Transportation Research Part B 81 (2015) 848–862

Contents lists available at ScienceDirect

Transportation Research Part B

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

Bounding tandem queuing system performance with variational theory

Jia Li [a], H. Michael Zhang [b,*]

[a] Center for Transportation Research, University of Texas at Austin, TX 78701, United States
[b] Department of Civil and Environmental Engineering, University of California at Davis, CA 95616, United States

A R T I C L E I N F O

Article history:
Received 28 April 2015
Received in revised form 17 July 2015
Accepted 18 July 2015
Available online 21 August 2015

Keywords:
Queuing systems
Traffic flow
Variational solution

A B S T R A C T

Queuing models are often used for traffic analysis, but analytical results concerning a system of queues are rare, thanks to the interdependence between queues. In this paper, we present an analysis of queuing systems to obtain bounds of their performance without studying the details of individual queues. Queuing dynamics is formulated in continuous-time, subject to variations of demands and bottleneck capacities. Our analysis develops new techniques built on the closed-form solution to a generalized queuing model for a single bottleneck. Taking advantage of its variational structure, we derive the upper and lower bounds for the total queue length in a tandem bottleneck system and discuss its implication for the kinematic wave counterpart. Numerical experiments are conducted to demonstrate the appropriateness of the derived upper and lower bounds as approximations in a stochastic setting.

1. Introduction

Understanding and quantifying the interplay between traffic and bottlenecks is a central focus of traffic flow theory, and numerous analytical and computational models were developed for this purpose. These models capture different level of details and adopt different mathematical representations. Two most widely used macroscopic continuous-time traffic flow models are the (first-order) kinematic wave model (Lighthill and Whitham, 1955; Richards, 1956) and the point queue model (Vickrey, 1969; Nie and Zhang, 2002; Shen and Zhang, 2008), etc. Conventionally, hyperbolic conservation equations (Dafermos, 2005; Toro, 2009) and ordinary differential equations are respectively the primary mathematical tool to analyze the former and latter. Discrete time formulations of corresponding dynamics (e.g. Daganzo, 1994; Lebacque, 1996) are derivable from continuous-time models, using numerical tools such as the finite volume approximations (LeVeque, 1992). In this way, we get consistent models in analytical and numerical domains. Alternative formulations of these models exist, including the Lagrangian models and variational models. In general, these models were developed to study link-level flow dynamics.

System-level traffic queuing dynamics and properties are of particular interests in many regards, especially in the context of traffic control and operations. For instance, a better understanding of gridlock mechanism allows more effective control of arterial traffic in urban areas, which triggers the analysis of large-scale networks and associated gating and decentralized strategies in recent years (see e.g. Keyvan-Ekbatani et al., 2012; Haddad and Shraiber, 2014; Li and Zhang, 2014 and references therein). As another example, in operations of a corridor, it is desirable to consider collective dynamics of tandem

* Corresponding author. Tel.: +1 530 754 9203.
E-mail address: hmzhang@ucdavis.edu (H.M. Zhang).

http://dx.doi.org/10.1016/j.trb.2015.07.013
0191-2615/

queues along sequential bottlenecks rather than addressing them isolatedly, due to the their underlying interdependence. Nonetheless, albeit the desirability to consider queuing systems as an entity, knowledge on analytical properties of such systems are limited and difficult to obtain, which constitutes a major hurdle to developing a sound systemic theory and related applications. Among others, a central problem to conquer in this agenda is how to establish aggregate queuing properties based on given queuing dynamics at link and node level. In the last several years, we have seen increasing studies along this direction, which attempted to build the linkage between local (link- and node-level) queuing dynamics and analytical properties of networks. The properties that have been investigated mainly include existence of macroscopic fundamental diagram in traffic networks and related properties (Daganzo and Geroliminis, 2008; Daganzo et al., 2011), existence of solution to a dynamical queuing system (Jin, 2012), stability of diverge-merge networks (Jin, 2013), and continuity of path delay operators on general networks (Han and Friesz, 2012; Han et al., 2015). In these studies, link models employed range from simple link queue model to the more full-fledged LWR model, but their focuses are exclusively on qualitative characters. Quantitative measures, such as level of service (LoS) and throughput of network, were not addressed. Along another line of research, which tackles the so-called morning commute problem, exact arrival profiles are derivable, thus giving full quantitative characterizations of the system (see e.g. Newell, 1988; Kuwahara, 1990). Nonetheless, similar analysis can become demanding and difficult to extend if more than two bottlenecks are involved.

We aim to bridge this gap. As a first step, in this paper, we consider the problem of performance bounding. This problem looks for upper and lower bounds of system performance, when demand and bottleneck capacity data of a queuing system are given. In other systems, e.g. communication networks, this problem is well known, and the network calculus theory (Cruz, 1991; Le Boudec and Thiran, 2001) was developed to tackle it. Some models in the network calculus, e.g. the leaky bucket model, are similar to the point queue model that we discuss below. But as will be clear from below, the approach developed in this paper is different from network calculus. Central to the network calculus are the concepts of service curve and max-plus algebra, based on which link-level dynamics are concatenated. In contrast, our analysis exploits a variational property of point queue dynamics in this paper. Performance bounds are established through making use of this special structure. Moreover, relations between the point queue model and kinematic wave model are examined, in order to extend the analysis concerning point bottleneck models.

We take the following steps. We start from a simple yet flexible queuing model, called generalized queuing model (Li and Zhang, 2015). We present this model and review its major properties. Among others, the variational property of this model is most interesting and relevant, which allows its solution be expressed in closed form, even when demand and bottleneck capacity are time-dependent and discontinuous. Furthermore, the closed form of this solution is mapping of nothing else but the so-called demand surplus function $h(t)$. Upon noting this link-level property, we demonstrate that the notion of demand surplus can be extended to a more general setting, i.e. a route consisting of tandem bottlenecks. Such an extension underpins upper and lower bounds for system performance, e.g. total queue length. Having bounding the performance of a general tandem queue system analytically, we consider approximation of first-order kinematic wave model with the generalized queuing model, in order to understand the impact of different notions of queue on resulted performance bounding. We derive their difference, in terms of queue length, when identical initial and boundary data are used. Along with these steps, we also prove the tightness of given bounds, and conduct numerical experiments to verify our analytical findings.

The remainder of this paper is organized as follows. In Section 2, we introduce the technical background of this study. In Section 3, we present the major result of this paper, performance bounds derived from the generalized queuing model, extension of the results to kinematic wave models. In Section 4, we conduct numerical experiments and verify derived analytical results. In Section 5, a discussion is presented concerning the modeling issues arising in previous sections. In Section 6, we summarize the findings and discuss future works.

2. Preliminaries

In this section, we provide an overview of queuing models, queuing system analysis, and the variational theory of traffic flow. We start with considering traffic dynamics on a link. The following notation will be used:

$t : t \in \mathbb{R}^+$, time
$x(t)$: queue length, i.e. number of vehicles on concerned link, at t
$u(t)$: inflow rate at t
$v(t)$: outflow rate at t
$U(t)$: cumulative inflow, i.e. $U(t) = \int_0^t u(s)ds$
$V(t)$: cumulative outflow, i.e. $V(t) = \int_0^t v(s)ds$
$C(t)$: time-dependent link discharging capacity
S: link storage capacity
$\tau(t)$: link traversal time for car entering at t
τ_0: link traverse time at free flow