Energy-saving metro train timetable rescheduling model considering ATO profiles and dynamic passenger flow

Hou, Zhuopu; Dong, Hairong; Gao, Shigen; Nicholson, Gemma; Chen, Lei; Roberts, Clive

DOI:
10.1109/TITS.2019.2906483

License:
Other (please specify with Rights Statement)

Document Version
Peer reviewed version

Citation for published version (Harvard):

Link to publication on Research at Birmingham portal

Publisher Rights Statement:
Checked for eligibility: 17/04/2019

© 2019 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.


https://doi.org/10.1109/TITS.2019.2906483

General rights
Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

• Users may freely distribute the URL that is used to identify this publication.
• Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
• User may use extracts from the document in line with the concept of ‘fair dealing’ under the Copyright, Designs and Patents Act 1988 (?)
• Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy
While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

Download date: 26. Oct. 2019
Energy-Saving Metro Train Timetable Rescheduling Model Considering ATO Profiles and Dynamic Passenger Flow

Zhuopu Hou1, Hairong Dong1,∗, Shigen Gao1, Gemma Nicholson2, Lei Chen2, Clive Roberts2

Abstract—For metro systems in over-crowded conditions, when an unexpected disturbance occurs, the operation of trains might be disturbed due to the high frequency and density of the metro traffic. A large number of passengers might be stranded on platforms due to service gaps and the limited free capacity of trains. In this paper, by introducing binary variables as selection indicators for ATO profiles which were preset in on-board ATO systems by metro signal suppliers, we develop a mixed integer programming (MIP) model for a metro train timetable rescheduling problem in order to jointly optimize the total train delay, the number of stranded passengers and the energy consumption of trains. We formulate the total energy consumption as the difference between the tractive energy consumption and the regenerated energy by considering the mass of in-vehicle passengers. Then, we adopt commercial optimization software CPLEX to solve the proposed model, which can obtain trade-off solutions in a short time. Finally, three numerical experiments based on real-world operational data are carried out to verify the effectiveness of the proposed method.

Index Terms—Metro train timetable rescheduling, mixed integer programming, ATO profile

I. INTRODUCTION

A. Motivation

An urban metro system is considered to be the pillar of urban public transportation systems because of its larger transport capacity, higher punctuality and lower energy consumption compared with other conventional transportation services [1]–[3]. Although equipped with Automatic Train Control (ATC) systems to increase the line capacity and transport efficiency, the dramatic increase in metro passenger numbers in recent years and limited infrastructural resources, means that many metro systems such as the Beijing Subway and the Shanghai Subway are facing high-load operations [4]. In such cases, metro operation is highly likely to be vulnerable when unavoidable disturbances occur during daily operation. Here disturbances refer to those events (e.g., obstruction of doors, medical emergencies) that cause trains to have a relatively short duration of delay, and the effect of the delay can be gradually eliminated by adjusting the timetable. A common effect of disturbances is that the train dwells at a station for longer time than planned in the original timetable. Delays might propagate to the next station and affect the operation of subsequent trains once the reserve headway capacity is used up [5], resulting in original timetable infeasibility, passengers accumulating and platform crowding, which might cause potential safety hazards for passengers on the platforms and affect service quality of the entire metro line. In such cases, the original train timetable needs to be rescheduled as soon as possible to allow affected trains to resume normal operation, which is the called train timetable rescheduling (TTR) problem [6].

In practical metro operations, when delays occur, the rescheduling process is carried out either manually by dispatchers using their experience or automatically by the Automatic Train Supervision (ATS) systems. For those metro lines equipped with ATO systems, a set of pre-programmed train speed profiles corresponding to different operation levels are preset in the on-board ATO systems by metro signal suppliers [7]–[9]. Different operation levels correspond to various running times. During metro operation, the ATS systems continuously compare the original timetable with the actual arrival and departure time of trains. When a timetable deviation is detected, the ATS system selects the operation level according to the rescheduled running time calculated by the ATR module and sends it to the on-board ATO system of the affected train. In the case of minor delays that only a single train needs to be rescheduled, the current ATS system can manage it well [8]. But when it comes to the situations which multi trains need to be rescheduled, this work is still manually done by human dispatchers with their professional knowledge and experiences [10]. This experience-based manual rescheduling mode has the following three disadvantages [4]. Firstly, it lacks rigorous computation and optimisation. Secondly, it cannot effectively integrate the information of dynamic passenger flow. Lastly, it is hard for human dispatchers to acquire high quality rescheduling plans in real time.

In view of the above issues, it is practically significant to propose an effective method based on the actual ATS and ATO system for solving the multi-train rescheduling problem in real...
time. This research will specifically investigate this problem.

B. Literature Review

As a vital issue in the daily operation of railway systems, the train timetable rescheduling problem has drawn tremendous attention in the past few decades. In general, the literature on train timetable rescheduling can be divided into two categories based on the extent of deviation to the train operation: (1) disruption management and (2) disruption management. The main difference is that the former copes with those events that cause trains to have a relatively short duration of delay by adjusting the timetable, while the latter deals with the relatively large incidents by modifying both the timetable and the duties for rolling stock and crew [6]. Most of the literature tends to formulate mathematical models with practical operational constraints, such as a quadratic programming (QP) model [11], a mixed integer programming (MIP) model [12] and a alternative graph (AG) model [13], which aim to reduce the deviation between the rescheduled and original timetable. Solution methodologies, such as heuristic algorithms [14], exact algorithms (e.g. branch and bound) [15] and commercial solvers (e.g. CPLEX, GAMS) [5], [16] are adopted to seek the optimal solution of the train rescheduling problem.

In case of an unexpected disturbance, the planned timetable may become infeasible and needs to be modified in real time. For mainline railways, the commonly used methods consist of rerouting, reordering and retiming of affected trains [6], according to the characteristics of the infrastructure of railways, such as single-track, double-track and junctions. Yang et al. [16] presented a two-stage integer programming model for double-track railway rescheduling. A fuzzy variable was introduced to denote the disturbance duration and the numerical experiments were solved using the commercial software GAMS. Chen et al. [17] developed a MIP model for a train rescheduling problem at a junction area, an improved algorithm (DE-JRM) was proposed to and the case study was performed on the core area of Thameslink route with Monte Carlo methods. Meng and Zhou [18] proposed a Lagrangian relaxation algorithm to solve the N-track railway network rescheduling problem simultaneously considering rerouting and retiming strategies. The alternative graph model was adopted by D’Ariano et al. [19] to represent the train dispatching process, and the speed coordination among consecutive trains was also considered.

In case of disruptions (e.g., section blockage, locomotives malfunction), for which the duration is much longer than disturbances, not only timetable rescheduling but also other measures such as cancellation of trains, skip-stop patterns, etc. need to be taken to resume trains to normal operations. To solve the train rescheduling problem in case of a blockage of a railway segment, Zhan et al. [20] formulated a MIP model which considered the station capacity and cancellation of trains. A bi-objective MILP (Mixed Integer Linear Programming) model and a heuristic iterative algorithm was proposed by Gao et al. [21] to solve train rescheduling problem for a metro line in the case of disruptions.

Different to the mainline railway system, overtaking between two successive trains is not possible in most cases due to the characteristics of the metro infrastructure [11]. The main method of handling delays caused by disturbances is to rearrange the arrival and departure time of affected trains. A quadratic programming model was developed to deal with the traffic regulation problem for metro loop lines by Fernandez et al. [5], who adopted the commercial software CPLEX to solve the proposed model. Bai et al. [22] aimed to reduce the total delay and improve the equilibrium of arrival and departure times of affected trains in cases of disturbance. A Genetic Algorithm (GA) was adopted to solve the problem. Gao et al. [10] developed a metro train rescheduling strategy which considers the information of fault handling. The rescheduled timetable was adjusted as the feedback to the fault handling information. Xu et al. [23] developed a metro train rescheduling framework which considers cross tracks to balance the service quality in cases of disturbance. A discrete event model was formulated and solved by combining the commercial software GAMS with an efficient heuristic algorithm. The information on the dynamic passenger flow, which is an important factor affecting the results of metro train timetable rescheduling, was not, however, taken into account in these studies. When it comes to the over-crowded situation of a metro line, the above methods will not be suitable anymore.

With the development of automatic fare collection (AFC) technology, adoption of passenger dynamic description in a metro traffic regulation model is more reasonable and realistic by analyzing historical passenger flow data. A constrained state-space model for the metro train rescheduling problem was investigated in [24], which considers the stochastic passenger arrival flow as a discrete Markovian process. Li et al. [11] investigated a joint optimal train regulation and passenger flow control strategy for metro lines. They solved this problem by applying a model predictive control method and verified the stability of the proposed control system based on Lyapunov stability theory. However, these studies did not investigate the stranded passengers on the platform due to service gaps and the limited free capacity of trains. Since reducing the passenger accumulation on the platform is a vital work in the metro operations [25]. Hou et al. [26] investigated the metro train rescheduling problem with the aim of reducing the number of stranded passengers on the platform after a disturbance. A new hybrid algorithm was developed to solve the model. However, this research did not consider the impact of the rescheduling strategy on the train energy consumption, since the variability in energy consumption of a train is mostly influenced by the running time in each section [4]. Yin et al. [4] proposed a stochastic programming model which aims to minimize the passenger delay time, traveling time and energy consumption of affected trains. An approximate dynamic programming-based algorithm was designed to solve the proposed problem. But the impact of the passenger flow on train dwell time was not considered in this research. It bears nothing that precise calculation of the train dwell time at the stations can improve the operation efficiency of the metro systems [27]. Sheu and Lin [28] designed a dual heuristic programming method for metro train regulation problem. The energy saving and traffic stability was achieved by adjusting the running time and dwell time of trains. However, the above-
mentioned methods are not suitable for those metro systems which use a set of pre-programmed train speed profiles in the on-board ATO system.

C. The Focuses of This Study

To the best of our knowledge, the existing literature for metro train timetable rescheduling problems rarely considers the pre-programmed recommended speed profiles in the on-board ATO system. To fill the gap between the practical application and theoretical research of the metro train rescheduling problem, this paper attempts to present an optimization model for those metro systems which equipped with ATO systems pre-programmed with a set of recommended speed profiles. In summary, considering the research gaps mentioned in Section I-B, we try to make the following contributions:

1) For trains operated in ATO mode, the train running time in sections is determined by the pre-programmed speed profiles on-board ATO systems [8]. In this research, by using binary variables to select train operation level profiles instead of restricting running time with lower and upper bounds, we formulate a novel mixed integer programming (MIP) model to solve the metro train timetable rescheduling problem.

2) We jointly optimize the total accumulated delay, number of stranded passengers and energy consumption of the affected trains in cases of disturbance. We evaluate the energy consumption of trains based on the selected operational profiles, which reduces the complexity of the model. The adoption of the commercial optimization software CPLEX allows the model to be solved in a reasonable time.

The remainder of this paper is organized as follows: Detailed descriptions of the considered problem and on-board ATO operation level profiles are described in Section II. The optimization model for solving metro train timetable rescheduling which considers the passenger flow and energy consumption is introduced in Section III. In Section IV, numerical experiments based on real operational data are carried out and solved by commercial software CPLEX to demonstrate the efficiency of the proposed model. The paper ends with conclusions and suggestions for future research.

II. PROBLEM DESCRIPTION

![Fig. 1. Typical metro line sketch](image)

In order to conveniently describe the metro train timetable rescheduling problem and formulate the model, the relevant sets, indices and parameters are listed in Table I. A typical metro line infrastructure considered in this paper is illustrated in Fig. 1, where stations are numbered as 1, 2, · · · , |M| and sections are numbered as 1, 2, · · · , |M| − 1.

<table>
<thead>
<tr>
<th>TABLE I</th>
<th>PARAMETERS AND NOTATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>M = {1, 2, ...,</td>
<td>M</td>
</tr>
<tr>
<td>N = {1, 2, ...,</td>
<td>N</td>
</tr>
<tr>
<td>K = {1, 2, ...,</td>
<td>M</td>
</tr>
<tr>
<td>i</td>
<td>Index of station i, i ∈ M</td>
</tr>
<tr>
<td>j</td>
<td>Index of train j, j ∈ N</td>
</tr>
<tr>
<td>k</td>
<td>Index of section k, k ∈ K and k = i</td>
</tr>
<tr>
<td>l</td>
<td>Index of operation level l, l ∈ L</td>
</tr>
<tr>
<td>Mt</td>
<td>Mass of a train</td>
</tr>
<tr>
<td>M_{\text{avg}}</td>
<td>Average weight of a passenger</td>
</tr>
<tr>
<td>a_{i,j}^{\text{sec}}</td>
<td>Acceleration rate of trains in section k</td>
</tr>
<tr>
<td>D_{d,\text{min}}</td>
<td>Minimum dwell time of train j at station i</td>
</tr>
<tr>
<td>D_{d,\text{max}}</td>
<td>Maximum dwell time of train j at station i</td>
</tr>
<tr>
<td>H_{\text{min}}</td>
<td>Minimum delay between two adjacent trains in the same section</td>
</tr>
<tr>
<td>p_{\text{ boa}}</td>
<td>Minimum delay between two adjacent trains at the same station</td>
</tr>
<tr>
<td>p_{\text{ boa}}</td>
<td>The number of boarding passengers of train j at station i</td>
</tr>
<tr>
<td>p_{\text{ al}}</td>
<td>The number of alighting passengers of train j at station i</td>
</tr>
<tr>
<td>\lambda_{i,j}</td>
<td>The arrival rate of passengers waiting for train j at station i per second</td>
</tr>
</tbody>
</table>

A. Model Assumptions

The following assumptions are made with regard to the proposed model.

1) There is only one train running in a section at a time, it is common for most urban rail lines [10]. Once the operational level is selected before train starts, the ATO profile cannot be changed until train stops at next station.

2) The delay duration is known. Trains can track the recommended speed profiles accurately. The regenerative energy is assumed to be stored in the stationary storage systems. The energy recovery rate between the kinetic and electricity energy is assumed to be constant [29]. Energy consumed by the on-board auxiliary equipment is disregarded in this paper.

3) The number of passengers arriving at the station is assumed to be distributed uniformly, which means that this number is proportional to the operational headway between two successive trains. The number of alighting passengers of the train is assumed to be proportional to the number of in-vehicle passengers [11]. The transfer passengers are not considered in this paper.
4) We assume that the disturbance has no effect on trains running in the other direction. The overtaking and crossing operation is not considered in this paper.

B. Description of Passenger Flows

Train delays caused by disturbances might result in passengers accumulating and platform overcrowding, which are potential safety hazards for passengers. In turn, accumulated passengers and the crowded platform will extend dwell times of affected trains, cause service gaps and aggravate delays. Therefore, understanding passenger flows is vital for metro train timetable rescheduling.

The number of passengers arriving at the platform of station \( i \) when train \( j \) departs from station \( i \) is given by

\[
P_{\text{arr}}^{i,j} = \lambda_{i,j} \cdot (td_{i,j} - td_{i,j-1}),
\]

where \( \lambda_{i,j} \) indicates the passengers’ arrival rate at station \( i \) during the rescheduling time horizon, which is assumed to be distributed uniformly. This assumption is considered reasonable for metro systems operated with a short headway [30]. It is introduced to indicate the number of passengers reaching the platform at station \( i \) in unit time.

The number of on-board passengers when train \( j \) departs from station \( i \) is equal to the number of on-board passengers when the train departs from the previous station plus the difference between the number of boarding and alighting passengers. Then the dynamic variance of in-vehicle passengers can be expressed as

\[
P_{\text{in}}^{i,j} = P_{\text{in}}^{i-1,j} + P_{\text{boa}}^{i,j} - P_{\text{ali}}^{i,j},
\]

where \( P_{\text{in}}^{i,j} \) represents the number of on-board passengers in train \( j \) when it departs from station \( i \); \( P_{\text{boa}}^{i,j} \) and \( P_{\text{ali}}^{i,j} \) are the number of boarding and alighting passengers when train \( j \) dwells at station \( i \), respectively. Based on the passenger demand origin-destination (OD) matrices during a short period of the day, the alighting ratio \( \omega_{i,j} \), which represents the proportion of alighting passengers to in-vehicle passengers when train \( j \) arrives at station \( i \), is introduced. Then the number of passengers getting off train \( j \) at station \( i \) is given by

\[
P_{\text{ali}}^{i,j} = P_{\text{in}}^{i-1,j} \cdot \omega_{i,j}.
\]

The number of passengers that can board a train depends on the train capacity, and the number of in-vehicle and alighting passengers. Therefore, the number of boarding passengers when train \( j \) stops at station \( i \) can be expressed as

\[
P_{\text{boa}}^{i,j} = \min\{ (P_{\text{cap}} - P_{\text{in}}^{i-1,j} + P_{\text{ali}}^{i,j}), (P_{\text{arr}}^{i,j} + P_{\text{str}}^{i,j-1}) \}. \tag{4}
\]

A proportion of the passengers waiting for train \( j \) may not get on the train due to the limited free capacity of the train and they are left to wait for the successive train \( j+1 \). We can thus denote the number of stranded passengers on the platform when train \( j \) departs from station \( i \) by

\[
P_{\text{str}}^{i,j} = P_{\text{str}}^{i,j-1} + P_{\text{str}}^{i,j} - P_{\text{boa}}^{i,j}.
\]

If the number of passengers waiting for train \( j \) is less than the free capacity of the train then no one will be stranded on the platform.

C. Calculation of Energy Consumption

Reducing energy consumption has become a global concern. Particularly, in metro systems, energy-saving is an important criteria for operators due to its association with operation costs as well as environmental concerns. The variability in energy consumption of a train is mostly influenced by the running time in each section [4]. During the train rescheduling process, the energy consumption of a train running in the section might be increased since the higher operation level corresponding to the shorter running time is selected to deal with a delay. The concept of operation levels of the ATO system is illustrated in Fig. 2. Each operation level corresponds to a unique train speed profile, and the lower operation level (numerical higher) corresponds to the longer running time. Given a specific train speed profile, the energy consumption of the train can be calculated. Howlett [31] validated that the optimal energy-saving operation mode sequence for a train running in a section should be maximum acceleration, cruising, coasting and maximum deceleration. In the acceleration phase, the energy consumption of train \( j \) running in section \( k \) with operation level \( l \) is denoted by \( E_{k,j}^{l,ac} \), which equals the work done by the traction force. The energy consumption in the cruising phase is denoted by \( E_{k,j}^{l,cr} \), to keep the train move at a constant speed the resultant forces acting on the train should be zero, thus the traction forces are equal to the resistance. In the coasting phase, the energy consumption of trains is 0 since no output of traction effort and braking effort. In the deceleration phase, the work done on the train by the braking force converts the kinetic energy into wasted heat and electric energy. The latter is known as regenerative braking energy, which can be either consumed immediately or stored until needed. In this research, we assume the regenerative braking energy can be stored in an energy storage system [32]. The energy generated by the braking force is represented by \( E_{k,j}^{l,de} \), the energy recovery rate \( \eta_{k,j} \) is introduced to denote the proportion of the stored regenerative energy. The total energy consumption of train \( j \) running in section \( k \) with operation level \( l \) can be expressed as

\[
E_{k,j}^{l} = E_{k,j}^{l,ac} + E_{k,j}^{l,cr} - \eta_{k,j} \cdot E_{k,j}^{l,de}.
\]

Considering the time \( t \) as the dependent variable, according
to Newton’s law of motion, equations of train motion in the acceleration phase and the deceleration phase can be formulated as

\[ M_{k,j} \frac{dv_{k,j}}{dt} = u \cdot f^a_{k,j} - (1-u) \cdot f^b_{k,j} - f^g_{k,j}, \]  

where \( v_{k,j} \) is the current train speed; \( f^a_{k,j} \) is the train traction force in the acceleration phase; \( f^b_{k,j} \) is the train braking force in the deceleration phase; \( u \) is introduced as a control signal indicator, equal to 1 when the forward traction force works and equal to 0 when the backward braking force works; \( M_{k,j} \) is the total train mass, equal to the pure mass of a train plus the total mass of on-board passengers in section \( k \), and can be calculated by equation (8); \( f^g_{k,j} \) is the frictional resistance, and it can be calculated by equation (9), where \( r_1 \), \( r_2 \) and \( r_3 \) are the resistance coefficients of the train and are usually provided by the rolling stock manufacturers [33]; \( f^g_{k,j} \) is the gravity component of the train along the track with an average gradient of \( \theta_k \), which is represented in equation (10).

\[ M_{k,j} = M_t + \sum_{i,j} M_p, \]  

\[ f^a_{k,j} = M_{k,j} \cdot (r_1 + r_2 \cdot v_{k,j} + r_3 \cdot v_{k,j}^2), \]  

\[ f^g_{k,j} = M_{k,j} \cdot g \cdot \sin \theta_k. \]

Thus the energy consumption of the acceleration phase can be calculated as

\[ E_{k,j}^{l,ac} = \int_{0}^{\tau_{k,j}^{l,ac}} f^a_{k,j} \cdot a_{k,j}^{ac} \cdot t \, dt, \]  

where \( \tau_{k,j}^{l,ac} \) is the operation time for acceleration phase when train \( j \) runs in section \( k \) with operation level \( l \); \( a_{k,j}^{ac} \) is the maximum acceleration rate. The energy consumption of the cruising phase can be given as

\[ E_{k,j}^{l,cr} = \int_{0}^{\tau_{k,j}^{l,cr}} (f^a_{k,j} + f^g_{k,j}) \cdot a_{k,j}^{ac} \cdot t \, dt, \]

where \( \tau_{k,j}^{l,cr} \) is the operation time for cruising phase. The generated kinetic energy in the deceleration phase can be given as

\[ E_{k,j}^{l,de} = \int_{0}^{\tau_{k,j}^{l,de}} f^b_{k,j} \cdot a_{k,j}^{de} \cdot t \, dt, \]

where \( \tau_{k,j}^{l,de} \) is the operation time for deceleration phase when train \( j \) runs in section \( k \) with operation level \( l \); \( a_{k,j}^{de} \) is the maximum deceleration rate.

**III. OPTIMIZATION MODEL FORMULATION**

**A. Decision Variables**

The metro train timetable rescheduling problem is essentially to rearrange the arrival and departure times of the trains affected by disturbances based on the interests of stakeholders. Thus the arrival time \( t_{a_{i,j}} \) and departure time \( t_{d_{i,j}} \) are considered as decision variables. Moreover, we also introduce a decision variable \( \epsilon_{k,j}^l \), which is used to indicate which operation level \( l \in L \) is selected.

\( \epsilon_{k,j}^l \) is a binary variable as indicator of metro train operation level, equal to 1 if and only if train \( j \) uses operation level \( l \) in section \( k \); \( t_{a_{i,j}} \) is a continuous variable as actual arrival time of train \( j \) at station \( i \); \( t_{d_{i,j}} \) is a continuous variable as actual departure time of train \( j \) at station \( i \).

**B. Objective Functions**

In this paper, we try to minimize the deviation between the original timetable and the rescheduled timetable, the total number of stranded passengers and the total train energy consumption by jointly adjusting operation levels and dwell times of affected trains, thus the proposed problem also is a multi-objective optimization problem.

We use \( T_{delay} \) to denote the total accumulated delay of affected trains, which is the criteria of the deviation between the original timetable and rescheduled timetable and can be expressed as

\[ T_{delay} = \sum_{i=1}^{M} \sum_{j=1}^{N} \{ (t_{d_{i,j}} - T_{d_{i,j}}) + (t_{a_{i,j}} - T_{a_{i,j}}) \}. \]

The number of passengers on the platform left by the train due to the limited free capacity when it departs from stations can be calculated according to equations presented in Subsection II-B. The total number of stranded passengers during the rescheduling period is \( N_{stranded} \), which can be given as

\[ N_{stranded} = \sum_{i=1}^{M} \sum_{j=1}^{N} P_{i,j}^{stranded}. \]

Given a specific operation level \( l \), the corresponding speed profile of train \( j \) running in section \( k \) is determined. Then the energy consumption of each train can be calculated according to the equations presented in Subsection II-C. To determine the speed profile of trains and optimize the energy consumption from a global perspective, the binary variable \( \epsilon_{k,j}^l \) is introduced to denote which operation level the train running in each sections with the specific operation level can be expressed as

\[ E_{total} = \sum_{k=1}^{M} \sum_{j=1}^{N} E_{k,j}^l. \]

In general, the total accumulated delay of affected trains and the number of stranded passengers can be reduced by making use of running time supplements and the reserve headway capacity, i.e., shortening the running times in a section and the operational headway. However, the energy consumption of affected trains might be increased. Therefore, these objectives are in conflict with each other. To obtain trade-off solutions among the above mentioned objectives, a wildly used approach called the the weighted sum method [34] is adopted and the objective function can be given as

\[ \min f = \alpha \cdot \frac{T_{delay}}{T_{delay}} + \beta \cdot \frac{N_{stranded}}{N_{stranded}} + \gamma \cdot \frac{E_{total}}{E_{total}}, \]

where \( \alpha, \beta \) and \( \gamma \) are three non-negative weight factors and their sum is equal to 1. The value of these coefficient are determined based on the intrinsic knowledge of the problem.
by the decision maker [35], in this context, the dispatchers. \(T_{\text{delay}}, N_{\text{stranded}}\) and \(E_{\text{total}}\) are three nominal values of total accumulated delay, total stranded passengers and total energy consumptions, respectively. Thus, we can transform these objectives into the same magnitude. These nominal values can be calculated based on a heuristic method which is introduced in Section IV.

C. Constraints

The objective is subject to various constraints to ensure train operational safety, enforce speed restrictions and permit train stops.

1) Headway time constraints: In order to guarantee operational safety, a certain time interval must be kept between two consecutive trains, i.e., the headway time. For two adjacent trains, the safety headway time in the same section and at the same station are ensured by restricting their arrival and departure times, which can be expressed as

\[
\begin{align*}
td_{i,j} - td_{i,j-1} & \geq h_{\text{sec}}^{\min}, \\
\min_a(ta_{i,j} - ta_{i,j-1}) & \geq h_{\text{sec}}^{\min}, \\
\min_a(ta_{i,j} - td_{i,j-1}) & \geq h_{\text{sta}}^{\min},
\end{align*}
\]

where \(h_{\text{sec}}^{\min}\) and \(h_{\text{sta}}^{\min}\) are used to denote the minimum headway time between two consecutive trains in the same section and at the same station, respectively.

2) Running time constraints: In many previous studies for train rescheduling in cases of disturbance, the train running time in a section is restricted with lower and upper bounds. That is, the train running time is considered to be in a continuous range. However, in practical metro lines equipped with ATO systems, several train speed profiles corresponding to different operation levels are preset in on-board ATO systems by metro signal suppliers [4]. Once a specific operation level is selected before the train departing from the station, the train travel time in the corresponding section is uniquely determined. Thus, running time constraints in this research are discrete sets instead of a continuous range, and can be formulated as

\[
\begin{align*}
ta_{i+1,j} - td_{i,j} & = \epsilon_{k,j} \cdot R_{k,j}^l, \\
\sum_{l=1}^{|L|} \epsilon_{k,j} & = 1,
\end{align*}
\]

where \(\epsilon_{k,j}\) is binary variable which is the selection indicator of train operation level \(l\); \(R_{k,j}^l\) is the running time of train \(j\) in section \(k\) corresponding to operation level \(l\); equation (21) links the departure time of train \(j\) from station \(i\) and the arrival time of train \(j\) at next station \(i+1\); equation (22) means that train \(j\) can be operated according to only one recommended speed profile when it runs in section \(k\), that is, only one of the \(|L|\) operation levels can be selected.

3) Dwell time constraints: The dwell time constraint of train \(j\) at station \(i\) is restricted with lower bound \(Dw_{i,j}^{\min}\) and upper bound \(Dw_{i,j}^{\max}\), and can be given as

\[
\begin{align*}
td_{i,j} - ta_{i,j} & \leq Dw_{i,j}^{\max}, \\
td_{i,j} - ta_{i,j} & \geq Dw_{i,j}^{\min}.
\end{align*}
\]

Although the dwell time of trains at each station are predefined during the planning phase, the actual metro train dwell time is mainly affected by the number of alighting and boarding passengers and the degree of crowdedness of the train [36]. In many studies for metro train scheduling/rescheduling problems [4], [10], [22], [23], a predefined minimum dwell time is usually adopted to be the lower bound of the train dwell time. In other words, they did not consider the actual train dwell time affected by the passenger. According to Zhuge et al. [27], the actual train dwell time can be estimated by

\[
E_{i,j} = a + b \cdot P_{\text{boa}}^{\text{min}} + c \cdot P_{\text{all}} + d \cdot \left( \frac{P_{\text{arr}}^{\text{min}}}{N_{\text{door}}} \right)^3 \cdot P_{\text{boa}},
\]

where \(a, b, c\) and \(d\) are dwell time coefficients that can be estimated based on historical passenger flow data. \(N_{\text{door}}\) is the total number of opened doors at each station. In this paper, we consider the lower bounds of the train dwell time as

\[
Dw_{i,j}^{\min} = \min\{S_{i,j}^{\min}, E_{i,j}\},
\]

where \(S_{i,j}^{\min}\) is the technical minimum dwell time predefined by the metro operator.

4) Train capacity constraints: The number of in-vehicle passengers of train \(j\) before it departs from station \(i\) should not exceed the maximum allowable capacity of the train, i.e. the maximum number of passengers on board,

\[
P_{\text{in},i,j} \leq P_{\text{cap}},
\]

\(P_{\text{in},i,j}\) is related to the operational headway of trains and can be calculated as shown in Subsection II-B.

5) Arrival and departure time constraints: In order to minimize the deviation between the original timetable and the rescheduled timetable, the following constraints are necessary conditions as explained in Gao et al. [10].

\[
\begin{align*}
ta_{i,j} & \geq Ta_{i,j}, \\
td_{i,j} & \geq Td_{i,j}.
\end{align*}
\]

Before the train departs from station \(i\) should not exceed the maximum allowable capacity of the train, i.e. the maximum number of passengers on board,
are 3 operation levels (i.e., 3 pre-programmed ATO profiles) for the trains in each section. It is clear that, the complexity of the proposed MIP model is fully dependent on the physical scale of the metro line, the number of affected trains and operation levels of the train control system.

### TABLE II
**Complexity analysis**

<table>
<thead>
<tr>
<th>Variables or constraints</th>
<th>The proposed model</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Binary variables $z_{k,j}$</td>
<td>$(</td>
<td>M</td>
</tr>
<tr>
<td>Continuous variables $t_{a,i,j}$, $t_{d,i,j}$</td>
<td>$2\cdot</td>
<td>M</td>
</tr>
<tr>
<td>Headway time constraints</td>
<td>$3\cdot</td>
<td>M</td>
</tr>
<tr>
<td>Running time constraints</td>
<td>$2\cdot(</td>
<td>M</td>
</tr>
<tr>
<td>Dwell time constraints</td>
<td>$2\cdot</td>
<td>M</td>
</tr>
<tr>
<td>Train capacity constraints</td>
<td>$</td>
<td>M</td>
</tr>
<tr>
<td>Arrival time constraints</td>
<td>$</td>
<td>M</td>
</tr>
<tr>
<td>Departure time constraints</td>
<td>$</td>
<td>M</td>
</tr>
</tbody>
</table>

### IV. Case Study

In this section, the effectiveness of the the proposed model in solving the metro train timetable rescheduling problem is verified by three experiments. The simulation is coded in MATLAB R2014a and run on a computer with Mac OS, 1.6 GHz Intel Core i5 processor and 8 GB RAM. We adopt the commercial optimization software CPLEX 12.6.2 to solve the proposed MIP model, and the YALMIP toolbox is used as the interface between CPLEX and MATLAB [37].

The parameters of CPLEX are set to be default values. CPLEX uses a built-in algorithm called branch-and-cut search to solve mixed integer programming (MIP) models. The time limit on the termination of the search procedure is set to be 10 seconds, in view of the real-time requirement of the proposed model.

#### A. Experiment 1: Comparison with FRM

In the first experiment, we consider a typical delay scenario where a disturbance occurs when train 4 dwells too long at station 3, resulting in a departure delay of 100 seconds.

Since the duration of delay exceeds running time margins and the buffer time, prolonged dwell times of train 4 lead to gaps in service and interference with successive trains. If no efficient action is taken, the delays will propagate throughout the entire line and perturb the operation of subsequent trains as shown in Fig. 3, where the grey rectangle denotes the duration of the disturbance and the blue solid lines represent the schedule without regulation. Thus, the proposed MIP model was adopted to solve the train timetable rescheduling problem.
TABLE IV  
THE ORIGINAL TIMETABLE AND PASSENGER FLOW DATA

<table>
<thead>
<tr>
<th>Operation data</th>
<th>Station index $i$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scheduled dwell time $D_{i,j}$ (s)</td>
<td></td>
<td>30</td>
<td>30</td>
<td>45</td>
<td>45</td>
<td>45</td>
<td>40</td>
<td>45</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>Minimum dwell time $S_{i,j}^{\text{min}}$ (s)</td>
<td></td>
<td>25</td>
<td>25</td>
<td>40</td>
<td>40</td>
<td>40</td>
<td>35</td>
<td>40</td>
<td>25</td>
<td>25</td>
<td>25</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>Maximum dwell time $D_{i,j}^{\text{max}}$ (s)</td>
<td></td>
<td>90</td>
<td>90</td>
<td>105</td>
<td>105</td>
<td>105</td>
<td>100</td>
<td>105</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Passenger arrival rate $\lambda_{i,j}$</td>
<td></td>
<td>1.40</td>
<td>1.51</td>
<td>1.49</td>
<td>1.54</td>
<td>1.57</td>
<td>1.51</td>
<td>1.48</td>
<td>1.31</td>
<td>1.43</td>
<td>1.40</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Passenger alighting ratio $\omega_{i,j}$</td>
<td></td>
<td>0</td>
<td>0.25</td>
<td>0.23</td>
<td>0.08</td>
<td>0.27</td>
<td>0.13</td>
<td>0.13</td>
<td>0.30</td>
<td>0.23</td>
<td>0.14</td>
<td>0.11</td>
<td>1</td>
</tr>
</tbody>
</table>

TABLE V  
RUNNING TIMES CORRESPONDING TO OPERATION LEVELS

| Operation level $l$ | Section index $k$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---------------------|------------------|---|---|---|---|---|---|---|---|---|----|----|----|
|                     |                  | 1 | 63 | 105 | 123 | 87 | 78 | 65 | 80 | 87 | 105 | 78 | 63 |
|                     |                  | 2 | 73 | 115 | 133 | 97 | 88 | 75 | 90 | 97 | 115 | 88 | 73 |
|                     |                  | 3 | 83 | 125 | 143 | 107 | 98 | 85 | 100 | 107 | 125 | 98 | 83 |
|                     |                  | 4 | 93 | 135 | 153 | 117 | 108 | 95 | 110 | 117 | 135 | 108 | 93 |
|                     |                  | 5 | 118 | 160 | 178 | 142 | 133 | 120 | 135 | 142 | 160 | 133 | 118 |
| Section length (m)  |                  | 839 | 1564 | 1649 | 1377 | 1188 | 983 | 1134 | 1377 | 1564 | 1188 | 839 |

Fig. 3. The timetable without rescheduling in case of a disturbance

Fig. 4. The comparison of three timetables

and compared with the heuristic method FRM. The nominal values of the total accumulated delay, passengers left by the trains and the energy consumption can be calculated by the FRM. To define the value of weights, we consulted some metro dispatchers about their preference when they cope with the considered delay scenario. The weight coefficients $\alpha$, $\beta$ and $\gamma$ are set as 0.4, 0.4 and 0.2, respectively, which reflect the preference and insight of the metro dispatchers to the delay scenario. The specific values of these coefficients will, actually, not have a substantial impact on the model of this paper.

Three train time-space diagrams are compared in Fig. 4, where the dotted lines represent the new timetable after the fixed regulation method was applied in case of disturbance and the green solid lines express the optimal rescheduled timetable with the proposed model. It can be seen that the number of following affected trains is 5 with the fixed regulation method, while it is 3 with the proposed model, and the extent of the effect weakened in the order of trains. The total accumulated train delay is 2053 s with FRM while it is 1570 s with the proposed model. The number of stranded passengers on the platform when the FRM was applied is shown in Fig. 5, where the total number of stranded passengers during the regulation period is 1605. The total number of stranded passengers was reduced to 1124 with the proposed model,
which is shown in Fig. 6. This is because the dwell time of affected trains was adjusted according to the passenger flow and the interval of two consecutive trains was optimized. The total energy consumption of trains is \(2.091 \times 10^9\) J with FRM while it is \(2.112 \times 10^9\) J with the proposed model. The energy consumption of the proposed model was increased a little because the running time of some affected trains was shortened to meet the reduction of the total accumulated delay.

The computational results can also reflect the preference of dispatchers who manage a metro line running frequent services during peak hours, that is, they are more inclined to give the same priority to reducing train delays and stranded passengers rather than energy consumption. This trade-off can be changed by adjusting the weight coefficients, the influence by different weight factors will be demonstrated in the second experiment.

### B. Experiment 2: Different Weights for Objectives

In the second experiment, we studied the effect of different weights in the objective function on reducing the effect of delay and saving energy. The objective function (17) is composed of three parts. The first term represents the deviation between the original timetable and the rescheduled timetable. The second term denotes the total number of passengers left by trains which mainly affected by the actual operational headway of trains. The last term means the total energy consumption of trains, which is mainly influenced by the running time of trains. For metro systems in over-crowded conditions, a delay may cause passengers to strand on the platform while the accumulating passengers will increase the dwell time thus exaggerate the delay. In other words, the total accumulated delay is positively related to the number of stranded passengers while negatively related to energy consumption. In this research, we give the same weight for reducing the total delay and stranded passengers, which also is in conformity with the preferences of dispatchers and actual requirements in over-crowded conditions.

The delay scenario is the same as in the first experiment, and we consider five cases with different weights in the objective function. As shown in Table VI, the weights for the total delay and stranded passengers are decreasing from case 1 to case 5, while the weights for total energy consumption are increasing from case 1 to case 5.

We can observe from Table VI that the total energy consumption of trains is reduced with increasing weights \(\gamma\) while the total delays are increased from case 1 to case 5. This is because the lower operation levels that correspond to a longer running time are selected to save energy, as illustrated in Fig. 7 and Fig. 8. These two figures visualize the selection of train operation levels during the rescheduling process. It can be seen from the comparison that the higher the weight for energy consumption is, the more lower operation levels (numerical

<table>
<thead>
<tr>
<th>Weights</th>
<th>(T_{\text{delay}}) (s)</th>
<th>(N_{\text{stranded}}) (pax)</th>
<th>(E_{\text{total}}) (J)</th>
<th>Optimal objective value</th>
<th>Computation time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. (\alpha=0.5, \beta=0.5, \gamma=0.0)</td>
<td>1482</td>
<td>1006</td>
<td>2.213 \times 10^9</td>
<td>0.674</td>
<td>7.58</td>
</tr>
<tr>
<td>2. (\alpha=0.4, \beta=0.4, \gamma=0.2)</td>
<td>1570</td>
<td>1124</td>
<td>2.112 \times 10^9</td>
<td>0.788</td>
<td>7.64</td>
</tr>
<tr>
<td>3. (\alpha=0.3, \beta=0.3, \gamma=0.4)</td>
<td>1735</td>
<td>1286</td>
<td>1.892 \times 10^9</td>
<td>0.855</td>
<td>7.55</td>
</tr>
<tr>
<td>4. (\alpha=0.2, \beta=0.2, \gamma=0.6)</td>
<td>1955</td>
<td>1432</td>
<td>1.763 \times 10^9</td>
<td>0.874</td>
<td>7.58</td>
</tr>
<tr>
<td>5. (\alpha=0.1, \beta=0.1, \gamma=0.8)</td>
<td>3160</td>
<td>2189</td>
<td>1.552 \times 10^9</td>
<td>0.884</td>
<td>7.64</td>
</tr>
</tbody>
</table>
higher) are selected. Additionally, we can see that there is a positive correlation between the stranded passengers and the total delay according to the computational results in Table VI. Thus, a trade-off among the total delay, stranded passengers and energy consumption can be obtained by choosing proper weights. For practical reasons only one solution shall be chosen according to the preference of the decision maker.

C. Experiment 3: Different Delay Scenarios

In the third experiment, several scenarios were tested with different delay occurrence locations and durations. We use \( d_t, d_l \) and \( t_d \) to denote the train, duration and location of the disturbance, respectively. For example, \( d_t=5, d_l=3 \) and \( t_d=70 \) mean that train 5 is delayed at station 3 for 70 seconds. The weight coefficients are set as the same as in the first experiment. The computational results of different scenarios with respect to the proposed model and FRM are illustrated in Table VII. The results show that the total accumulated delay and stranded passengers can be efficiently reduced by the proposed model. It can be seen from the comparison between scenario 1 and 2 that the total delay is similar, the number of stranded passengers of scenario 2 is less than in scenario 1, while the situation of the energy consumption is reverse, this is because the mass of on-board passengers is taken into account to calculate the energy consumption in this paper. Without considering the dynamic passenger flow, the computational results of scenario 1 and 3 should be same since the space-distance diagram corresponding to the original timetable is parallel as shown in Fig. 3. In practice, the dynamic passenger demand is a main factor to affect the metro train timetable rescheduling results. Thus, the proposed model considering the dynamic passenger flow is more realistic. It can be observed that the energy consumption of trains is increased with the extension of the delay duration. This is because the longer the delay duration, the more affected trains need to shorten the running time in sections to reduce the impact of the delay. The computation times of these five scenarios using CPLEX are only about 6 to 8 seconds, which satisfy the practical requirements of metro train timetable rescheduling in cases of disturbance.
In this paper, we solved the metro train timetable rescheduling problem in cases of disturbance. By introducing binary variables to select the train speed profile and taking account of the dynamic passenger flow and limited train capacity, we proposed a novel mixed integer programming (MIP) model to generate trade-off solutions between the total train delay, stranded passengers and energy consumption of trains from a global perspective. Three numerical experiments based on real operational data were carried out and solved by the state of the art optimization software CPLEX. The simulation results showed that the proposed method can efficiently reduce the total accumulated train delay and stranded passengers compared with a heuristic method derived from a commonly used strategy in practice. The computational time to generate trade-off solutions by the proposed method was around 8 seconds, which meets the real-time requirement for metro train timetable rescheduling to deal with small-scale delays.

In future research, the connections with other lines and the transfer passengers should be investigated from the metro network perspective. In over-crowded conditions, the passenger boarding limiting strategies are carried out in many metro systems to reduce the stranded passengers, which will be studied in our research. Furthermore, we will develop a stochastic programming model to cope with the uncertain delay durations.

REFERENCES


[33] Y. Zhou, Y. Bai, J. Li, B. Mao, and T. Li, “Integrated optimization on train control and timetable to minimize net energy consumption of metro lines,” 2018.


Zhuopu Hou received the B.S. degree in automation from Tianjin Polytechnic University, Tianjin, China, in 2014.

He is currently working towards the Ph.D. degree with State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University. He is currently a Visiting Ph.D. Student with the Center for Railway Research and Education, School of Electronic, Electrical and Computer Engineering, University of Birmingham, Birmingham, U.K.

His current research interests include railway traffic management and optimization techniques.

Hairong Dong (SM’12) received the B.S. and M.S. degrees in automatic control and basic mathematics from Zhengzhou University, Zhengzhou, China, in 1996 and 1999, respectively, and the Ph.D. degree in general and fundamental mechanics from Peking University, Beijing, China, in 2002.

She is currently a Professor with the State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University. She was a Visiting Scholar with the University of Southampton, Southampton, U.K., in 2006; the University of Hong Kong, Pokfulam, Hong Kong, in 2008; the City University of Hong Kong, Kowloon, Hong Kong, in 2009; and the Hong Kong Polytechnic University, Kowloon, in 2010. KTH Royal Institute of Technology, Sweden, in 2011, as well as University of Birmingham, UK, in 2013, and so on.

Prof. Dong is a Senior Member of IEEE, a Member of IEEE Intelligent Transportation Systems Society, the chair of Technical Committee on Railroad Systems and Applications, and the vice general secretary of Chinese Automation Congress.

Her research interests include intelligent transportation systems, automatic train operation, cooperative control of multiple trains, driver assistance system, parallel control and management for high-speed railway systems, fault diagnosis and fault-tolerant control for high-speed trains. She is the Chair of the Technical Committee on Railroad Systems and Applications of the IEEE Intelligent Transportation Systems Society.

Shigen Gao received the B.S. degree in Electrical Engineering and Automation from Tianjin University of Technology, Tianjin, China in 2009, and the Ph.D. degree in Traffic Information Engineering and Control from the State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing, China in 2016.

He is currently an Associate Professor with the State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University. He was a research assistant in the City University of Hong Kong, Hong Kong SAR, in 2011, and visiting PhD in the University of Birmingham, UK, in 2013. His primary research interests are nonlinear control, neural adaptive control, and their applications.

Gemma Nicholson received the MMath and M.Sc. degrees in mathematics from University of Bath, Bath, U.K., in 2004 and 2005, respectively. She received her Ph.D. in 2009 for research into the design of microwave bandpass filters using optimization algorithms.

She is currently a Research Fellow with the Birmingham Centre for Railway Research and Education. Her research interests lie broadly in the application of mathematical techniques to a range of applied research problems using modeling, simulation, and signal processing techniques. In particular, she is involved in the visualization and performance evaluation of traffic management and timetabling methods. She is a Visiting Scholar with the University of Southampton, Hampshire, UK, in 2011.

Lei Chen received the B.S. degree in automation engineering from Beijing Jiaotong University, Beijing, China, in 2005 and the Ph.D. degree in civil engineering from University of Birmingham, UK, in 2012.

He is currently a Birmingham Fellow (research focused Lecturer) in railway traffic management and train control. He leads the Railway Control and Simulation Subgroup at the Birmingham Centre for Railway Research and Education where his research work is delivered. His research interests mainly cover railway traffic management and train control, railway simulation and testing, and railway safety critical system design.

Clive Roberts received the Ph.D. degree from the University of Birmingham, Birmingham, U.K. His Ph.D. dissertation was on condition monitoring of railway infrastructure.

He is currently a Professor of railway systems at the University of Birmingham and Director of the Birmingham Centre for Railway Research and Education. Over the last 14 years he has developed a broad portfolio of research aimed at improving the performance of railway systems. He leads the University’s contribution in a number of large EPSRC, European Commission and industry funded projects. He works extensively with the railway industry in Britain and overseas. He currently leads a team of 11 research fellows and 16 Ph.D. students.