aimed at restricting node model solution spaces. Flow maximization states that the total flows through the node at any time instant should be maximized. The principle is motivated by the physical requirement that flows through a network node should not involve *traffic holding*, which is the undesirable situation in which vehicles do not proceed through the node despite availability of supply in the downstream. In this paper, it is demonstrated that traffic holding and flow maximization are **not** one and the same except in special cases, namely, merge and diverge nodes. For general intersections, flow maximized solutions are only sufficient conditions for holding-free solutions.

One of the themes in the node modeling literature is the derivation of formulas for calculating node flows which satisfy the aforementioned restrictions, e.g., (Bliemer, 2007; Jin and Zhang, 2004). Such closed form expressions for node flows are desirable: the link dynamics are represented by systems of differential equations, the node flows represent a coupling between the differential equations. Hence, closed form solutions not only facilitate the analysis of the network dynamics, but can also help to dramatically reduce the computational effort for dynamic network loading tools. There exist many efficient techniques in the literature for computing the forward dynamics, e.g., (Gentile, 2010; Raadsen et al., 2015). However, a challenge that persists today is the analytical intractibility of the network models. For example, using the most efficient loading tools, the effect of a perturbation to the inputs can, in many circumstances, only be investigated by re-solving the entire problem. As a result, their complexity can be prohibitive for real-time applictions such as signal control, traffic assignment, and real-time estimation. An exception to this can be found in recent work that applies the link transmission model for link dynamics; the reader is referred to Himpe et al. (2016) and references therein.

Lebacque and Khoshyaran (2005) demonstrated that some solution approaches can result in unrealistic dynamics on the links, namely, the possibility of waves propagating in the wrong direction. To overcome this behavior, they proposed a principle, referred to as the *invariance principle*, which states that when restricted by the demands (supplies), the node flow solutions should be invariant to increases in the supplies (demands). Unlike the other conditions mentioned above, the undesirable behavior that results from violations of the invariance principle is dynamical in nature. Hence, invariance is a property of node model solutions and the algorithms used to compute them. Violations of the invariance principle tend to arise when there are multiple solutions and the algorithm picks out a "bad" solution. The more recent solution techniques in the literature tend to honor the invariance principle; some do so implicitly (Corthout et al., 2012; Flötteröd and Rohde, 2011; Tampère et al., 2011) and some have employed other mechanisms to ensure that waves propagate in the correct direction (Jin, 2012).

Despite all of the restrictions that have been imposed on node models, the majority of (realistic) models and their associated algorithms can still produce different solutions (Corthout et al., 2012; Smits et al., 2015) to the same problem. These differences can result in dramatically different network dynamics under the same intial and boundary conditions. This non-uniqueness can be attributed to the node flow problems being insufficiently constrained to produce unique solutions for any set of boundary conditions. Different approaches in the literature incorporate different physical insights in order to pick out one of the possibly many solutions that all honor the general requirements mentioned above. The physical insight incorporated in this paper to overcome this non-uniqueness issue is that conflicting movements should not be allowed to proceed through the node simultaneously. This couples with knowledge of the signal phasing scheme results in trivial movement priorities. This, in turn, promotes unique solutions.

The rest of this paper is organized as follows: Section 2 gives an overview of first-order link models and summarizes the principles that node models enforce. Section 3 formulates the demand, supply, and demand distribution conditions and Section 4 formally defines holding-free solutions and their relation to flow maximizing solutions. Section 5 presents the invariance principle and an invariant greedy algorithm that produces holding-free solutions. The non-simultaneity of