Transportation Research Part C 35 (2013) 286–304

Contents lists available at SciVerse ScienceDirect

Transportation Research Part C

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

Reliability-based stochastic transit assignment with capacity constraints: Formulation and solution method

W.Y. Szeto [a],*, Yu Jiang [a], K.I. Wong [b], Muthu Solayappan [c]

[a] Department of Civil Engineering, The University of Hong Kong, Pokfulam Road, Hong Kong, PR China
[b] Department of Transportation Technology and Management, National Chiao Tung University, 1001 Ta Hsueh Road, Hsinchu 30010, Taiwan
[c] Department of Industrial and Systems Engineering, National University of Singapore, 1, Engineering Drive 2, Singapore 117576, Singapore

ARTICLE INFO

Article history:
Received 13 August 2010
Received in revised form 9 April 2011
Accepted 1 September 2011

Keywords:
Transit assignment
Reliability-based user equilibrium
Linear Complementarity Problem
Column generation
Capacity constraint
Chance constraint

ABSTRACT

This paper proposes a Linear Complementarity Problem (LCP) formulation for the reliability-based stochastic transit assignment problem with capacity constraints and non-additive link costs, where in-vehicle travel times and waiting times are uncertain. The capacity constraints are developed via the notions of effective capacity and chance constraints. An equivalent route-based linear program (LP) for the proposed problem is formulated to determine the patronage of each line section, critical links, critical service frequencies, unmet demand and the network capacity, which considers the risk-aversive behavior of travelers. A solution method is developed, utilizing the K-shortest path algorithm, the column generation technique, and the revised simplex method, to solve the proposed LP with guaranteed finite convergence. Numerical experiments are also set up to illustrate the properties of the problem and the application of the proposed model for reliability analysis.

1. Introduction

The planning of urban transit services relies on the use of transit assignment models for predicting the way in which transit travelers choose routes from their origins to their destinations (Cepeda et al., 2006). As a result, transit assignment models have received considerable attentions over the last two decades. Earliest models (e.g., Dial, 1967; Fearnside and Draper, 1971; Le Clercq, 1972; Nguyen and Pallottino, 1988; De Cea and Fernández, 1989; Spiess and Florian, 1989) assumed that all transit lines have unlimited capacity to accommodate any amount of transit demand, and in-vehicle congestion effects over transit networks were not considered. This limitation does not permit an accurate representation of transit networks with high congestion levels due to insufficient capacity of services (De Cea and Fernández, 1993).

To deal with capacity-related congestion, two approaches have been developed in the literature: the congestion cost function approach and the capacity constraint approach. The congestion cost function approach (e.g., De Cea and Fernández, 1993; Wu et al., 1994; Bouzaïene-Ayari et al., 1995; Cominetti and Correa, 2001) adopts an unbounded increasing convex function to model the effect of in-vehicle congestion due to insufficient capacity on waiting time. The concept of effective frequency is often used together with this approach. Earlier models such as De Cea and Fernández (1993), Wu et al. (1994), and Bouzaïene-Ayari et al. (1995) allowed the development of convergent solution algorithms under some monotonicity or uniqueness conditions, but these models allow the flow on a link to be greater than its capacity, which is unrealistic. The model of Cominetti and Correa (2001) rectified this problem by introducing queue theoretic models, but the algorithm was not necessary convergent.

---

* Corresponding author.
E-mail address: ceszeto@hku.hk (W.Y. Szeto).

0968-090X/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.trc.2011.09.001

The capacity constraint approach (e.g., Lam et al., 1999, 2002; Nguyen et al., 2001; Yin et al., 2003; Poon et al., 2003, 2004; Cepeda et al., 2006; Yang and Lam, 2006; Tian et al., 2007a; Teklu, 2008) traditionally incorporates capacity constraints in transit assignment models to disallow the flow on a link to be greater than the corresponding capacity. In the frequency-based method, the excess flow is usually assigned either to a pedestrian arc with unlimited capacity connecting to the destination directly (e.g., Cepeda et al., 2006) or to a failure node (e.g., Kurauchi et al., 2003; Schmöcker et al., 2008, 2011). In the schedule-based method, the passengers who fail to board a vehicle are kept waiting until they can board the next arriving vehicle with sufficient capacity (e.g., Poon et al., 2004; Hamdouch and Lawphongpanich, 2008). Recently, seat capacity has even been considered in the models of Tian et al. (2007b), Sumalee et al. (2009), Leurent and Liu (2009), Schmöcker et al. (2011) and Leurent (2011). This capacity constraint approach is more realistic. However, the resultant models are usually solved by the method of successive averages, which can guarantee convergence under specific conditions. However, these conditions may not be fulfilled by the models.

Other than the congestion effect, most existing transit assignment models do not consider network stochasticity. In fact, due to supply side uncertainty, in-vehicle travel times and waiting times, especially for buses and mini-buses, are highly uncertain. Moreover, most existing transit assignment models do not consider the influence of the variances of travel times on route choice. Indeed, empirical studies like Abdel-Aty et al. (1997) and Jackson and Jucker (1982) point out that travel time variability, which is one of the measures of travel time reliability, plays a major role in influencing the trip makers' route choice behavior. The trip makers select their routes by considering the trade-off between travel time (or cost) and its uncertainty (Yin et al., 2004). It is essential to capture this realistic travel behavior into the transit modeling framework. To our best knowledge, only Yang and Lam (2006), Li et al. (2008, 2009), Szeto and Solayappan (2009a,b), Sumalee et al. (2011), and Szeto et al. (2011b) considered this behavior in their models. These studies adopted the congestion cost function approach to model the effect of congestion. However, no convergent solution techniques have been developed for their models. Moreover, the capacity constraint approach has not been considered simultaneously with the uncertainties of in-vehicle travel time and waiting time.

This paper proposes a stochastic transit assignment problem with capacity constraints that takes into account the variabilities of in-vehicle travel times and waiting times. In our model, both in-vehicle travel times and waiting times are modeled as random variables. Their means and variances are incorporated in the modeling framework through the concepts of effective travel cost and reliability-based user equilibrium such that the network uncertainty and risk-taking behavior (including risk-aversive behavior) of travelers can be captured. The capacity constraints are developed by the chance constraints, which are formulated based on the notion of effective capacity introduced in this paper.

The proposed problem is formulated as a Linear Complementarity Problem (LCP), which is later reformulated as a route-based linear programming problem. This reformulation allows us to determine critical links, critical service frequencies, met and unmet demand, and the transit network capacity. The network capacity considers the risk-averse behavior of travelers, and it is not required to have any reference origin–destination (OD) matrix for determining the capacity. This capacity contrasts to the classical maximum network capacity (e.g., Ahuja et al., 1993) that only considers one OD pair and ignores the route choice behavior of travelers. Our proposed network capacity also contrasts to the definition of network capacity that relies on a reference OD matrix (e.g., Asakura, 1992; Wong and Yang, 1997; Morlok and Riddle, 1999; Liu and Deng, 2010). Moreover, our proposed capacity differs from those considering route choice behavior of travelers (Asakura, 1992; Ahuja et al., 1993; Akamatsu and Miyawaki, 1995; Wong and Yang, 1997; Chen et al., 1999; Yang et al., 2000; Ziyou and Yifan, 2002; Ge et al., 2003; Kasikitwiwat and Chen, 2005; Lee et al., 2006) in the sense that the risk averse behavior of travelers is taken into account.

This paper proposes a new convergent solution approach, utilizing the column generation technique, the K-shortest path algorithm, and the revised simplex method, to solve the proposed linear programming problem. Numerical examples are also set up to illustrate the properties of the proposed problem and the application of the proposed model for reliability analysis.

The rest of the paper is organized as follows. Section 2 describes the general representation of transit networks, the particular network used for the illustration, and the assumptions that are relevant to our work. Section 3 elucidates the various cost components of the travel cost. Section 4 describes the idea of effective travel costs, and Section 5 postulates the problem formulation. The proposed solution method is detailed in Section 6, whereas Section 7 discusses the computational results. Finally, Section 8 provides concluding remarks and identifies directions for future research.

2. Network representation and assumptions

Following the method in Lam et al. (1999), a general transit network is represented by a set of transit lines and a set of stations (nodes) where passengers can board, alight or transfer. For the simplicity of presentation, a walk link is also represented as a transit line with high frequency. To handle the common line problem and reduce the number of paths handled by the model proposed in Section 3, the line-node transit network is transformed into a link-node network.

To facilitate the discussion, a transit network adopted for the purpose of analysis is shown in Fig. 1, which approximates the existing bus network in Singapore. JE, BL, HF, TP, and EU denote five nodes, namely Jurong East, Boon Lay, Harbour Front, Toa Payoh, and Eunos, respectively. Fig. 2 shows the same network coded by links (or route sections). In this example, there are four origin–destination (OD) pairs (with HF being a transfer hub) connected by ten routes, and the ten routes are serviced