Transportation Research Part C 36 (2013) 212–230

Contents lists available at ScienceDirect

Transportation Research Part C

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

Optimizing urban rail timetable under time-dependent demand and oversaturated conditions

Huimin Niu [a,*], Xuesong Zhou [b]

[a] School of Traffic and Transportation, Lanzhou Jiaotong University, Lanzhou 730070, China
[b] School of Sustainable Engineering and the Built Environment, Arizona State University, Tempe, AZ 85287, United States

ARTICLE INFO

Article history:
Received 7 March 2013
Received in revised form 29 August 2013
Accepted 29 August 2013

Keywords:
Urban rail line
Train timetable
Time-dependent demand
Oversaturated condition
Transit service optimization
Genetic algorithm

ABSTRACT

This article focuses on optimizing a passenger train timetable in a heavily congested urban rail corridor. When peak-hour demand temporally exceeds the maximum loading capacity of a train, passengers may not be able to board the next arrival train, and they may be forced to wait in queues for the following trains. A binary integer programming model incorporated with passenger loading and departure events is constructed to provide a theoretic description for the problem under consideration. Based on time-dependent, origin-to-destination trip records from an automatic fare collection system, a nonlinear optimization model is developed to solve the problem on practically sized corridors, subject to the available train-unit fleet. The latest arrival time of boarded passengers is introduced to analytically calculate effective passenger loading time periods and the resulting time-dependent waiting times under dynamic demand conditions. A by-product of the model is the passenger assignment with strict capacity constraints under oversaturated conditions. Using cumulative input–output diagrams, we present a local improvement algorithm to find optimal timetables for individual station cases. A genetic algorithm is developed to solve the multi-station problem through a special binary coding method that indicates a train departure or cancellation at every possible time point. The effectiveness of the proposed model and algorithm are evaluated using a real-world data set.

1. Introduction

As a key component of public transit operations and management, the passenger train timetable design problem on an urban rail transit line aims to determine the arrival and departure times for each train at each station. This complex scheduling task requires a systematic consideration of time-varying and stochastic passenger demand patterns, available train-unit fleet and many practical regulations.

A fundamentally important element for designing public transit timetables is how to estimate passenger waiting times. Using continuum fluid-flow models to approximate the passenger loading patterns on a single origin–destination route, a number of early studies, such as Newell (1971), Osana and Newell (1972), and Hurdle (1973), calculated the total passenger waiting time by various analytical formulas in order to derive optimal dispatching policies. Daganzo (1997) presented a cumulative flow count-based approach to capture transient queues under overflow conditions for a general scheduled transportation system, and he also described a simplified simulation process to derive the time-dependent passengers delay on a corridor with different origin-to-destination pairs.

* Corresponding author. Tel./fax: +86 931 4938015.
E-mail addresses: hmniu@mail.lzjtu.cn (H. Niu), xzhou74@asu.edu (X. Zhou).

0968-090X/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.trc.2013.08.016

In general, an even schedule with a constant headway between consecutive vehicles can reduce total waiting time when the passenger arrival pattern at stations follows some particular probability distributions, such as uniform and Poisson distributions. LeBlanc (1988) introduced a method to optimize frequencies of each transit line in a network using a modal-split assignment programming model with distinct transit routes to capture the effects of increases or decreases in individual transit line frequencies. Banks (1990) studied the transit scheduling problem to determine net-benefit maximizing headways for multi-route transit systems which are subject to constraints on vehicle capacity, subsidy and fleet size. Ceder (2001) proposed a graphical method to improve the synchronization of vehicle departure times with passenger demand while minimizing the number of required buses. Under time-varying demand conditions, however, an even headway timetable may lead to longer passenger waiting time during oversaturated periods due to some passengers may not be able to board the next arrival train, or ineffective train capacity utilization during undersaturated periods due to the under-capacity boarded passengers.

A multi-phase, semi-regular timetable, which divides a day into several time blocks and applies the even vehicle-departing interval for each period, may somehow help to accommodate peak-hour demand while maintaining a certain level of service for passengers boarding at non-peak hours. Guihaire and Hao (2008) presented a global review of the crucial strategic and tactical steps of transit planning, and also discussed the scheduling problem with phase-regular for a transit corridor. Ceder (2009) provided a comprehensive modeling framework for determining vehicle departure times with either even headways or even average loads, with a special focus on smoothing the transitions between time periods. These studies provide useful methods for optimizing frequency for a particular time period, while a unified framework is critically needed for scheduling methods that can consider uneven headways and time-dependent demand patterns.

A number of recent studies have put more attention to developing optimization models for constructing periodic transit timetables and synchronized schedules. Odijk (1996) built a model consisting of periodic time window constraints associated with arrival and departure times and further described a constraint generation-based solution algorithm. Carey and Crawford (2007) designed a series of heuristics for finding and resolving trains conflicts so as to satisfy various operational constraints and objectives. Zhou and Zhong (2005), Zhou and Zhong, 2006 formulated train scheduling models which consider segment and station headway capacities as limited resources, and developed algorithms to minimize both the expected passenger waiting times and total train travel times. Goverde (2007) described a railway timetable stability measure by using a max-plus system theory and analyzed train delay propagation processes. Focusing on reducing passenger waiting time at stops and transfers, Liebchen (2008) adapted a periodic event-scheduling approach and a well-established graph model to optimize the Berlin subway timetable. Wong et al. (2008) concentrated on the synchronization between the different lines of an urban rail transit network to minimize passengers' transfer times. Caimi et al. (2011) addressed the problem of generating conflict-free train schedules on a microscopic model of the railway infrastructure, where each considered train path corresponds to a vertex, and edges represent pairwise conflicts so that a conflict-free schedule corresponds to a maximum independent set. Corman et al. (2012) developed a bi-objective model of minimizing train delays and missed connections in order to provide a set of feasible non-dominated schedules. From the system effectiveness and efficiency point of view, Dundar and Sahin (2013) recently developed a decision support tool to resolve inter-train conflicts for the single-track railway train re-scheduling problem.

Within an optimization modeling framework, it is possible to further consider many other transit operational strategies and practical constraints. For example, Eberlein et al. (1998) studied a real-time deadheading problem in transit operations control. Haghani and Banihashemi (2002) developed a mathematical programming model and a heuristic approach to solve large-scale transit vehicle scheduling problems with routing time constraints. Fu et al. (2003) examined a bus operations control strategy in which stop skipping is applied to every other bus dispatched from the terminal. Recently, Leiva et al. (2010) developed a non-linear optimization model for designing limited-stop services, and the goal is to minimize the waiting time, in-vehicle travel time and operator cost of an urban bus corridor.

In general, existing studies have not fully integrated the consideration of the time-dependent origin-to-destination (OD) passenger demands in the train timetabling process, partly due to the difficulty of obtaining the dynamic OD demand table. Thanks to the continuing deployment of Intelligent Transportation Systems (ITS) technologies, automatic fare collection systems (AFCS) have been widely used to obtain individual trip records for many urban transit systems (e.g., Zhao et al., 2007; Farzin, 2008). These ridership measurements provide a data-rich environment for developing more efficient transit services, especially under oversaturated conditions. Note that in most freeway traffic assignment or transit scheduling studies, time-varying demand matrices use five-min or 15-min departure time intervals to characterize aggregated traveler behavioral choices. In comparison, available AFCS data offer a much finer temporal resolution (e.g., second-by-second) for greatly enhancing congestion modeling capabilities.

As discussed above, a limited number of studies have been devoted to the transit-scheduling problem under heavily congested conditions and limited train/vehicle fleet availabilities, where some of the passengers must wait for an extended period of time and then board the next or the third arriving trains. To design service-oriented, transit timetables that utilize emerging time-dependent, origin-to-destination ridership data, this paper is intended to address the following modeling issues: (1) analytical formulations for computing time-varying and extended passenger waiting time under oversaturated conditions; (2) optimal solution algorithms for considering both passenger delay and limited fleet availability; and (3) heuristic algorithms for finding robust solutions for medium or large-scale data sets with multi-station samples.

The remainder of this paper is organized as follows. The overall problem statement and underlying assumptions for passenger boarding behavior are first presented in Section 2. Section 3 constructs a binary programming model with cumulative