Transportation Research Part C 20 (2012) 16–33

Contents lists available at ScienceDirect

Transportation Research Part C

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

A schedule-based assignment model with explicit capacity constraints for congested transit networks

Agostino Nuzzolo, Umberto Crisalli *, Luca Rosati

"Tor Vergata" University of Rome, Department of Civil Engineering, via del Politecnico 1, 00133 Rome, Italy

A R T I C L E I N F O

Article history:
Received 24 September 2009
Received in revised form 29 January 2011
Accepted 24 February 2011

Keywords:
Transit networks
Schedule-based
Dynamic assignment
Congestion
Capacity constraints

A B S T R A C T

This paper presents a schedule-based dynamic assignment model for transit networks, which takes into account congestion through explicit vehicle capacity constraints. The core of this assignment model is the use of a joint choice model for departure times, stops and runs that defines a space-time path in which users decide to leave at a given time, to access the network at a given stop and to board a given run to reach their destination. The assignment model is defined through a dynamic process approach in which the within-day network loading procedure allocates users on each transit run according to user choice and to the residual capacity of vehicles arriving at stops. The proposed model, albeit general, is specified for frequent users, who constitute a particularly congestion-sensitive class of users. Finally, an application to a real-size test network (part of the Naples transit network in southern Italy) is illustrated in order to test the proposed approach and show the ability of the modelling framework to assess congestion effects on transit networks.

1. Introduction

In the context of transit networks, congestion usually refers to the decrease in on-board comfort as the on-board load increases up to a maximum threshold (vehicle capacity), after which users are not allowed to board (oversaturation) and have to wait for next arriving vehicle. In the literature, the Transit Capacity and Quality of Service Manual (Kittelson & Associates et al., 2003) describes this problem as passenger capacity, initially studied using different approaches based on service frequency (Nguyen and Pallottino, 1988; Spiess and Florian, 1989). Such approaches comprise the effective frequency approach (De Cea and Fernàndez, 1993; Cominetti and Correa, 2001; Cepeda et al., 2006), the penalty-function approach (Lam et al., 1999), the failure-to-board approach (Bell, 2003; Kurauchi et al., 2003; Bell and Schmocker, 2005; Schmocker et al., 2008) and that of ordered preference (Hamdouch et al., 2004).

As in the frequency-based approach transit services are represented by lines, single vehicles (runs) are not explicitly considered. This results in an approximation in calculating single vehicle loads: the higher the variation in the demand profile in peak hours, the greater is the approximation. Such approximation also arises in the case of irregular arrival of vehicles at stops, which is why peaks of boarding users at stops cannot be explicitly considered.

Hence the use of a schedule-based approach (Tong and Richardson, 1984; Nuzzolo and Russo, 1994; Hickman and Wilson, 1995) seems to be more suited to examining congestion in high-frequency transit networks, as each run with its vehicle capacity can be considered, and a generic temporal demand profile can also be taken into account.

A general classification of schedule-based models should consider the disaggregate or aggregate approach for both supply and demand. On the supply side, the disaggregate approach considers performance (e.g. travel time) for each vehicle

* Corresponding author. Tel./fax: +39 06 72597053.
E-mail address: crisalli@ing.uniroma2.it (U. Crisalli).

0968-090X/$ - see front matter
doi:10.1016/j.trc.2011.02.007

depending on the characteristics of the vehicle concerned and its explicit interaction with others on the same link; the aggregate approach (even if the schedule-based approach explicitly considers single vehicles) takes into account average performance, calculated on the basis of all vehicles travelling on the same link in the same period. On the demand side, the disaggregate approach explicitly considers single users with their own characteristics, for which single user behaviour and choices are simulated; by contrast, the aggregate approach considers groups of users with common characteristics, for which user group choices are carried out by considering the whole user group choice probabilities of each alternative (e.g. origin–destination pair and departure time). According to this general classification, schedule-based models specified in a disaggregate way for both supply and demand can be classified as schedule-based microsimulation models, while the others may be grouped as schedule-based macrosimulation models.

This paper investigates the schedule-based macrosimulation approach. In this context we present an up-to-date review of models dealing with congestion by using explicit vehicle capacity constraints. For further details on the schedule-based microsimulation approach the reader can refer to Morgan et al. (2004) and Shalaby and Wahba (2009).

Transit networks usually provide different access stops and lines/runs for users travelling on a given origin–destination pair. In this case, user arrival at stops is not related to the schedule due to high service frequency. However, user behaviour is greatly affected by congestion, for which a specific choice model has to be used.

Passenger congestion on public transport can be modelled by using an implicit or explicit approach. The implicit approach is derived from road network modelling, for which strictly non-decreasing continuous link-cost functions with respect to link flows are defined for particular links. The main drawback of this approach is the approximation in assessing the disutility for boarding users at stops with respect to users already on board. In other words, the effect of congestion is the same for all users and is mainly represented by a discomfort increase. Among papers using this approach the reader can refer to Crisalli (1999), Nielsen (2004), Wong and Tong (1999), Nuzzolo et al. (1999, 2001, 2003), and Tong and Wong (2004). In order to overcome such an approximation, congestion may be modelled through an explicit approach: by introducing vehicle capacity constraints, users board the arriving run according to its residual capacity. This means that assignment models are able to capture changes in user behaviour when oversaturation and failure-to-board events occur and can produce more precise outputs. This approach can be considered at the forefront of the research in this field, even if theoretical properties of this modelling approach are under investigation. Among recent papers that use this approach, mention should be made of Poon et al. (2004), Nuzzolo et al. (2007), Hamdouch and Lawphongpanich (2008), Papola et al. (2009) and Sumalee et al. (2009). A concise description of their features is reported below.

Poon et al. (2004) develop a method that explicitly models the available capacity of each vehicle at each boarding station and the queuing time for each passenger, assuming First-Come–First-Served (FCFS) queuing discipline at the station. They propose a user equilibrium assignment approach for a congested schedule-based transit network, which can be considered assignment model updating of the modelling framework presented in Tong and Wong (2004). This paper focuses on the route choice problems of travellers in a schedule-based transit network, whereas the other choice dimensions, such as departure time, and entry, transfer and exit stations, are not considered. In particular, route choice is based on the assumption that all users have full predictive information (which has been gained through past experience) about present and future network conditions and select paths that minimise a generalized cost function made of in-vehicle time, waiting time, walking time, as well as line change penalty. It also considers passengers queuing at platforms based on an FCFS rule to board transit vehicles, which have a fixed capacity and operate precisely (i.e. no delays occur in service functioning) according to the timetable. Route choice is obtained by a simulation-based approach and stochastic user equilibrium assignment is carried out by using a Method of Successive Averages (MSA) algorithm: incremental assignment of two different classes of users (those re-routed due to congestion and otherwise) ensures algorithm convergence. The authors propose this model as a tool to evaluate the performance of transit networks subject to peak period loading; it can be used to simulate the performance of an existing transit system operating with pre-announced schedules under variable passenger demand conditions, or to evaluate the effects of changes in schedules, lines or passenger demand on system performance.

Hamdouch and Lawphongpanich (2008) present a user equilibrium transit assignment model that explicitly considers transit schedules and individual vehicle capacities. For each origin–destination (O/D) pair passengers are divided into groups according to the desired arrival time interval of the passengers in the group, even if groups with desired departure times may also be distinguished. This model assumes that a passenger uses travel strategies conceptually similar to hyperpaths (Nguyen and Pallottino, 1988) by providing, at each transit station and each point in time, an ordered list of transit lines he/she prefers to use to continue his/her journey. This strategy can be adaptive over time and can be graphically represented as sub-graphs on a time-expanded network (Ahuja et al., 1993). This supply model describes both passengers and vehicles over the planned time horizon as a set of equally spaced points in time representing access to, egress from, and runs of various transit lines in the system. The strategy sub-graph is characterised by specific link penalties for both access links (to levy a penalty for early departure, or consider the latest possible time a passenger can depart from his or her destination and arrive within the desired interval) and egress links (to account for lost opportunities associated with arrivals outside the desired interval). Passengers behave by minimising their expected travel cost made of in-vehicle travel time, fare and additional opportunity costs associated with having to leave home early and/or arriving at the destination outside the desired arrival interval. The authors assume that a passenger can provide several user-preference sets to be used when he/she reaches a station at different times. Strategies can include many crowded transit lines and may result in long waiting times because the passengers may not be able to board their preferred lines. When loading a transit vehicle at a station, on-board passengers continuing to the next station remain on the vehicle and waiting passengers are loaded in an FCFS or a random