Optimizing Urban Rail Timetable under Time-dependent Demand and Oversaturated Conditions

Huimin NIU
School of Traffic and Transportation
Lanzhou Jiaotong University, Lanzhou, 730070, China
Email: hmniu@mail.lzjtu.cn
Tel: +86-0931-4938015
Fax: +86-0931-4938015
(Corresponding Author)

Xuesong ZHOU
Department of Civil and Environmental Engineering
University of Utah, Salt Lake City, UT 84112-0561
Email: zhou@eng.utah.edu
Tel: +1-801-585-6590
Fax: +1-801-585-5477

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 unit, passengers may not be able to board the next arrival train, and they may be forced to wait in queues for the following trains. Based on time-dependent, origin-to-destination trip records from an automatic fare collection system, a nonlinear optimization model is developed to capture the overall passenger delay, subject to resource constraints associated with a limited number of electronic multiple units. A first-in-first-out queuing assumption is introduced to analytically calculate effective passenger loading time periods and the resulting time-dependent waiting times for given dynamic and stochastic demand patterns. Based on cumulative input-output diagrams, two novel solution algorithms, namely, local improvement and dynamic programming methods, are presented 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.

Keywords: Urban rail line; train timetable; time-dependent demand; transit service optimization; automatic fare collection system

1. Introduction
As a key component of public transit operations and management, the passenger train timetable design on an urban rail transit line must 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 locomotive and train unit resources, as well as many practical regulations.

Designing public transit timetables has received much attention in recent decades. For the problem of constructing periodic timetables of public rail systems under unsaturated conditions, many studies concentrate on how to model train delays and resolve conflicts, with limited focus on passenger waiting times. 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. Zhou and Zhong (2005) embedded a random incidence theorem in a train-scheduling model to minimize both the expected waiting times for high-speed passenger trains and the total travel times of high-speed and medium-speed trains. 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.

In a more specific problem of transit timetabling, a majority of the existing studies have focused on bus scheduling with even headways. In general, an even schedule with a constant train-departing interval can reduce total waiting time when the passenger arrival pattern at stations follows a uniform distribution. However, the evenly spacing train timetable, under time-varying and oversaturated conditions, may lead to longer passenger waiting time during peak travel hours, or ineffective fleet capacity utilization during non-peak or normal hours. A multi-phase, semi-regular timetable, which divides a day into several time blocks and applies the even-train 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. 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), constructed various analytical formulas to derive optimal dispatching policies that could minimize the total passenger waiting time and the cost associated with vehicle operation. Eberlein et al. (1998) considered a real-time deadheading problem in transit operations control. Ceder (2001) proposed a graphic method to improve the synchronization of vehicle departure times with passenger demand while minimizing the number of required buses. Haghani and Banilhashemi (2002) developed a mathematical programming model and a heuristic approach to solve large-scale transit vehicle scheduling problems. Fu et al. (2003) studied a bus operations control strategy in which stop skipping was applied to every other bus dispatched from the terminal. 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. Recently, Leiva et al. (2010) developed a non-linear optimization model for designing limited-stop services that aim to minimize the waiting time, in-vehicle travel time and operator cost of an urban bus corridor.

It is critically important to estimate the passenger waiting times at stations when designing a practical and efficient transit timetable under oversaturated conditions. Without considering the passenger loading capacities of buses, the random incidence theorem described by Larson and Odoni (1981) can be used to calculate expected delay under random passenger arrival times, given bus arrival-time interval distributions. Daganzo (1997) presented a cumulative flow, counts-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.

In the general field of transportation systems modeling, a wide variety of studies have been devoted to optimizing signal- timing plans for a transportation network with oversaturated nodes or links, for example, an optimal control formulation by D’Ans and Gazis (1976), a mixed integer programming formulation by Gartner et al. (1975), and genetic algorithms by Girianna and Benekohal (2004). The above studies also adapted cumulative flow, counts-based analytical models for single or multiple intersections. An embedded or external traffic simulator, such as TRANSYT (Wong, 1996), is typically required to describe the complex delay propagation and queue spillback in a hybrid simulation-optimization framework.

To estimate time-varying transit service patterns and optimize passenger train timetables, transit planners need to collect dynamic origin-to-destination demand information. 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 traffic 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 small 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. Many fundamental modeling issues need to be addressed to design service-oriented, transit timetables that utilize emerging time-dependent, origin-to-destination ridership data. These issues include the following requirements: (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-day 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 proposes a detailed analytical queueing model for train timetabling on a heavily congested urban rail line. Section 4 presents two novel solution algorithms, namely, local improvement and dynamic programming methods, to find an optimal timetable for single station cases. The genetic algorithm, with a special binary coding method, is described in Section 5, followed by a numerical example with real-world data.

2. Problem statement
2.1. Notation

This study considers train operations over one day on a double-direction urban rail line with \( N \) stations, as shown in Fig.1. The stops are marked as \( 1, 2, \ldots, 2N \), sequentially, where stations 1 and \( N \) denote the start terminal and the return terminal, respectively. Geographically, stop 1 corresponds to stop \( 2N \), 2 corresponds to \( 2N-1 \), \ldots, \( N \) corresponds to \( N+1 \), respectively, and each pair of stops are in opposite directions but with the same station names. Without loss of generality, our model views them as different stations. Each train departs from station 1, makes a u-turn at station \( N/N+1 \), and then comes back to station \( 2N/1 \) for preparing a new departure. All trains are assumed to be scheduled according to the same speed between two consecutive stations and the same dwell time at each station. As a result, the optimization model needs to focus on how to determine the departure times of trains at the start terminal. The detailed arrival and departure times at the other stations can be accordingly generated for timetabling applications.
The urban rail transit system can be modeled as a hybrid process with both discrete and continuous events, where trains are scheduled in terms of discrete time stamps and passengers arrive at stations semi-continuously according to a certain time-dependent pattern. More specifically, the arriving times of passengers are determined by their own individual trip purposes (e.g., home-to-work or home-to-shopping) and are affected by many other external random factors, such as freeway conditions and passenger walking times from home or office to terminals.

A first-in-first-out (FIFO) queuing assumption is introduced in this paper to analytically calculate effective passenger loading time periods and the resulting time-dependent waiting times for given dynamic and stochastic demand patterns. To introduce this FIFO queuing principle, this paper assumes that all passengers arrive at stations individually and that no more than two passengers arrive at a station at the same time. Such an assumption can be ensured by decomposing the study time horizon into very small intervals, so that there is only one passenger arriving at a station at a given time stamp. If the raw data contain records with simultaneous arrivals, we can simply randomly assign their arrival sequences without loss of generality. To consider the time-dependent demand matrix along the rail corridor over a one-day horizon, the number of passengers that arrive at station $u$ at time $t$ toward station $v$ is denoted as $P^{u,v}(t)$. We set $P^{u,v}(t) = 0$, with station indices $u \in \{1,2,\ldots,N\}$ and $v \in \{N+1,N+2,\ldots,2N\}$, by assuming that all boarded passengers will get off at the return terminal.

The following notation and parameters are used throughout this paper.

**Index:**
- $u,v$: index of stations;
- $j$: index of trains;
- $t$: index of times.

**Given Input:**
- $P^{u,v}(t)$: trip demand table, i.e., number of passengers who arrive at station $u$ at time $t$ heading to station $v$;
- $m$: number of supplied electronic multiple units (EMU) during the study period (e.g., one day);
- $c$: EMU capacity, that is, the maximum number of passengers who could be accommodated by a set of EMU (e.g., 900 persons);
- $h$: pre-specified safety time interval between two consecutive trains at the same station (e.g., two min);
- $s^u$: dwell time at station $u$;
- $r^u$: running (task processing) time from station $u$ to station $u+1$ ($u = 1,2,\ldots,2N-1$).

**Timetabling Variables:**
- $A^u_j$: arrival time of train $j$ at station $u$;
- $D^u_j$: departure time of train $j$ at station $u$;
- $K$: total number of trains departing from the starting terminal over one day.

**Queueing variables:**
- $B^u_j$: number of boarded passengers arriving at station $u$ toward station $v$ for train $j$;
- $X^u_j$: number of remaining passengers in train $j$ after the train departs from station $u$;
- $O^u_j$: number of alighted passengers in train $j$ after the train arrives at station $u$, that is, outflow count;
- $I^u_j$: number of boarded passengers in train $j$ during the train stay at station $u$; that is, inflow count;
- $W^u_j$: number of passengers waiting for train $j$ during the train stay at station $u$;
- $L^u_j$: the latest time of boarded passengers arriving at station $u$ for train $j$; that is, the passengers could get on train $j$ if they arrive at station $u$ before and at this time stamp, and it is also the ending time of the effective loading period for train $j$ at station $u$. 

![Diagram](image-url)  
Fig. 1. The representation of an urban transit rail line.
2.2 Queueing-based Representation

The cumulative input-output diagram shown in Fig. 2 illustrates the relationship between passenger demand and capacity supply, as well as the time-varying congestion phenomena. There are four consecutive trains, \( j, j+1, j+2 \) and \( j+3 \), departing from station \( u \). All passengers waiting for trains \( j \) and \( j+3 \) could board the corresponding trains, but some of the passengers waiting for trains \( j+1 \) and \( j+2 \), respectively, must wait for the second stage next train because the vehicle capacity \( c \) cannot accommodate all passengers, where \( c_{\text{over}} \) is used to denote the number of passengers who cannot get on the current train. To capture the effective passenger loading time period for each train, this paper introduces a new parameter, \( L_j^u \), as the latest time of boarded passengers arriving at station \( u \) for train \( j \), that is, the ending time of the effective loading period. Obviously, for trains \( j \) and \( j+3 \), their departure times correspond to the latest arrival times. However, in the case of trains \( j+1 \) and \( j+2 \), the effective loading time periods are earlier than their departure times. In particular, under the illustrated heavily congested conditions, train \( j+1 \) can only accommodate passengers between times \( D_j^u \) and \( L_{j+1}^u \), and the effective loading time period for train \( j+2 \) covers from \( L_{j+1}^u \) to \( L_{j+2}^u \).

![Fig. 2. Cumulative input-output diagram at station \( u \).](image)

When the effective time window for each train is given from Fig. 2, Fig. 3 further shows how individual passengers with different destinations can board different trains. In this specific example, the arriving times for the first and sixth passengers with destinations \( u+1 \) and \( u+2 \) correspond to the latest times \( L_{j+1}^u \) and \( L_{j+2}^u \) for trains \( j+1 \) and \( j+2 \).

![Fig. 3. Illustration of the passenger arriving process at station \( u \).](image)
Assuming that the leftover capacity for train \( j+2 \) at station \( u \) is five, only five passengers are sequentially picked up at their arrival times between the time intervals \( t \in \{ L^u_{j+1}, L^u_{j+2}\} \). However, the seventh and eighth passengers, with destinations \( u+3 \) and \( u+1 \), where \( t \in \{ L^u_{j+2}, D^u_{j+2}\} \), must wait for the next train.

In the cumulative input-output diagram for a single station case shown in Fig. 2, the cumulative arrival flow curve is externally given by the demand profile \( P^{u,v}(t) \), while the cumulative departure flow count curve can be computed according to train departure time sequences \( D^u_j \). Under both undersaturated and oversaturated conditions, we must derive the value of \( L^u_j \) and other parameters recursively, and load passengers at station \( u \) to train \( j \) if the passenger’s arrival time \( t \) satisfies \( P^{u,v}(t)=1 \) and \( t \in \{ L^u_{j-1}, L^u_j\} \). In the case of multiple stations, we need to sequentially calculate the effective passenger loading period \( (L^u_{j-1}, L^u_j) \), starting from the first train \( j=1 \) and the start terminal \( u=1 \). It is worth noting that the number of supplied EMU \( (m) \) is a predetermined input for this study. However, the actual total number of trains departing from the starting terminal \( (K) \) is a decision variable to be optimized as the by-product of final arrival/departure time sequences \( A^u_j \) and \( D^u_j \) for all trains at different stations.

3. Optimization Model

This section presents all constraints and objective functions of the urban rail-timetabling problem under consideration, and our focus is on how to capture the essential constraints for multiple stations with different origin-to-destination passenger flows.

3.1. Queueing system-related constraints

Timetable constraints

Given dwell time \( s^u \) and section running time \( r^u \), the arrival and departure times should satisfy the following equations for each train at each station.

\[
A^u_j = D^u_{j-1} + r^u \quad \text{for} \ u=2, 3, \ldots, 2N; \ j=1, 2, \ldots, K
\]

\[
D^u_j = A^u_j + s^u \quad \text{for} \ u=2, 3, \ldots, 2N; \ j=1, 2, \ldots, K.
\]

For each train at the starting terminal, the safety interval \( h \) should be ensured as follows.

\[
D^1_j - D^1_i \geq h \quad \text{for} \ j=1, 2, \ldots, K-1.
\]

Number of boarded or alighted passengers

For each effective passing loading time window \( (L^u_{j-1}, L^u_j) \), \( B^{u,v}_j \) for the number of passengers from one station \( u \) to another station \( v \), the available boarding a given train \( j \) can be calculated as Eq. (4).

\[
B^{u,v}_j = \int_{t \in [L^u_{j-1}, L^u_j]} P^{u,v}(t) dt \quad \text{for} \ u=1, 2, \ldots, 2N-1; \ v= u+1, u+2, \ldots, 2N; \ j=1, 2, \ldots, K.
\]

This calculation leads to the number of alighted passengers at a given station \( v \) and the number of boarded passengers at station \( u \).

\[
O^v_j = \sum_{u=1}^{2N} B^{u,v}_j \quad \text{for} \ u=2, 3, \ldots, 2N; \ j=1, 2, \ldots, K
\]

\[
I^u_j = \sum_{v=u+1}^{2N} B^{u,v}_j \quad \text{for} \ u=1, 2, \ldots, 2N-1; \ j=1, 2, \ldots, K.
\]

As the boundary condition, parameters \( O^1_j \) and \( I^{2N}_j \) equal zero for each train \( j \). We can calculate the number of waiting passengers at station \( u \) for train \( j \) as shown below.

\[
W^u_j = \sum_{v=u+1}^{2N} \int_{t \in [L^u_{j-1}, L^u_j]} P^{u,v}(t) dt \quad \text{for} \ u=1, 2, 3, \ldots, 2N; \ j=1, 2, \ldots, K.
\]
The waiting passengers at a station will get on the current train until the number of onboard passengers exceeds its capacity. After a train $j$ departs from a station, the number of remaining passengers in the train is as follows:

$$ X_j^u = \min\{c, X_j^{u-1} - O_j^u + W_j^u\} \quad \text{for } u=1, 2, \ldots, 2N; j=1, 2, \ldots, K. $$

(8)

According to the illustration of the train operation and passenger queue, as shown in Fig. 4, the flow conservation constraint is then represented as follows:

$$ X_j^u = X_j^{u-1} + I_j^u - O_j^u \quad \text{for } u=2, 3, \ldots, 2N; j=1, 2, \ldots, K. $$

(9)

Eqs. (8) and (9) show that if $L_j^u < D_j^u$, the passengers staying at station $u$ numbering $W_j^u - I_j^u$ cannot get on train $j$ and have to wait for the next train $j + 1$.

![Diagram of train operations and passenger queuing process](image)

Fig. 4. Illustration of train operations and the passenger queuing process.

Using a space-time graph, Fig. 4 shows the flow balance of inflow count, outflow count, and effective loading count for two adjacent trains. An arrival or departure node represents that trains arrive at or depart from the station. A node of the latest arrival time (corresponding to $L_j^u$) separates the time coordinate into several mutually exclusive intervals for each effective passenger-loading period for train $j$. A service arc represents the train traversing from a departure node to an arrival node in the time graph, Fig. 4. An alighting or boarding arc represents the passengers getting off or onto the train. A dwell arc represents the dwelling service process of a train at the station. An alighting or boarding arc represents the passengers getting off or on the train. A queue arc represents the passengers waiting at the station.

### 3.2. Determining effective passenger loading time periods

After some passengers leave the train while it is at the station, the residual capacity of the train determines how many passengers can be further accommodated by the train, which, in turn, decides the calculation of the boundary points of the effective time-loading period. Recall that, this paper assumes that all passengers arrive alone and get on the train according to the FIFO rule.

It is obvious that the maximum number of (in-flow) passengers at station $u$ who can be accommodated by train $j$ is $c - X_j^{u-1} + O_j^u$. The last time stamp of boarded passengers arriving at station $u$ for train $j$ is thus set to $L_j^u = D_j^u$, if there is enough capacity for all passengers arriving before the departure time of train $j$, that is, $W_j^u \leq c - X_j^{u-1} + O_j^u$. However, only a portion of the waiting passengers can board train $j$ if $W_j^u > c - X_j^{u-1} + O_j^u$. The latest arrival time is obtained as follows:

$$ L_j^u = \max\{h | \sum_{v=1}^{2N} \int_{t_{v+1}(L_j^u, h)}^{t_v} P^{v, u}(t) dt \leq c - X_j^{u-1} + O_j^u\} \quad \text{for } u=1, 2, \ldots, 2N; j=1, 2, \ldots, K. $$

(10)

where $X_j^{u-1} = 0$ for $u=1$ is satisfied as the boundary condition.
It is noted that the number of alighted passengers \( O_j^u \) and number of boarded passengers \( X_j^{u-1} \) have already been computed by the recursive formula (10) used to calculate the latest time \( L_j^u \). Thus, the latest time of boarded passengers arriving at station \( u \) for train \( j, L_j^u \), can be also presented as follows:

\[
L_j^u = \min \left\{ D_j^u, \max \{ h \left| \sum_{t=0}^{2N_k} \int_{t}^{t+h} P_t \, dt \leq c - X_j^{u-1} + O_j^u \right. \} \right\} \quad \text{for } u=1, 2, \ldots, 2N; j=1, 2, \ldots, K.
\] (11)

It should be remarked that Eq. (11) is valid under both unsaturated and oversaturated conditions.

In the single station case shown in Fig. 2, there is no passenger leaving the train. Therefore, we need only to calculate the cumulative departure flow curve based on train departure times and available constant capacity. That is, \( O_j^1 = 0 \) and \( X_j^0 = 0 \) for any train \( j \). Thus, we need to load arriving-passenger demand \( P_{u=1,v=2}^j(t) \) up to time \( L_j^1 \). When considering a much more complex case involving multiple stations, as described in this section, we need to precisely take into count the passengers remaining in and leaving a train at each station. Those two sets of system status variables, \( X_j^{u-1} \) and \( O_j^u \), are, in turn, determined by the multi-origin/destination incoming passenger demand profile \( P_t \) and the corresponding effective passenger loading-time periods for different trains.

### 3.3. Resource constraints and objective function

After presenting the queueing system-related constraints, we now discuss the remaining constraints related to the limited resources.

#### The available number of EMU

For an urban rail transit system with high-frequency train schedules, it is very important to have enough EMU available for dispatching at any moment. This paper considers that the EMU are the major resource constraint in our timetable design problem. The following constraint is obtained according to the departure order of trains at the start terminal.

\[
D_j^1 \geq D_{j-m}^1 + T_{j-m} \quad \text{for } j = m+1, m+2, \ldots, K,
\] (12)

where \( T_{j-m} = T_0 + A_{j-m}^{2N} - D_{j-m}^1 \) denotes the cycle time that consists of the running time, dwell time and preparing time. \( T_0 \) represents the time that trains need to be prepared at the start terminal. Inequality (12) can ensure that the departure time for a train at the start terminal is not earlier than the allowable departure time, which is supplied by EMU, or ensure that each train is matched by a prepared EMU.

#### Passenger loading requirements

The transit operation company requires a certain threshold of passenger-loading standards for most of trains to ensure its benefit, which can be quantified by the overall number of on-board passengers to the total utilized capacity. At the same time, few trains are allowed to operate with lower load rates than required for passenger welfare. This loading requirement constraint can be expressed as follows, taking into account that the trains run in a no-load mode within sections \([N, N+1]\) and \([2N, 1]\).

\[
\frac{1}{(2N - 2) \times c} \times \sum_{u=1}^{2N-1} X_j^u \geq \alpha \quad \text{for } j = 1, 2, \ldots, K_0; \quad K_0 < K.
\] (13)

where \( \alpha \) denotes the minimum passenger load rate and \( K_0 \) denotes the number of trains with higher passenger load levels.

#### Train operating time period

This study considers that the first and last trains depart from the start terminal at the initial and end moments over an operation day. For example, the first train departs at 6:00 a.m., and the last leaves at 21:00 p.m. from the start terminal. We also assume that passengers waiting for the last train could board the corresponding run, considering the lower demands associated with the end period, or we assume that the train supply could meet the total passenger demands. The operating time period constraint is then represented as follows:

\[
W_k^u \leq c - X_k^{u-1} + O_k^u \quad \text{for } u=1, 2, \ldots, 2N-1.
\] (14)
**Objective function**

The objective is to minimize the total waiting times of passengers at the stations over one day, shown in Eq. (15). At station \( u \), the waiting time of a passenger is associated with \( D_{j}^{u} - t \) if the demand \( P^{u,v} (t) = 1 \) and the arriving time is \( t \in (L_{j-1}^{u}, L_{j}^{u}] \).

\[
\min Z = \sum_{j=1}^{K} \sum_{u=1}^{2N-1} \sum_{v=u+1}^{2N} \int P^{u,v} (t) \times (D_{j}^{u} - t) dt ,
\]

(15)

where \( L_{0}^{u} \) denotes the time for the first passenger arriving at station \( u \) during the day.

**4. Optimization algorithm for single station**

**4.1 Local improvement algorithm**

By precisely accounting for the total waiting times, available resources and practical regulations, the following section describes a comprehensive model to capture the time-dependent, transient queues and compute the corresponding passenger waiting times in an urban rail corridor with oversaturated stops. To further provide insights into the train timetabling problem under oversaturated conditions, we first demonstrate the process of adjustment of the departure times of trains for a simple case with a single station.

Two propositions based on an approximate numerical gradient are presented to better understand the system-wide impact of adjusting a train schedule for an overflow queue. For notation convenience, we ignore the superscripts of \( u \) and \( v \) for the key parameter, such as \( P^{u,v} (t) \) and \( D_{j}^{u} \), in this single-station case.

![Diagram](image)

**Fig. 5.** Gradient information when moving the departure time of a train forward before an oversaturated time period.

**Proposition 1.** The overall net benefit is obtained by \((E_j - (D_j + \rho))\)\( \times \int_{t \in (D_j, D_{j+1})} P(t) dt - X_{j} \times \rho \) when moving forward \( \rho \) units of departure time \( D_j \) before the end time \( E_j \) of an oversaturated time period for train \( j \) subject to

\[
D_{j+1} - (D_j + \rho) \geq h
\]

(16)

\[
D_{j+m} \geq (D_j + \rho) + T_j
\]

(17)

\[
X_{j} + \int_{t \in (D_j, D_{j+1})} P(t) dt \leq c
\]

(18)

**Proof.** Firstly, the safety time interval for the moved trains is ensured by constraint (16) and the available EMU are satisfied by constraint (17).

For an undersaturated train \( j \) shown in Fig. 5, we consider the numerical gradient when moving forward \( \rho \) units at departure time \( D_j \). The total delay is increased by the area of block \( R_0 \) with \( X_j \times \rho \), and decreased by the area of block \( R_1 \) with
$H(R_i) \times W(R_i)$ for passengers waiting for the following train $j+1$, where height $H(R_i) = \int_{t_0(D_j,D_j+p]} P(t)dt$ and width

$W(R_i) = D_{j+1} - (D_j + \rho)$. In Fig. 5, trains $j+1$ and $j+2$ are currently oversaturated, so train $j+1$ now can accommodate an additional $H(R_i)$ number of passengers (who originally were waiting for train $j+2$), illustrated by the height of block $R_2$. That is, if the additional passengers and the remaining passengers $H(R_i) + X_j$ are less than the capacity and if the next trains are saturated, then the total waiting time during this overflow time period can also be decreased by the same height for $D_{j+2} - D_{j+1}$ and $D_{j+3} - D_{j+2}$. Three blocks, $R_1$, $R_2$ and $R_3$, in Fig. 5 have the same height, and their time boundaries are connected until the ending time of oversaturated period $E_j$. Therefore, the overall net benefit is $H(R_i) \times [E_j - (D_j + \rho)] - X_j \times \rho$, and the proposition is proved by the above discussions.

Proposition 1 means that the adjustment is beneficial when moving the departure time of an under-capacity train forward, if this train is scheduled before the oversaturated time period and the safety interval and EMU supply are ensured. It is worth notice that the value $\rho$ should be taken as $\max\{\rho \mid X_j + \int_{t_0(D_j,D_j+p]} P(t)dt \leq c\}$ by proposition 1. It is well known that the train departure times at the start terminal are usually taken as an integer multiple of half or one min. However, the passengers can arrive at the terminal at any moment. Thus, the value $\rho$ should be revised as $\min\{\rho \mid X_j + \int_{t_0(D_j,D_j+p]} P(t)dt \geq c\}$ to minimize the overall passenger delay at the start terminal.

Fig. 6. Gradient information when moving the departure time of a train backward during an oversaturated time period.

**Proposition 2.** The overall net benefit is obtained by $c \times \rho$ when moving backward $\rho$ units of departure time $D_j$ for train $j$ during an oversaturated time period subject to

\begin{align*}
(D_j - \rho) - D_{j-1} & \geq h \\
D_j - \rho & \geq D_{j-m} + T_{j-m} \\
D_j - \rho & \geq L_j.
\end{align*}
Proof: Similarly, the safety time interval for the moved trains is ensured by constraint (19), and the available EMU are satisfied by constraint (20). Constraint (21) ensures that the corresponding effective passenger loading-time period \( L_{j+1}, L_j \) is not changed when moving backward \( \rho \) units of departure time \( D_j \) for train \( j \) during an oversaturated time period. It is obvious that the waiting times can be reduced by the area of block \( R \) with \( c \times \rho \). □

According to Proposition 2, the adjustment is beneficial when moving the departure time of a train backward if this train is scheduled during an oversaturated time period, given that the safety interval and EMU supply are ensured.

Based on these two proposed propositions, a local improvement algorithm can be presented below to find the optimal solution to this simple case. For each step, we calculate the numerical gradient of each train, select a departure time with the maximum net benefit while satisfying safety time interval and vehicle fleet constraints, and, accordingly, move forward or backward or stop at the current location, until no over improvement can be made.

Algorithm 1. Local improvement algorithm

Step 1: (Initialization)
- Generate a regular timetable based on the safety time interval and vehicle fleet size constraints;

Step 2: (Iterative computation)
- For train index \( j = 1 \) to \( j = K \)
  - If train \( j \) is under capacity, then
    - Move forward \( \rho_{opt} \) units of departure time \( D_j \), subject to the constraints (16) and (17),
      
      \[
      \rho_{opt} = \max\{ \rho \mid X_j + \int_{t_{D,j}}^{t_{D,j}+\rho} P(t)dt \leq c \}
      \]
    
  - Otherwise, if train \( j \) is oversaturated, then
    - Move backward \( \rho_{opt} \) units of departure time \( D_j \), subject to the constraints (19) - (21),
      
      \[
      \rho_{opt} = \min\{ \rho \mid X_j + \int_{t_{D,j}}^{t_{D,j}+\rho} P(t)dt \geq c \}
      \]
  
Endfor//\( j \)

4.2. Dynamic programming algorithm

We then design a dynamic programming algorithm, based on the recursive computation, to derive the optimal solution to the simple case. Consider train \( j \) departing at \( t_0 \) and the remaining queue size \( Q[j, t_0] \). For train \( j + 1 \) , departing at \( t_1 \) , and the remaining queue size \( Q[j + 1, t_1] \), we can analytically calculate the total arrival flow \( \int_{t_0}^{t_1} P(t)dt \) , and the residual queue size after train \( j + 1 \) leaves as \( Q[j + 1, t_1] = Q[j, t_0] + \int_{t_0}^{t_1} P(t)dt - c \).
Denote the total waiting time when train \( j \) departs at time \( t_0 \) with a residual queue size \( r_0 \) as \( F(j, t_0, r_0) \), where \( r_0 = Q[j, t_0] \). Using Fig. 7, we can derive the marginal waiting time increase as \[ \int_{t_0}^{t_1} P(t)(t_1 - t) \, dt + Q[j, t_0] \times (t_1 - t_0) \] from \( F(j, t_0, r_0) \) to \( F(j + 1, t_1, r_1) \), where \( r_1 = Q[j + 1, t_1] \). This result is possible because the area of block \( R_1 \) is \( Q[j, t_0] \cdot (t_1 - t_0) \) and the area of block \( R_2 \) is \[ \int_{t_0}^{t_1} P(t)(t_1 - t) \, dt. \]

By using the above recursive formulation, the following dynamic programming algorithm aims to systemically break the entire complex problem into much simpler calculations and sub-problems, through which we can determine the total waiting time increase for a pair of partial solutions denoted by \((j, t_0, Q[j, t_0])\) and \((j+1, t_1, Q[j+1, t_1])\). The Bellman equation is then established as follows:

\[
F(j + 1, t_1, r_1) = \min \{ F(j, t_0, r_0) + \int_{t_0}^{t_1} P(t)(t_1 - t) \, dt + Q[j, t_0] \times (t_1 - t_0) \} \text{ for all feasible } t_0 \text{ of train } j \text{ and for all feasible residual queue size } r_0 = 0, 1, \ldots, R_{\text{max}}, \text{ where } R_{\text{max}} \text{ is the maximum queue size.}
\]

In this case, the state variables of this dynamic programming problem are indexed by two dimensions. The first one is the time index \( t \) and the second one is the size of the residual queue \( r \) after the train leaves at time \( t \). We define the earliest and latest departure times of train \( j \) as \( ts(j) \) and \( te(j) \), respectively. The previous train optimal departure time corresponding to the current train \( j+1 \) at time \( t_1 \) with a residual queue size \( r_1 \) is \( pt(j+1, t_1, r_1) \). The detailed solution algorithm can be described as follows:

**Algorithm 2**: Dynamic programming algorithm

**Step 1**: (Initialization)
- Initialize the cost label for total waiting time as \( F(j, t, r) = \infty \), and set \( F(j=0, t=0, r=0) = 0 \).
- For train \( j = 1 \),
  - For feasible departure time \( t_0 \),
    - calculate the total waiting time up to time \( t_0 \) as \( F(j, t_0, r_0) = \int_{t_0}^{t_1} P(t)(t - t_0) \, dt \)
    where \( r_0 = \max \{ 0, \int_{t_0}^{t_1} P(t) \, dt - c \} \).

**Step 2**: (Recursive computation)
- For train \( j+1 = 2, \ldots, K \)
  - For each feasible departure time \( t_1 \) of train \( j+1 \), \( t_1 = ts(j+1), \ldots, te(j+1). \)
    - For each feasible departure time \( t_0 \) of the previous train \( j \), \( t_0 = ts(j), \ldots, \min\{t_1 - 1, te(j)\}. \)
For each feasible residual queue size, \( r_0 = Q[j, t_0] = 0, 1, \ldots, R_{\text{max}}, \) at time \( t_0 \) for train \( j \).

Step 2.1: Calculate the residual queue size for train \( j+1 \).
\[
   r_i = Q[j+1, t_i] = \max\{0, Q[j, t_0] + \int_{t_0}^{t_i} P(t) \, dt - c\}
\]

Step 2.2: Calculate the total cumulative waiting time
\[
   F_{\text{temp}}(j+1, t_i) = F(j, t_0, r_0) + \int_{t_0}^{t_i} P(t)(t_i - t) \, dt + Q[j, t_0] \times (t_i - t_0)
\]

Step 2.3: Update
If \( F_{\text{temp}}(j+1, t_i, r_i) < F(j+1, t_i, r_i) \), then
\[
   F(j+1, t_i, r_i) = F_{\text{temp}}(j+1, t_i, r_i)
\]
Record processor \( pd(j+1, t_i, r_i) = t_0 \)
Endif
Endfor // \( r_0 \)
Endfor // \( t_0 \)
Endfor // \( t_1 \)
Endfor // \( j+1 \).

Step 3: (Retrieve optimal solution)
Find optimal solution as the minimum value of \( F[K, T^*, r=0] \), where \( T^* \) belongs to a set of the desirable departure times for the last train train. Used the recorded information at predecessor \( pd(j+1, t_i, r_i) \) to back trace the optimal departure time of the previous train \( j \), recursively, for \( j=K-1, K-2, \ldots, 1 \).

4.3. A simple case
We present a very simple case with a single station and 5 trains to demonstrate the above algorithms. The passenger demands are recorded at two-min intervals, with the number of passengers for each interval and the arriving moment for each passenger at the station shown in Table 1.

<table>
<thead>
<tr>
<th>Record Number</th>
<th>Begin Time</th>
<th>End Time</th>
<th>Volume</th>
<th>Passenger Arrival Timestamp</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>2</td>
<td>4</td>
<td>0.5/1.0/1.5/2.0</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>4</td>
<td>6</td>
<td>2.3/2.7/3.0/3.3/3.7/4.0</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>6</td>
<td>7</td>
<td>4.2/4.5/4.8/5.1/5.4/5.7/6.0</td>
</tr>
<tr>
<td>4</td>
<td>6</td>
<td>8</td>
<td>8</td>
<td>6.2/6.5/6.7/7.0/7.2/7.5/7.7/8.0</td>
</tr>
<tr>
<td>6</td>
<td>10</td>
<td>12</td>
<td>10</td>
<td>10.2/10.4/10.6/10.8/11.0/11.2/11.4/11.6/11.8/12.0</td>
</tr>
<tr>
<td>7</td>
<td>12</td>
<td>14</td>
<td>7</td>
<td>12.2/12.5/12.8/13.1/13.4/13.7/14.0</td>
</tr>
<tr>
<td>8</td>
<td>14</td>
<td>16</td>
<td>6</td>
<td>14.3/14.7/15.0/15.3/15.7/16.0</td>
</tr>
<tr>
<td>9</td>
<td>16</td>
<td>18</td>
<td>6</td>
<td>16.3/16.7/17.0/17.3/17.7/18.0</td>
</tr>
<tr>
<td>10</td>
<td>18</td>
<td>20</td>
<td>4</td>
<td>18.5/19.0/19.5/20.0</td>
</tr>
</tbody>
</table>

The basic settings consider a vehicle capacity of 14 and a safety interval of two mins. We also assume that the last train is scheduled at the end of the study horizon. Using the local improvement algorithm, we can obtain an optimal timetable by adjusting the even-spacing timetable shown in Fig. 8, where an abbreviation of PAX is used to denote the number of passengers loaded in each train.
Secondly, we solve the problem using the dynamic programming algorithm, and the same solution is achieved. The overall waiting time for the regular timetable is 201 min, leaving out four passengers. However, all passengers board the trains, and the overall waiting delay is 154 min for the optimal timetable.

Theoretically, when there are multiple stations under consideration, we can first identify the bottleneck among the corridors and apply the constructed solution algorithm for the single-station case until no system-wide improvement or waiting time saving can be made. However, the complexity will be increased dramatically due to the interaction of passenger boarding and alighting at the stations. Another effective algorithm has to be designed in the next section.

5. Genetic algorithm for multiple stations

5.1. Representation of chromosome

The model proposed in this paper is a complex, nonlinear programming problem, which is especially useful when dealing with numerical integrals and multiple stations. This problem could be difficult to solve with conventional, gradient-based methods or commercial optimization solvers. The genetic algorithm, a search method based on natural selection, is, therefore, adopted in this paper. The main idea is to calculate the total passenger delay and other parameters, simultaneously, while the key variables, namely, the arrival and departure times at each station for each train, have been predetermined by chromosomes.

Considering the discrete characteristic for the train operation, a special binary coding approach is introduced here to solve the proposed model. Each gene location corresponds to a possible time point for representing a decision result at the start terminal during the operation period, where ‘1’ indicates departure and ‘0’ indicates cancellation at the corresponding moment. For example, if we set half a minute as the length of modeling time intervals and use 6:00 to 8:10 a.m. as the reduced operation period, a chromosome then associates with 260 possible time points as shown in Fig. 9. At the same time, the first and last trains departing at 6:00 and 8:10 a.m., respectively, are predetermined in this study.

A scenario shown in Fig. 9 corresponds to chromosome (1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1), which denotes that the departure times of trains at the starting terminal are 6:00:00, 6:01:00, 6:01:30, 6:03:00..., 8:09:00, 8:10:00 a.m., respectively. New train operation scenarios will be obtained continually by applying the genetic operators into the chromosomes coded by the above-proposed approach.

5.2. Modification toward solution feasibility

The generated scenarios may not satisfy the safety time interval constraint in Eq. (3) and the available number of EMU in
Inequality (12). As a result, the following algorithm is performed to modify a scheme to enable its feasibility, where \( \tau \) denotes the length of time interval between two adjacent genes.

**Algorithm 3.**

For train index \( j = 1 \) to \( K \).
Check to see if the headway constraint \( D^1_j - D^1_{j-1} \geq h \) is not satisfied. If the fleet size condition \( j \leq m \) or \( D^1_j \geq D^1_{j-m} + T_{j-m} \) is not satisfied, perform the following:

If \( D^1_{j+1} - D^1_j = \tau \), set \( D^1_j = D^1_{j+1} \) and move forward one label of following trains from \( j \).
Otherwise, set \( D^1_j = D^1_j + \tau \).
Endfor

5.3. Calculation of fitness function for finding a robust solution under input data uncertainty

The total waiting time and the other parameters are computed simultaneously by solving the objective and equality constraints, recursively, from the first train and the start terminal, step by step. However, the loading-requiring constraint (13) and the operating-period constraint (14) might not be feasible. In that case, a function is defined as the following to obtain an unconstrained optimization problem that integrates the loading and period constraints.

\[
E_{\omega} = Z_{\omega} + \sum_{j=1}^{K_n} \beta_j \cdot \max \left\{ 0, 2(N - 1)c\alpha - \sum_{u=1}^{2N-1} X^u_j \right\} + \sum_{u=1}^{2N-1} \gamma_u \cdot \max \left\{ 0, W^u_{K} - (c - X^{u-1}_K + O^{u}_K) \right\},
\]

where \( \omega \) is the day scenario index, \( \beta_j \) \( (j = 1, 2, \cdots, K_0) \) and \( \gamma_u \) \( (u = 1, 2, \cdots, 2N - 1) \) denotes penalty factors. \( E_{\omega} \) is the revised profile of objective on day \( \omega \). In formula (22), the first item, \( Z_{\omega} \) is the total passenger delay on day \( \omega \) from Eq. (15). The second term corresponds to the penalty of the loading requirement constraint (13), and the third one is the penalty associated with the operation time period constraint (14).

To recognize the input data uncertainty associated with highly complex and dynamic passenger trip desires over different days of a week, this study, in particular, incorporates different probabilistic passenger arrival profiles, different months of the year, or even, potentially, the impact of special events during regular weekdays. By doing so, we can also filter out the potential inherent randomness in the raw passenger arrival profiles collected from the AFCS. That is, the final problem aims to find a robust timetable solution so as to minimize the expected total system costs (or passenger waiting time) over multiple days with different demand trip records.

The enhanced algorithm for finding robust solutions is constructed by the following simple revision. We can still use a single-solution representation in GA. However, the fitness function will be extended to an expected total value over different days (shown in objective function (23)), and the value for each day will require a separate evaluation, using that particular day’s demand profile. Finally, we can just select the solution that optimizes the expected total value over different days.

\[
E = \sum_{\omega} E_{\omega}
\]

Finally, the fitness function in GA is

\[
\text{fitness} = \frac{E_{\max} - E}{E_{\max} - E_{\min}},
\]

where \( E_{\max} \) and \( E_{\min} \) denote the maximum and minimum value of \( E \).

The other steps and approaches of genetic algorithms are similar to the standard GA algorithm, and interested readers are referred to the related references (e.g., Niu and Hu, 1998; Gen and Cheng, 2000).

5.4. Time-based decomposition method for large-scale problems

By following the spirit of the proposed dynamic programming algorithm, we develop a decomposing method in this paper to solve the proposed model, effectively, for a relatively longer period (e.g., one day). The idea is to divide the solution process into two stages. The first stage focuses on a sub-problem, and the second stage solves the problem based on the results from the first
stage.

(1) First stage

We first decompose the study period $[T_b, T_e]$ into two sub-periods, $[T_b, T_m]$ and $[T_m, T_e]$, where $T_b$ is the begin time, $T_e$ is the end time and $T_m$ is the intermediate time for the study period. It is necessary to count the waiting delay incurred by passengers who cannot board the last train $K$ during period $[T_b, T_m]$. As a result, this paper assumes that a dummy train $K + 1$ departs from the start terminal with the same interval as previous trains. The departure time is determined as follows:

$$D_{K+1}^1 = \max \{ D_k^1, (D_k^1 - D_{K-1}^1), T_m \}$$

(25)

$$D_{K+1}^u = D_k^1 + \xi_u$$

for $u = 2, 3, \cdots, 2N - 1$, (26)

where $\xi_u = \sum_{i=1}^{u-1} (r^i + s^{i+1})$ denotes the value of the departure time at station $u$ minus the departure time at the starting terminal.

The latest arrival time, $L_{K+1}^u$, for the dummy train, based on departure time $D_{K+1}^u$, can be computed by using formula (11). The waiting time, $W(t)$, for passengers who cannot board the last train $K$, is calculated as follows:

$$W(t) = \begin{cases} D_{K+1}^u - t & \text{if} \ L_{K+1}^u < t \leq L_{K+1}^u \\ D_{K+1}^u + (D_k^1 - D_{K-1}^1) - t & \text{if} \ L_{K+1}^u < t \leq T_m + \xi_u \end{cases}$$

(27)

Then, we apply the proposed algorithm for sub-period $[T_b, T_m]$. A subset consists of the best 20% of solutions in the final population of GA, and the corresponding objectives are then obtained for the second stage.

(2) Second stage

At first, let $\zeta_1, \zeta_2, \cdots, \zeta_I$ denote the values of objectives for the obtained subset. The probability is calculated as follows:

$$Pr(\tau) = \frac{\zeta_\tau}{\sum_k \zeta_k} \quad \text{for} \quad \tau = 1, 2, \cdots, I.$$

(28)

Then, we design a chromosome by connecting two sub-chromosomes, where the preceding part is generated by solutions from the optimal subset from the first stage and the subsequent part is generated by a randomization approach. The number of occurrences in chromosomes for a solution from the first stage is determined by the corresponding probability. The calculation is performed by applying the saluting-building method to the first, smaller period. The optimal timetable over one day can be obtained by moving the calculation from a shorter period forward to a longer period, recursively.

6. Numerical Example

6.1. Input data and parameter settings

The presented solution procedure is applied to design the timetable for the No. 8 subway line in Guangzhou city of China, which consists of 13 stations. The numbers of stations and the dwell times are given in Table 2, and the dwell time at station 3 also includes the required time of return operation. The running times between two adjacent stations are given in Table 3.

<table>
<thead>
<tr>
<th>Station Number</th>
<th>Station</th>
<th>Dwell Time</th>
<th>Station Number</th>
<th>Station</th>
<th>Dwell Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Fenghuang Xincun</td>
<td>0.50</td>
<td>8</td>
<td>Kecun</td>
<td>0.75</td>
</tr>
<tr>
<td>2</td>
<td>Shuyuan</td>
<td>0.50</td>
<td>9</td>
<td>Chigang</td>
<td>0.50</td>
</tr>
<tr>
<td>3</td>
<td>Baogang Dadao</td>
<td>0.50</td>
<td>10</td>
<td>Modiesha</td>
<td>0.50</td>
</tr>
<tr>
<td>4</td>
<td>Changgang</td>
<td>0.50</td>
<td>11</td>
<td>Xingangdong</td>
<td>0.75</td>
</tr>
<tr>
<td>5</td>
<td>Xiaogang</td>
<td>0.50</td>
<td>12</td>
<td>Pazhou</td>
<td>0.50</td>
</tr>
<tr>
<td>6</td>
<td>Sun Yat-sen University</td>
<td>0.75</td>
<td>13</td>
<td>Wanshengwei</td>
<td>5.00</td>
</tr>
<tr>
<td>7</td>
<td>Luijiang</td>
<td>0.50</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Adjacent Station Pair</th>
<th>Running Time</th>
<th>Adjacent Station Pair</th>
<th>Running Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-2</td>
<td>2.20</td>
<td>7-8</td>
<td>1.80</td>
</tr>
<tr>
<td>2-3</td>
<td>1.70</td>
<td>8-9</td>
<td>2.30</td>
</tr>
<tr>
<td>3-4</td>
<td>1.90</td>
<td>9-10</td>
<td>2.45</td>
</tr>
<tr>
<td>4-5</td>
<td>2.15</td>
<td>10-11</td>
<td>2.50</td>
</tr>
</tbody>
</table>
The values of the dwell times and the running times along one direction are the same as those of the other direction that are shown in Tables 2 and 3. With a particular focus on the morning peak hours, the study period below considers a relatively short interval, from 6:00 to 8:10 a.m. at the starting terminal. The time-dependent demand patterns in terms of total number of passengers over five workdays from Monday to Friday are shown in Fig. 10.

This study sets the number of EMU as 13 and the pull-out time at the start terminal as four min, according to the operation practice. Parameters $h$, $c$ and $\alpha$ are set as two, 700 and 60%, respectively. The number of trains that satisfy the load-rate constraints, $K_0$, is set as 70% of the total number of departed trains. The penalty factors, $\beta_j$ and $\gamma_u$, are set as the maximum value of objectives with the chromosomes of the current population. In addition, we set the size of population as 30, the number of total iterations as 1000, the crossover probability as 0.90 and mutation probability as 0.10.

6.2. Solutions by different timetabling strategies

The optimal departure times of trains at the start terminal, based on five day demand patterns, are given in Table 4, and the timetabling scenario is shown in Fig. 11. We have also solved the problem by the proposed decomposing algorithm. The improvement method demonstrates evident advantages over the original method due to the former method discovering the best available solution at 523 iterations and the latter at 992 iterations.

<table>
<thead>
<tr>
<th>Train</th>
<th>EMU</th>
<th>Departure Time</th>
<th>Train</th>
<th>EMU</th>
<th>Departure Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>6:00:00</td>
<td>10</td>
<td>10</td>
<td>7:17:00</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>6:10:00</td>
<td>11</td>
<td>11</td>
<td>7:24:00</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>6:20:00</td>
<td>12</td>
<td>12</td>
<td>7:30:30</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>6:29:00</td>
<td>13</td>
<td>13</td>
<td>7:35:30</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>6:36:30</td>
<td>14</td>
<td>14</td>
<td>7:41:30</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>6:46:30</td>
<td>15</td>
<td>15</td>
<td>7:52:00</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>6:56:00</td>
<td>16</td>
<td>3</td>
<td>7:55:30</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td>7:03:30</td>
<td>17</td>
<td>4</td>
<td>8:04:30</td>
</tr>
<tr>
<td>9</td>
<td>9</td>
<td>7:10:30</td>
<td>18</td>
<td>5</td>
<td>8:10:00</td>
</tr>
</tbody>
</table>
The intervals of trains in Fig. 11 are essentially consistent with the demand patterns in Fig. 10. For example, the larger headway corresponds to the lower demand before 7:00 a.m., while the larger demand is associated with the smaller headway around 7:30 a.m. We have also calculated the total passenger waiting times associated with the obtained timetable and the schedule with an even headway that satisfy all the constraints. The comparison results are given in Table 5. Obviously, with the proposed method in this paper, the average waiting time is reduced by 19.08%, and the average load rate is increased by 5.34%. This paper also demonstrates that the demand-based timetabling method can effectively reduce the waiting time of passengers and improve the utilization rate of EMUs during the study period.

Table 5. The comparison between the even and demand-based timetabling solutions

<table>
<thead>
<tr>
<th>Timetabling Scenario</th>
<th>Total Objective(min)</th>
<th>Average Waiting Time(min)</th>
<th>Average Load-rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demand-based headway scenario</td>
<td>919017.30</td>
<td>4.79</td>
<td>60.94</td>
</tr>
<tr>
<td>Even headway scenario</td>
<td>1134256.93</td>
<td>5.92</td>
<td>57.85</td>
</tr>
</tbody>
</table>

In order to demonstrate the proposed robust solution algorithm in section 5, we also calculated the individual-day-based objectives based on demand patterns from Monday to Friday. Fig. 12 illustrates the robustness and effectiveness of the algorithm from the decreasing trend of the across-day results.
To examine the benefit of a robust solution, we apply a single-day solution to five different days and calculate its average weekly waiting time over five different days. The final timetables have been obtained using the time-dependent demand records from Monday to Friday, separately. Then, the scheme for one day is applied to the other four days. The average waiting time and the other constraint-related statistics over five different days are shown in Table 6, where BC denotes the benefit constraint related to the loading requirement and SC denotes the fleet size constraint related to the train operating time period.

<table>
<thead>
<tr>
<th>Timetabling Scenario</th>
<th>Monday</th>
<th>Tuesday</th>
<th>Wednesday</th>
<th>Thursday</th>
<th>Friday</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>BC</td>
<td>SC</td>
<td>WT (min)</td>
<td>BC</td>
<td>SC</td>
</tr>
<tr>
<td>Monday</td>
<td>Yes</td>
<td>Yes</td>
<td>3.92</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Tuesday</td>
<td>Yes</td>
<td>Yes</td>
<td>5.23</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Wednesday</td>
<td>No</td>
<td>Yes</td>
<td>—</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Thursday</td>
<td>No</td>
<td>No</td>
<td>—</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Friday</td>
<td>No</td>
<td>No</td>
<td>—</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Optimal</td>
<td>Yes</td>
<td>Yes</td>
<td>4.77</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

This study also solves the problem according to different numbers of supplied EMU, without considering the loading constraint. Three parameters, the total number of departing trains and the average waiting time of passengers, as well as the average load-rate of trains, are computed based on varying numbers of EMUs from 10 to 16. Table 7 lists the computational results, and Fig. 13 shows the corresponding tendency graph.

<table>
<thead>
<tr>
<th>Number of EMU</th>
<th>Number of trains</th>
<th>Average waiting time(min)</th>
<th>Average load-rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>19</td>
<td>5.16</td>
<td>56.55</td>
</tr>
<tr>
<td>11</td>
<td>21</td>
<td>4.28</td>
<td>52.21</td>
</tr>
<tr>
<td>12</td>
<td>22</td>
<td>3.37</td>
<td>49.86</td>
</tr>
<tr>
<td>13</td>
<td>23</td>
<td>3.09</td>
<td>47.70</td>
</tr>
<tr>
<td>14</td>
<td>25</td>
<td>2.99</td>
<td>43.88</td>
</tr>
<tr>
<td>15</td>
<td>27</td>
<td>2.68</td>
<td>40.63</td>
</tr>
<tr>
<td>16</td>
<td>29</td>
<td>2.56</td>
<td>37.83</td>
</tr>
</tbody>
</table>

![Average waiting time - Average load-rate](image)

Fig. 13. The comparison results using different numbers of supplied EMU.

As expected from Fig. 13, the average waiting time for passengers declines with an increase in the number of supplied EMU. At the same time, the average load-rate of trains declines, which slightly downgrades the efficiency of EMU.

7. Conclusions
In this paper, we analyze the characteristic of time-dependent demands and the behavior of passengers arriving and waiting at the stations of a heavily congested urban rail line. We introduce a number of variables for effective passenger loading periods to calculate the number of boarded passengers at possibly oversaturated stations. Taking the overall waiting times as the objective, this paper formulates an optimization model of demand-based and congestion-sensitive timetables. The EMU fleet resource constraint is also considered systematically. A number of optimal and heuristic solution algorithms are developed to solve the proposed model. In particular, based on cumulative input-output diagrams, two novel solution algorithms, namely, local improvement and dynamic programming methods, are presented to find an optimal timetable for single station cases. A customized genetic algorithm is developed to solve the proposed model with a special binary coding method that indicates a train departure or cancellation at every possible time point. The proposed model and algorithm have been successfully tested through a real-world case study, and it is shown that the proposed method could effectively solve the timetabling problem of an urban rail line, based on time-dependent, multiple origin-to-destination demand profiles. Our future study will consider the response of passengers to the optimized timetable and extend our model to a network case.

Acknowledgements

This research work of the first author was supported, in part, by the National Natural Science Foundation of China (No. 50968009) and the Research Fund for the Doctoral Program of Higher Education of China (No. 20096204110003).

References


