Timetable Recovery After Disturbances in Metro Operations
An Exact and Efficient Solution

Gkiotsalitis, Konstantinos; Cats, Oded

DOI
10.1109/TITS.2020.3041151

Publication date
2020

Document Version
Final published version

Published in
IEEE Transactions on Intelligent Transportation Systems

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Timetable Recovery After Disturbances in Metro Operations: An Exact and Efficient Solution

Konstantinos Gkiotsalitis, *Member, IEEE*, and Oded Cats, *Member, IEEE*

***Abstract*— This study proposes an exact model for timetable recovery after disturbances in the context of high-frequency public transport services. The objective of our model is the minimization of the deviation between the actual headway and the respective planned value. The resulting mathematical program for the rescheduling problem is nonlinear and non-smooth; thus, it cannot be solved to optimality. To rectify this, we reformulate the model using slack variables. The reformulated model can be solved to global optimality in real-time with quadratic programming. We apply the model to real data from the red metro line in Washington D.C. in a series of experiments. In our experiments, we investigate how many upstream trips should be rescheduled to respond to a service disturbance. Our findings demonstrate an improvement potential of service regularity of up to 30% if we reschedule the five upstream trips of a disturbed train.**

***Index Terms*— Timetabling, high-frequency services, disturbance management, metro recovery, regularity-based services.**

I. Introduction

THE planning process of metro services consists of a sequence of stages: strategic (determination of stops and routes), tactical (frequency settings, timetable design, crew and vehicle schedules), and operational (timetable rescheduling, holding, short-turning/stop-skipping) [1]–[3]. Due to the discrepancy between planning and operations, a plan made at the tactical stage might need to be modified at the operational stage. This is often performed so as to maintain a regular service in order to mitigate the impacts of travel or dwell time disturbances (see [4], [5]). A failure to re-plan the timetable after a disturbance is likely to lead to increased passenger waiting times and schedule sliding [6].

Given that metro lines operate in dense urban areas, they typically operate under regularity-based schemes that aim at maintaining the planned headway between successive trips [7]. To achieve this, timetables are developed while considering their robustness to travel time variations. Notwithstanding, these timetable are subject to frequent rescheduling in an attempt to adapt to operational disturbances [8]. The real-time rescheduling of services is considered in the context of *disturbance* management and differs from *disruption* management. The latter deals with large incidents (e.g., station or track closures) which involve service cancellations [9]–[11]. Disturbances are typically addressed by applying local timetable amendments, while disruptions require more involved control measures, such as the cancellation of trips, re-routing, or short-turnings [12], [13].

In this study, we focus on the problem of disturbance management and propose a model that can efficiently react to disturbances and re-plan the dispatching times of trips in real-time, where the dispatching time of a trip is the time it departs the first station of the line. Unlike past works that resort into manual control, rule-based approaches or heuristics to re-plan a timetable in the event of disturbances, we introduce in this work an easy-to-solve mathematical program that can be solved to global optimality in real-time. This facilitates the disturbance management process and improves the regularity of high-frequency metro lines.

The remainder of this study is structured as follows: in section II we discuss the relevant literature and position the contribution of our work. In section III, we formulate our problem and introduce the objectives and constraints of our main mathematical model. Thereafter, the mathematical model is presented in section IV and is reformulated to ensure its feasibility after relaxing its soft constraints. This mathematical model is non-smooth and its objective function is not differentiable at every point of its domain - prohibiting the application of an exact solution method. To rectify this, in section V we propose a model reformulation with the introduction of slack variables. The reformulated program is proven to have a globally optimal solution and can be easily solved with exact optimization methods. Next, in section VI, the application of our approach to the red metro line in Washington D.C. is presented, demonstrating that we can achieve a significant benefit if we reschedule the dispatching times of up to 5 upstream trips of a train that exhibit a disturbance. Finally, we provide in section VII concluding remarks and discuss future research directions.

II. Related Studies

In this study, we focus on disturbance management. Since extensive disruptions and disruption management are not the primary focus of our work, we refer the interested reader to the comprehensive literature study of [14] on disruption management. We also refer to the works of [15] and [16] who devised methods for reliable disruption length estimation. Such estimations can then be used as part of a mixed integer linear programming (MILP) for making real-time decisions on skipping/adding stops or performing short-turnings when a

Manuscript received November 27, 2019; revised June 17, 2020; accepted November 23, 2020. The Associate Editor for this article was W. Jin. *(Corresponding author: Konstantinos Gkiotsalitis.)*

Konstantinos Gkiotsalitis is with the Department of Civil Engineering, University of Twente, 7500 Enschede, The Netherlands (e-mail: k.gkiotsalitis@utwente.nl).

Oded Cats is with the Department of Transport and Planning, Technical University of Delft, 2600 Delft, The Netherlands (e-mail: o.cats@tudelft.nl).

Digital Object Identifier 10.1109/TITS.2020.3041151

disruption occurs [7] or devising contingency plans by making alterations to the original timetable [17].

The aim of this study is to produce an efficient model that can react to disturbances and re-set the dispatching times of trips in real-time. Solutions to this problem commonly adopt local rescheduling to adjust the timetable, see [18]–[20]. D'Ariano *et al.* [18] aimed at improving the punctuality of trains by routing and sequencing trains in an iterative manner - first, an optimal train sequencing was produced for the given train routes, and thereafter this solution was improved by locally rerouting some trains. Their solution method was based on local search and Branch and Bound (B&B) given the discrete nature of their mathematical problem. This work was extended in [19] by incorporating effective rescheduling algorithms and local rerouting strategies in a Tabu search scheme. Corman *et al.* [19] alternated between a fast heuristic and a truncated B&B algorithm for computing train schedules. Their train schedules were computed within a short computation time, but their solution method did not guarantee the convergence to a globally optimal solution. Pellegrini *et al.* [21] aimed at minimizing delays after an unexpected disturbance perturbs the operations by searching for the best train routing and scheduling. The proposed model was a mixed integer linear program, representing the infrastructure with fine granularity. Solving the model of [21] to global optimality is not feasible in real-time for most of the cases, and the reported computational costs were typically beyond one and a half minute, even with the use of heuristics.

Most relevant to the problem addressed in this study, [22] focuses solely on the timetable re-setting after a disturbance. In [22], the rescheduling problem was not solved to global optimality. Instead, a greedy heuristic was introduced to ensure that a (hopefully) good-enough solution is obtained within a short time (i.e. within 30 seconds). To this end, [22] introduced a heuristic solution method without formulating the timetable rescheduling problem as a mathematical program.

Apart from the works on disturbance management for rail operations, several works have addressed this problem in the context of bus operations that operate in mixed-traffic environments, e.g. [23] or [24] who developed a library of operational tactics to avoid bunching. Bus operators apply dynamic control strategies such as stop-skipping [25]–[28], holding for single-line reliability and cross-line transfer synchronization [29]–[33], or rescheduling [34], [35] in the event of disturbances. A combination of holding and speed changing is also common, as presented in the work of [36] which focused on autonomous vehicles. Nevertheless, rescheduling typically resorts to rule-based techniques or meta-heuristics to obtain a solution in real-time given the computational complexity of the problem [37].

As evident from the above review of the literature, since most works on disturbance management formulate complex mixed-integer programs (MIPs) that cannot be solved to global optimality, they resort to heuristic solution methods to solve the respective MIPs. To fill this research gap, in this study we propose a novel quadratic programming formulation for the local timetable re-setting problem in the case of disturbance(s). Our proposed mathematical program is proved to be convex and can re-compute a timetable each time a disturbance occurs. It can also provide a fast solution even at large problem instances due to its quadratic formulation. The contributions of our work to the state-of-the-art are:

- a regularity-based mathematical program for service rescheduling to mitigate the impacts of disturbances that can be solved in real-time;
- a problem-specific formulation and a solution approach that guarantees convergence towards a globally optimal solution;
- the investigation of how many trips need to be rescheduled after a disturbance occurs with the use of real data from a metro operator.

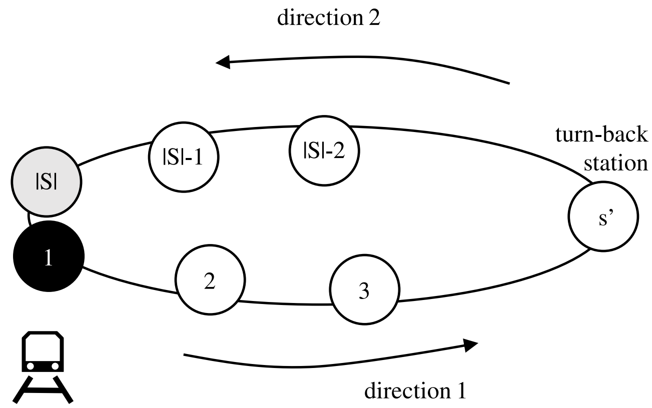

Fig. 1. Illustration of loop-formed metro line.

III. Problem Definition

A. Trip Re-Indexing After a Disturbance

We consider a metro line that serves a set of ordered stations $S = \langle 1, 2, \ldots, s', \ldots, |S|\rangle$ with 1 being the dispatching station and s' the turn-back station (see Fig.1). Note that we consider here the case of a double-track corridor where train traffic for the two directions can be assumed independent. The ordered set of daily trips operating in this line is $\mathbf{N}$. When the operation of trip $m \in \mathbf{N}$ is disturbed, the set of its subsequent trips (henceforth referred to as *upstream* trips) is $\mathbf{N}_m = \{m+1, m+2, \ldots\}$ with $\mathbf{N}_m \subset \mathbf{N}$. To alleviate the effects of the disturbance, the dispatching times of all trips $j \in \mathbf{N}_m$ are re-set by performing a headway-based optimization process as described in the following.

The choice of the length of set $\mathbf{N}_m$ has significant practical implications. For instance, if $\mathbf{N}_m = \{m+1, m+2, \ldots\}$ contains all remaining daily trips, the computation cost of re-setting their dispatching times increases due to the increased number of decision variables. In addition to the increase in computation costs, modifying the dispatching times of trips that are expected to be dispatched in the far future might be proved unnecessary if further disturbances occur in the near future and there is a need to re-optimize. For this reason, the sensitivity of the solution performance to the number of rescheduled upstream trips, $\mathbf{N}_m$, is investigated in our case study (this is presented in section VI).

Let us re-index set $\mathbf{N}_m = \{m+1, m+2, \ldots\}$ into $\mathbf{N}_m = \{0, 1, \ldots, n\}$ where trip 0 has already been dispatched and exhibits a disturbance (e.g., unexpected travel time), and n is the last trip in $\mathbf{N}_m$. Thus, trips 0 and $n+1$ are the "boundaries" of our re-timing problem because their dispatching times cannot be modified since: (i) trip 0 has already been dispatched, and (ii) trip $n+1$ does not belong to the set of upstream trips that need to be re-timed.

Let $\delta_j \in \mathbb{R}_+$ denote the originally planned dispatching time of each trip $j \in \{1, 2, \ldots, n\}$ (the time reported at the planned timetable). Then, $\delta_1 < \delta_2 < \ldots < \delta_j < \ldots < \delta_n$. The decision variable is an n-valued vector $\mathbf{x} = \{x_1, x_2, \ldots, x_n\}$ which expresses the dispatching time offset for each trip $j \in \mathbf{N}_m$, where $\mathbf{x} \in \mathbb{R}^n$. Thus, the adjusted dispatching times of trips $\{1, 2, \ldots, n\}$ are $\{\delta_1 + x_1, \delta_2 + x_2, \ldots, \delta_n + x_n\}$.

Before introducing the vehicle motion law that governs the movement of trains, we list the main assumptions used in the formulation of our mathematical model:

1) Dwell times at stations are fixed. This can be due to negligible observed variations in dwell times or in case this is dictated by the operational policy, for instance if the doors close automatically when the allotted dwell time elapsed due to the control principles of an automated passenger door system (see [38]).
2) Service supply is determined at the frequency settings stage and ensures that the passenger demand can be accommodated even at the maximum load point [39]–[42]. That is, dispatching time changes will not lead to overcrowding because the pre-planned service supply includes sufficient margins.

B. Vehicle Trajectories

With assumptions 1-2, we can outline a set of rules governing the vehicle movements. For this, we briefly introduce the notation in Table I.

TABLE I

NOMENCLATURE

Sets	
$S = \{1, ..., s', ...\}$	ordered set of consecutive stations of the metro line, where s' is the turn-back station;
$\mathbf{N}_m = \{1, 2, ..., n\}$	ordered set of future trips under re-timing consideration when trip 0 is disturbed;
Indices	
j	index of vehicles;
s	index of stations;
Parameters	
h_{min}, h_{max}	minimum and maximum headways (agency's requirements) between two successive dispatches from the first station
h'_{min}	minimum headway at the turn-back station to comply with its vehicle capacity limitations
β_j	latest possible dispatching time of any trip $j \in N_m$ to avoid schedule sliding and/or disrupting crew schedules
π_j	the earliest possible dispatching time of any trip $j \in N_m$ to ensure vehicle circulation
δ_j	originally planned dispatching time of trip $j \in N_m$
h^*_{js}	scheduled (target) headway of trip $j \in N_m$ at station $s \in \{2, 3, ..., \lvert S\rvert - 1\}$
$t_{j,s}$	expected inter-station travel time of trip $j \in N_m$ from station s to station $s+1$
$k_{j,s}$	pre-determined dwell time of trip $j \in N_m$ at station $s \in \{2, 3, ..., \lvert S\rvert\}$ after which the doors close automatically
$\bar{a}_{0,s}$	the realized arrival times of the disturbed trip 0 at stations $s = \{2, 3, ..., \lvert S\rvert\}$
$\bar{\delta}_0$	the realized dispatching time of trip 0
$c_{j,s}$	arrival time of trip j at station s if we follow the originally planned dispatching time of trip j
Decision Variables	
$\{x_1, ..., x_n\}$	the dispatching time offsets of trips $j \in \{1, 2, ..., n\}$ from δ_j;
Variables	
$a_{j,s}$	arrival time of trip j at station s, where $s \in \{2, 3, ..., \lvert S\rvert\}$
h_{js}	headway between trips j and $j-1$ at station s.

The expected arrival time $a_{j,s}$ of a trip $j \in \{1, 2, \ldots, n\}$ at station $s \in \{2, 3, \ldots, |S|\}$ is

$$a_{j,s} := (\delta_j + x_j) + \sum_{\phi=1}^{s-1} \tau_{j,\phi} + \sum_{\phi=2}^{s-1} k_{j,\phi} \tag{1}$$

where δ_j is the planned dispatching time, x_j the dispatching offset, $\tau_{j,\phi}$ the travel time from station ϕ to $\phi + 1$ and $k_{j,\phi}$ the dwell time at station ϕ. In addition, $\sum_{\phi=1}^{s-1} \tau_{j,\phi}$ is the total travel time from the first station until station s and $\sum_{\phi=2}^{s-1} k_{j,\phi}$ is the accumulated dwell times from stations $2, 3, \ldots, s-1$. Note that the dwell times are aggregated starting from station 2 because the departure time from station 1 is $\delta_j + x_j$.

From Eq.(1), the arrival time of each trip $j \in N_m$ at each station s varies according to the decision variable values of x_j. Therefore, Eq.(1) can be succinctly written as:

$$a_{j,s} := x_j + c_{j,s}, \quad \forall j \in N_m, \ \forall s \in \{2, 3, \ldots, |S|\} \tag{2}$$

where

$$c_{j,s} := \delta_j + \sum_{\phi=1}^{s-1} \tau_{j,\phi} + \sum_{\phi=2}^{s-1} k_{j,\phi}, \quad \forall j \in N_m, \quad \forall s \in \{2, 3, \ldots, |S|\} \tag{3}$$

The inter-arrival time headway $h_{j,s}$ of two successive trips $j-1$ and $j \in N_m \setminus \{1\}$ at the time of their arrival at station $s \in S \setminus \{1\}$ is

$$h_{j,s} := a_{j,s} - a_{j-1,s} = (x_j + c_{j,s}) - (x_{j-1} - c_{j-1,s}), \quad \forall j \in \{2, 3, \ldots, n\}, \quad \forall s \in \{2, 3, \ldots, |S|\} \tag{4}$$

C. Boundary Condition

Note that the arrival times $\bar{a}_{0,s}$ of trip 0 at stations $s \in \{2, \ldots, |S|\}$ are not affected by the decision variables since trip 0 has already been dispatched. Incorporating this initial condition, the time headway between trip 1 and its preceding one, 0, is:

$$h_{1,s} := x_j + c_{j,s} - \bar{a}_{0,s}, \quad \forall\ s \in \{2, 3, \ldots, |S|\} \tag{5}$$

Thus, Eq.(5) links the time headway between trip 1 and 0 with the dispatching time offset of trip 1, x_1.

D. Constraints

Firstly, a constraint is imposed by the latest possible dispatching time of a trip. While some works compute the vehicle and crew schedules together with the departure times of trips [43], most works treat them separately (see [22]). Therefore, it is not practical to delay the dispatching time of a trip further than a certain threshold because that would lead to the sliding of the schedule with possible ramifications for vehicle and crew schedules. If β_j is the pre-planned latest possible dispatching time set by operators based on the additional costs associated with the sliding of the timetable (e.g. overtime labor costs, contractual tender commitments), then

$$\delta_j + x_j \leq \beta_j, \quad \forall j \in N_m \tag{6}$$

At this point, we should note that this constraint cannot be always satisfied. Namely, in cases of external disruptions there might be no x_j that can guarantee the adherence to the pre-planned dispatching threshold β_j. We therefore choose to specify Eq.(6) as a soft constraint that we aim to satisfy whenever possible and it will be later replaced by a penalty term.

Secondly, agencies have specific requirements on the minimum and maximum allowable dispatching headway, h_{min}, h_{max}, to ensure a minimum level of service [44]. This constraint can be expressed as:

$$h_{min} \leq (\delta_j + x_j) - (\delta_{j-1} + x_{j-1}) \leq h_{max} \quad \forall j \in \{2, 3, \ldots, n\} \tag{7}$$

and, in the boundary case where $j = 1$,

$$h_{min} \leq (\delta_j + x_j) - \bar{\delta}_0 \leq h_{max} \tag{8}$$

where $\bar{\delta}_0$ is the realized dispatching time of trip 0.

In case that the vehicle capacity at the turn-back station s' is limited, one can impose a pre-defined minimum headway h'_{min} that should be maintained among all vehicles that arrive at the turn-back station in order to comply with its capacity limitations. This results in the following set of constraints:

$$a_{j,s'} - a_{j-1,s'} \geq h'_{min}, \quad \forall j \in N_m \tag{9}$$

Lastly, to ensure the circulation of vehicles, each trip j should depart after time π_j - where π_j is the time by which the vehicle and driver assigned to perform trip j have completed their previous trip and are ready to start trip j. The vehicle circulation constraint is modeled as:

$$\delta_j + x_j \geq \pi_j \quad \forall j \in \{1, 2, \ldots, n\} \tag{10}$$

This constraint ensures that a vehicle is available to operate the respective trip.

E. Objective Function of Our Model

In high-frequency services, each one of the trips $\{1, 2, \ldots, n\}$ has a target headway $h^*_{j,s}$ in relation to its leading train at any station $s \in \{2, \ldots, |S| - 1\}$ [45]. The target headway is determined at the tactical planning stage and should be maintained during the daily operations [46].

When striving to maintain the target (ideal) headway, the objective is to minimize the inter-arrival headway variability around the target values because this results in a reduction of passenger waiting times in high-frequency services (see [47]). To achieve that, the optimal dispatching offset $x = \{x_1, x_2, \ldots, x_n\}$ should be the solution of:

$$\min_h \sum_{s=2}^{|S|-1} \sum_{j=1}^{n} \left(h_{j,s} - h^*_{j,s}\right)^2 \tag{11}$$

which expresses the discrepancy of the inter-arrival headways in relation to their target values measured as the sum of squared errors.

Eq.(11) can be equivalently expressed as:

$$\begin{aligned} &\min_x f(x) \\ &:= \sum_{s=2}^{|S|-1} \left((x_1 + c_{1,s}) - \bar{a}_{0,s} - h^*_{1,s}\right)^2 \\ &\quad + \sum_{s=2}^{|S|-1} \sum_{j=2}^{n} \left((x_j + c_{j,s}) - (x_{j-1} + c_{j-1,s}) - h^*_{j,s}\right)^2 \end{aligned} \tag{12}$$

IV. Mathematical Program

Combining the expected trajectories of future trips and the objective function yields our continuous mathematical program:

$$\begin{aligned} (Q):\ \min_x\ & \sum_{s=2}^{|S|-1} \left((x_1 + c_{1,s}) - \bar{a}_{0,s} - h^*_{1,s}\right)^2 \\ & + \sum_{s=2}^{|S|-1} \sum_{j=2}^{n} \left((x_j + c_{j,s}) - (x_{j-1} + c_{j-1,s}) - h^*_{j,s}\right)^2 \\ \text{s.t. } & h_{min} \leq (\delta_j + x_j) - (\delta_{j-1} + x_{j-1}), \quad \forall j \in \{2, \ldots, n\} \\ & (\delta_j + x_j) - (\delta_{j-1} + x_{j-1}) \leq h_{max}, \quad \forall j \in \{2, \ldots, n\} \\ & h_{min} \leq (\delta_1 + x_1) - \bar{\delta}_0 \\ & (\delta_1 + x_1) - \bar{\delta}_0 \leq h_{max} \\ & a_{j,s'} - a_{j-1,s'} \geq h'_{min}, \quad \forall j \in \{1, 2, \ldots, n\} \\ & (\delta_j + x_j) \leq \beta_j, \quad \forall j \in \{1, 2, \ldots, n\} \\ & (\delta_j + x_j) \geq \pi_j, \quad \forall j \in \{1, 2, \ldots, n\} \\ & c_{j,s} = \delta_j + \sum_{\phi=1}^{s-1} \tau_{j,\phi} + \sum_{\phi=2}^{s-1} k_{j,\phi}, \quad \forall j \in N_m, \\ & \qquad \forall s \in S \setminus \{1\} \\ & x_j \in \mathbb{R}, \quad \forall j \in \{1, 2, \ldots, n\} \end{aligned} \tag{13}$$

At this point we would like to note that our model (Q) can be amended to modify the departure times of buses from any station by considering vehicle holding. In vehicle holding models, this is achieved by considering an additional decision variable $\tilde{x}_{j,s}$ that determines the holding time of any trip j at any station $s \in \{2, \ldots, |S| - 1\}$ that should be added to the respective dwell time $k_{j,s}$. This holding time should be always positive, $x_{i,j} \in \mathbb{R}_+$, and can only delay the completion of a trip that might result in schedule sliding and the delayed dispatching of future trips. In addition, specific attention should be given to maintain a safety distance among trains when implementing holding by imposing a constraint to ensure a minimum required headway among all trips at all stations.

A. Infeasibility and Soft Constraints

Program (Q) can be succinctly written as:

$$(Q): \min_{x} f(x) \quad \text{s.t. } x \in \mathcal{F} := \{x \mid x \text{ satisfy Eqs.(6)-(10), (12)}\} \tag{14}$$

where $\mathcal{F}$ is the feasible set. Note that the inequality constraints of Eqs.(6)-(8) cannot be always satisfied at the same time and some of them should be prioritized at the expense of others. Indeed, program (Q) might not have a feasible solution for some values of the inequality constraints of Eqs.(6)-(8) yielding an empty feasibility set $\mathcal{F}$ (refer to Lemma A.1 in the Appendix). To rectify this, we relax the inequality constraints of schedule sliding presented in Eq.(6). With this relaxation, they become soft constraints that may be violated under certain circumstances. As soft constraints, they are treated as penalty terms and they are added to the objective function (see [48]). In this way, program (Q) that, under certain circumstances, has no feasible solution is transformed to program $(\bar{Q})$.

This approach ensures that the constraints of Eq.(6) are satisfied when possible, or violated as little as possible when there is otherwise no feasible solution. Their relative importance is weighted by introducing a very large number $M \in \mathbb{R}_{\geq 0}$ which ensures that the satisfaction of the schedule sliding constraints is prioritized:

$$(\bar{Q}): \min_{x} f(x) + \sum_{j \in N_m} M \max(\delta_j + x_j - \beta_j, 0) \quad \text{s.t. } x \in \mathcal{F} := \{x \mid x \text{ satisfy Eqs.(7)-(10), (12)}\} \tag{15}$$

The penalty term $M \max(\delta_j + x_j - \beta_j, 0)$ ensures that the soft constraint $\delta_j + x_j \leq \beta_j$ is prioritized over $f(x)$. Indeed, if $\delta_j + x_j \leq \beta_j$ for some x_j, then x_j does not add any penalty to the objective function since $M \max(\delta_j + x_j - \beta_j, 0) = 0$. In reverse, when $\delta_j + x_j > \beta_j$ for some x_j, then the penalty term penalizes the objective function by a very large number $M \max(\delta_j + x_j - \beta_j)$ and directs the solution search towards another solution that reduces the value of $M \max(\delta_j + x_j - \beta_j, 0)$.

B. Properties of the Mathematical Program

Program $(\bar{Q})$ is a nonlinear programming (NLP) problem. Additionally, our new objective is a non-smooth function because of the non-smooth term $\sum_{j \in N_m} M \max(\delta_j + x_j - \beta_j, 0)$; hence, the objective function of $\bar{Q}$ is not differentiable at every point of its domain. This results in a non-linear, non-convex function that cannot be solved to global optimality with exact optimization methods. As a remedy, we propose a reformulation to cast the problem as an easier-to-solve quadratic program.

V. Reformulation to a Quadratic Program and Exact Solution

The "max" term of $\sum_{j \in N_m} M \max(\delta_j + x_j - \beta_j, 0)$ makes the objective function of program $\bar{Q}$ non-smooth. As a remedy, we convert the "max" penalty into a new set of slack variables $\nu_j,\ j \in N_m$ that, due to their bounds and the direction of optimization, will take the value $\sum_{j \in N_m} M \max(\delta_j + x_j - \beta_j, 0)$ at the solution. The reformulated program is:

$$\begin{aligned}(\tilde{Q}): \min_{x,\nu}\ & f(x) + \sum_{j \in N_m} M\nu_j \\ \text{s.t. } & x \in \mathcal{F} := \{x \mid x \text{ satisfy Eqs.(7)-(10), (12)}\} \\ & \nu_j \geq 0, \quad \forall j \in N_m \\ & \nu_j \geq \delta_j + x_j - \beta_j, \quad \forall j \in N_m \end{aligned} \tag{16}$$

which is reduced to a quadratic program (QP). As shown in Theorem 1, program $(\tilde{Q})$ is strictly convex and can be easily solved to global optimality since any locally optimal solution returned by a quadratic programming solver is also a globally optimal one.

Theorem 1: A local optimum of program $(\tilde{Q})$ is also a globally optimal solution

Proof: A local minimizer of $\tilde{Q}$ is the global minimizer of $\tilde{Q}$ if the objective function is strictly convex and the feasible region is a convex set. The feasible region is defined by linear (in)equalities (affine functions) and is a polyhedron. Thus, it is also a convex set. Further, we prove that the objective function $f(x) + \sum_{j \in N_m} M\nu_j$ is strictly convex with respect to x, ν.

Let $\tilde{f}(x, \nu) := f(x) + \sum_{j \in N_m} M\nu_j$. Then, the Hessian matrix of $\tilde{f}(x, \nu)$ is a matrix $\mathbf{H} \in \mathbb{R}^{2n \times 2n}$ with elements:

$$\mathbf{H} = \begin{bmatrix} \frac{\partial^2 \tilde{f}(x,\nu)}{\partial x_1^2} & \frac{\partial^2 \tilde{f}(x,\nu)}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 \tilde{f}(x,\nu)}{\partial x_1 \partial x_n} & \frac{\partial^2 \tilde{f}(x,\nu)}{\partial x_1 \partial \nu_1} & \cdots & \frac{\partial^2 \tilde{f}(x,\nu)}{\partial x_1 \partial \nu_n} \\ \frac{\partial^2 \tilde{f}(x,\nu)}{\partial x_2 \partial x_1} & \frac{\partial^2 \tilde{f}(x,\nu)}{\partial x_2^2} & \cdots & \frac{\partial^2 \tilde{f}(x,\nu)}{\partial x_2 \partial x_n} & \frac{\partial^2 \tilde{f}(x,\nu)}{\partial x_2 \partial \nu_1} & \cdots & \frac{\partial^2 \tilde{f}(x,\nu)}{\partial x_2 \partial \nu_n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 \tilde{f}(x,\nu)}{\partial x_n \partial x_1} & \frac{\partial^2 \tilde{f}(x,\nu)}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 \tilde{f}(x,\nu)}{\partial x_n^2} & \frac{\partial^2 \tilde{f}(x,\nu)}{\partial x_n \partial \nu_1} & \cdots & \frac{\partial^2 \tilde{f}(x,\nu)}{\partial x_n \partial \nu_n} \\ \frac{\partial^2 \tilde{f}(x,\nu)}{\partial \nu_1 \partial x_1} & \frac{\partial^2 \tilde{f}(x,\nu)}{\partial \nu_1 \partial x_2} & \cdots & \frac{\partial^2 \tilde{f}(x,\nu)}{\partial \nu_1 \partial x_n} & \frac{\partial^2 \tilde{f}(x,\nu)}{\partial \nu_1^2} & \cdots & \frac{\partial^2 \tilde{f}(x,\nu)}{\partial \nu_1 \partial \nu_n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 \tilde{f}(x,\nu)}{\partial \nu_n \partial x_1} & \frac{\partial^2 \tilde{f}(x,\nu)}{\partial \nu_n \partial x_2} & \cdots & \frac{\partial^2 \tilde{f}(x,\nu)}{\partial \nu_n \partial x_n} & \frac{\partial^2 \tilde{f}(x,\nu)}{\partial \nu_n \partial \nu_1} & \cdots & \frac{\partial^2 \tilde{f}(x,\nu)}{\partial \nu_n^2} \end{bmatrix} \tag{17}$$

The gradient of $\tilde{f}(x, \nu)$ is an $\mathbb{R}^{2n}$ vector:

$$\begin{aligned}\nabla \tilde{f}(x, \nu) = \Bigg(& \sum_{s=2}^{|S|-1} (4x_1 - 2x_2 + \rho_1), \\ & \sum_{s=2}^{|S|-1} (4x_2 - 2x_1 - 2x_3 + \rho_2), \\ & \ldots, \sum_{s=2}^{|S|-1} (4x_{n-1} - 2x_{n-1} - 2x_n + \rho_{n-1}), \\ & \sum_{s=2}^{|S|-1} (2x_n - 2x_{n-1} + \rho_n), \underbrace{1, \ldots, 1}_{\text{n}} \Bigg) \end{aligned} \tag{18}$$

where $\rho_1, \rho_2, \ldots, \rho_n$ are parameter values consisting of travel times, dwell times and target headway which do not vary with x or ν.

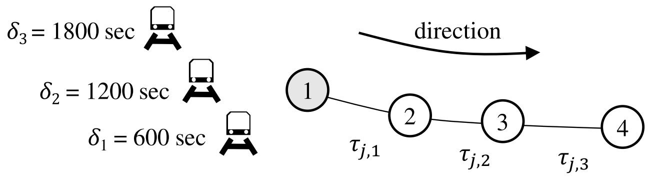

Fig. 2. Toy metro line.

To simplify the notation, let us set $\zeta := |S| - 2$ with $\zeta > 0$. This yields the Hessian:

$$\mathbf{H} = \begin{bmatrix} 4\zeta & -2\zeta & 0 & \dots & 0 & 0 & \dots & 0 \\ -2\zeta & 4\zeta & -2\zeta & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 2\zeta & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \\ 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{bmatrix} \tag{19}$$

$\tilde{f}(x, \nu)$ is strictly convex if the Hessian matrix is positive definite. That is, $\mathbf{z}^{\mathsf{T}}\mathbf{H}\mathbf{z} > 0$ for any non-zero vector $\mathbf{z} \in \mathbb{R}^{2n} \setminus 0$. $\mathbf{z}^{\mathsf{T}}\mathbf{H}\mathbf{z}$ yields:

$$\begin{aligned}\mathbf{z}^{\mathsf{T}}\mathbf{H}\mathbf{z} &= \begin{bmatrix} z_1 & z_2 & \dots & z_{n-1} & z_n \end{bmatrix} \mathbf{H} \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_{n-1} \\ z_n \end{bmatrix} \\ &= \zeta\big(4z_1^2 - 2z_1z_2 - 2z_1z_2 + 4z_2^2 - 2z_2z_3 - \dots \\ &\qquad - 2z_{n-2}z_{n-1} + 4z_{n-1}^2 - 2z_{n-1}z_n - 2z_{n-1}z_n + 2z_n^2\big) \\ &= \zeta\big(4z_1^2 - 4z_1z_2 + 4z_2^2 - 4z_2z_3 + 4z_3^2 - \dots \\ &\qquad + 4z_{n-2}^2 - 4z_{n-2}z_{n-1} + 4z_{n-1}^2 - 4z_{n-1}z_n + 2z_n^2\big) \\ &= \zeta\Big(2z_1^2 + (z_1\sqrt{2} - z_2\sqrt{2})^2 + (z_2\sqrt{2} - z_3\sqrt{2})^2 + \dots \\ &\qquad + (z_{n-2}\sqrt{2} - z_{n-1}\sqrt{2})^2 + (z_{n-1}\sqrt{2} - z_n\sqrt{2})^2\Big)\end{aligned}$$

Hence, $\mathbf{z}^{\mathsf{T}}\mathbf{H}\mathbf{z} > 0$ for any $\mathbf{z} \in \mathbb{R}^{2n} \setminus 0$ and thus $\tilde{f}(x, \nu)$ is *strictly* convex. This proves that a local optimum of the reformulated program $(\tilde{Q})$ is also its *unique* global minimizer. □

VI. Numerical Experiments

A. Demonstration for a Toy Network

To demonstrate the application of our mathematical program $(\tilde{Q})$ for timetable recovery, we introduce a small-scale idealized scenario using a toy network. Trip 0 has already been dispatched at time $\bar{d}_0 = 0$ sec and has exhibited a travel time disturbance. Its arrival times at the four stops included in our toy network are $a_{0,2} = 900$ sec and $a_{0,3} = 1600$ sec. Note that we only report the arrival times at the second and the third station because the service regularity in $f(x)$ is not measured at the first and last stations (Fig.2).

To proceed with the timetable recovery, we allow to modify the dispatching times of its three following trips (namely 1, 2 and 3). The originally planned dispatching times of trips 1, 2 and 3 are: $\delta_1 = 600$ sec, $\delta_2 = 1200$ sec and $\delta_3 = 1800$ sec.

TABLE II

Iterations Until Convergence and Computational Cost of Obtaining the Optimal Dispatching Times With Gurobi

Iteration	$\tilde{f}(x,\nu)$	Computational Time
0	1.70E+09	0 sec
1	7.79E+07	0 sec
2	1.48E+07	0 sec
3	1.41E+05	0 sec
4	3.71E+04	0 sec
5	1.92E+04	0 sec
6	1.04E+04	0 sec
7	8.23E+03	0 sec
8	8.08E+03	0 sec
9	8.08E+03	0 sec
10	8.08E+03	0 sec
Value of global minimum:	8.08E+03	

The expected inter-station travel times of trips 1, 2 and 3 are:

$$\begin{aligned}(\tau_{1,1}, \tau_{1,2}, \tau_{1,3}) &= (900, 720, 800) \text{ sec} \\ (\tau_{2,1}, \tau_{2,2}, \tau_{2,3}) &= (920, 700, 800) \text{ sec} \\ (\tau_{3,1}, \tau_{3,2}, \tau_{3,3}) &= (880, 640, 800) \text{ sec}\end{aligned}$$

The target headway is 10 minutes, thus $h^*_{j,s} = 600$ sec, $\forall j \in \{1, 2, 3\}$, $\forall s \in \{2, 3\}$. In addition, the minimum and maximum dispatching headways are $(h_{min}, h_{max}) = (300$ sec, 900 sec) and, given that Fig.2 does not include turnbacks, $h'_{min} = 0$. The dwell times at metro stations are set to $k_{j,s} = 30$ sec $\forall j \in \{1, 2, 3\}, \forall s \in \{2, 3\}$. Due to vehicle circulation constraints, the earliest possible dispatching times of trips 1,2,3 are $(\pi_1, \pi_2, \pi_3) = (600$ sec, 1220 sec, 1820 sec) in order to ensure vehicle availability. Finally, to avoid schedule sliding, the latest dispatching times of our trips are $(\beta_1, \beta_2, \beta_3) = (660, 1260, 1860)$ sec.

Our mathematical model $(\tilde{Q})$ is programmed in Python 3.7 and the experimental tests are performed in a general-purpose computer with Intel Core i7-455 7700HQ CPU @ 2.80GHz and 16 GB RAM. To solve our model to global optimality, we use Gurobi. To facilitate the reproduction of our methodology, our source code is publicly released at [49] and can be applied to other case studies. Starting from an initial solution guess, Gurobi converged to a globally optimal solution within 10 iterations in this demonstration (see Table II).

The globally optimal solution is:

$$\begin{aligned}x^* &= \{x_1 = 2.5, x_2 = 20, x_3 = 60\} \text{ sec} \\ \nu^* &= \{\nu_1 = 0, \nu_2 = 0, \nu_3 = 0\} \text{ sec}\end{aligned}$$

and the resulting vehicle trajectories when applying this rescheduling solution are presented in Fig.3. Note that the rescheduled operations result in even headways at the locations of the metro stations which are close to the target headway of 600 sec.

In Fig.4 we show how this solution improves the squared headway deviations at stations 2 and 3 in relation to the target value. Fig.4 indicates a potential improvement of up to 47% in terms of service regularity compared to the do-nothing case where rescheduling is not applied.

The average excessive passenger waiting time improvement considering random passenger arrivals at stations given the

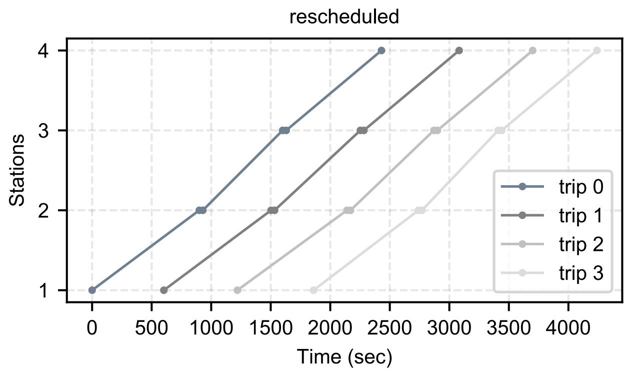

Fig. 3. Train trajectories before and after rescheduling.

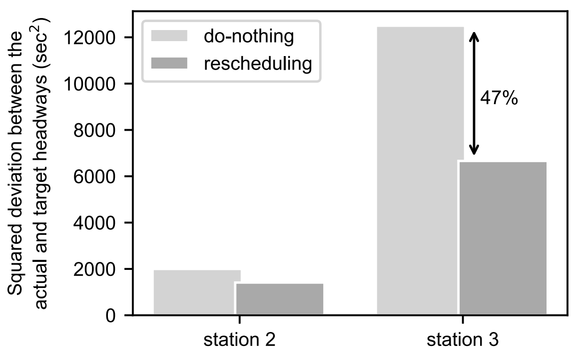

Fig. 4. Squared deviation between the actual and target headway at stations 2 and 3 when our rescheduling solution is applied and in its absence (do-nothing case).

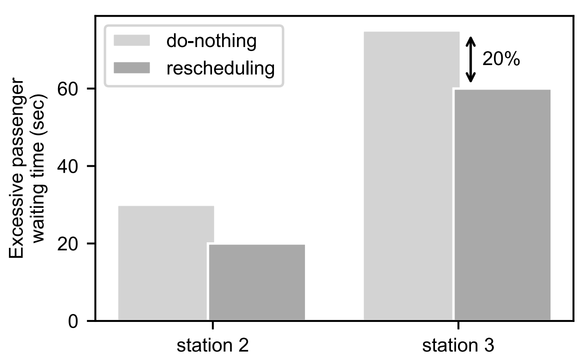

Fig. 5. Excessive passenger waiting times at stations 2 and 3 when our rescheduling solution is applied and in its absence (do-nothing case).

high service frequency is provided in Fig.5. Fig.5 shows the difference between the actual waiting time of passengers and the planned passenger waiting time of $h^*_{j,s}/2 = 300$ sec. As shown in Fig.5, the excessive passenger waiting time improvement is 20% at station 3.

We finally note that if we would not have considered the schedule sliding constraints (that is, $\beta_j = +\infty, \forall j \in \{1, 2, 3\}$), the globally optimal solution would have been:

$$x^* = \{x_1 = 2.5, x_2 = 20, x_3 = 90\} \text{ sec}$$
$$\nu^* = \{\nu_1 = 0, \nu_2 = 0, \nu_3 = 0\} \text{ sec}$$

with a solution performance of 6.28E+03. Hence, if we had additional resources (i.e., reserve trains) to perform the next trips when our dispatching time adjustments result in delays, service regularity would have been further improved by 22.3%.

TABLE III
SOLUTION PERFORMANCE AND COMPUTATIONAL COSTS WHEN SOLVING PROBLEM (Q) AND ITS REFORMULATED VERSION ($\tilde{Q}$)

	derived solution	solution performance	computation time (sec)
solve ($\tilde{Q}$)	$x^* = \{0, 20, 20\}$	4.02E+06	0.04
solve (Q)	$x^* = \{20, 20, 20\}$	6.01E+06	2.21
improvement		33.1%	98.2%

Finally, to demonstrate how our mathematical model handles the case where schedule sliding cannot be avoided, let us consider the same scenario with $(\beta_1, \beta_2, \beta_3) = (600, 1200, 1800)$ sec. Obviously, for such values of β_j the schedule will slide because the earliest possible dispatching times of trips to ensure vehicle circulation are $(\pi_1, \pi_2, \pi_3) = (600, 1220, 1820)$ sec. Hence, our program returns a globally optimal solution that slides the schedule as little as possible:

$$x^* = \{x_1 = 0, x_2 = 20, x_3 = 20\} \text{ sec}$$
$$\nu^* = \{\nu_1 = 0, \nu_2 = 20, \nu_3 = 20\} \text{ sec}$$

with a performance of 4.02E+06. Note that the schedule sliding is indicated by the positive values of ν_2, ν_3 which are always equal to zero when a feasible solution for (Q) exists. Because of its infeasibility, program (Q) cannot be always solved with an exact optimization solver. Therefore, we employ a heuristic to find an approximate solution of program (Q) and we report the performances when solving programs (Q) and ($\tilde{Q}$) for the aforementioned case in Table III. Note that program (Q) is solved with the heuristic differential evolution method described in [50] because it cannot be solved by an exact solver.

B. Case Study and Sensitivity Analysis

We apply our model to the case study of the red metro line in Washington D.C. The red line is a rapid transit line of the Washington Metro system, consisting of 27 stations in Montgomery County, Maryland, and Washington, D.C., in the United States. It is a primary line through downtown Washington and forms a long, narrow “U"-shaped, capped by its terminal stations at Shady Grove and Glenmont. Its configuration is presented in Fig.6.

Our data covers the period from 11 March 2018 to 29 March 2018 and includes the arrival time and departure time from each station along with the dwell time at stations, i.e. elapsed time between door opening and closing times. The dwell time exhibits a slight variation from the median as presented in Fig.7. Fig.7 presents the observed dwell times from the 11th until the 29th of March, 2018 using the Tukey boxplot convention [51]. The upper and lower boundaries of the boxes indicate the upper and lower quartiles (i.e. 75th and 25th percentiles denoted as Q3 and Q1, respectively). The black lines vertical to the boxes (whiskers) show the maximum and minimal values that are not outliers. The whiskers are determined by plotting the lowest datum still within 1.5 of the interquartile range (IQR) Q3-Q1 of the lower quartile, and the highest datum still within 1.5 IQR.

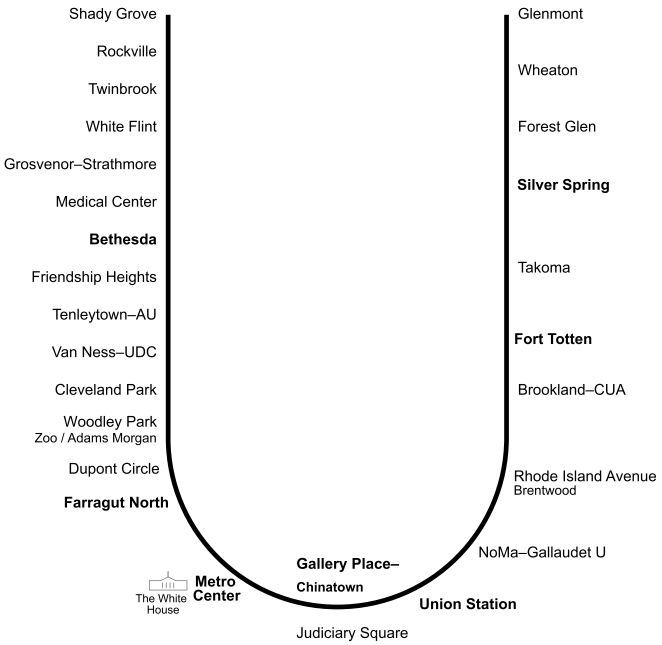

Fig. 6. "U"-shaped Red metro line in Washington Metro.

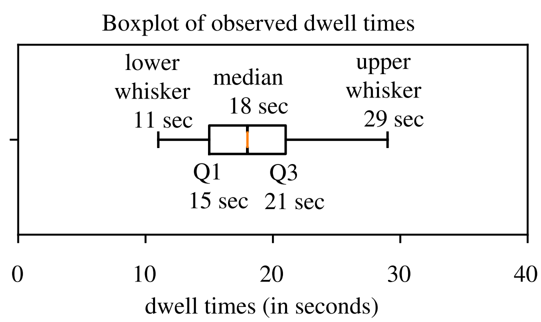

Fig. 7. Tukey boxplot of observed dwell times at all stations from 11/3/2018 until 29/3/2018.

TABLE IV

TARGET HEADWAYS AT DIFFERENT DAYS OF THE WEEK AND PEAK/OFF-PEAK HOURS

	Weekdays	Saturday	Sunday
AM Rush (5am-9:30am)	4 min	-	-
Midday (9:30am-3pm)	6 min	6 min	8 min
PM Rush (3pm-7pm)	4 min	6 min	8 min
Evening (7pm-9:30pm)	10 min	6 min	8 min
Late Night (9:30pm-close)	15 min	15 min	15 min

From Fig.7 one can note that the median dwell time at all stations is 18 s and the inter-quartile ranging between the 25th and the 75th percentile is only 6 sec. The coefficient of variation (CV) of the dwell time is 0.22. Therefore, the red metro line is deemed suitable for the application of our model because of its relatively stable dwell times, which do not vary significantly for all stations and trips included in the dataset.

The target headway of the red metro line at each station varies for peak and off-peak periods and for weekdays and weekends. Table IV summarizes the expected headway which the operator strives to deliver in a regular fashion.

In our experiments, we focus on the PM rush hours of one weekday where the target headway at the metro stations is set to 4 minutes. To investigate the improvement potential of our method, we compare the current regularity of services, extracted from the train movement observations, and the regularity when using our dispatching time modifications. Our day of interest is the 11 March 2018, and our time period of interest is 3pm-7pm.

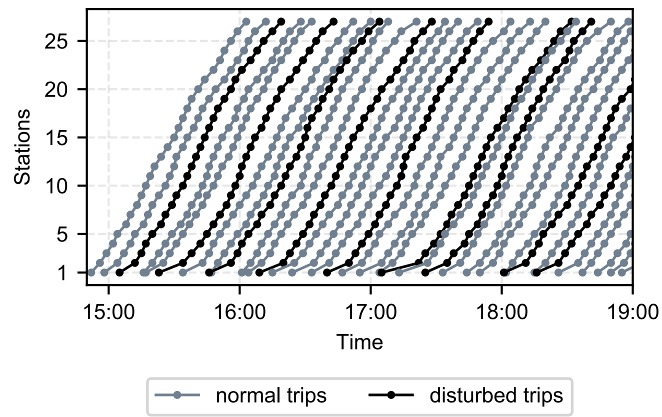

Fig. 8. Actual trajectories of trains operating from 3pm until 7pm on the 11th of March, 2018.

In addition, we perform a sensitivity analysis with regard to the number of trips considered, $\mathbf{N}_m$, i.e. the effect of the number of trips that we are allowed to modify their dispatching times after a disturbance occurred. On one hand, a limited number of dispatching time modifications is likely to be sufficient to smooth the operations after a disturbance. On the other hand, if we are allowed to modify the dispatching time of a single trip only (i.e., the trip that follows the train that is subject to a disturbance), the potential impact on the service regularity might be very limited. In our experiments, we use realistic travel times and dwell times as extracted from the operational data (see actual trajectories in Fig.8). Our assumption is that if we modify the dispatching time of a trip, its travel times between stations remain unchanged because trains do not operate in mixed traffic.

In our experiments, all trips that operate from 3pm until 7pm depart according to their actual departure times. Those dispatching times may be modified by our model to improve the service regularity. We examine the performance of our method when varying the number of rescheduled trips after a disturbance from a single trip ($\mathbf{N}_m = \{1\}$) up to 12 trips ($\mathbf{N}_m = \{1, 2, \ldots, 12\}$). That is to say, we perform 12 different experiments to investigate the importance of considering fewer or more "upstream" trips when re-setting the dispatching times after a disturbance occurs. The results of this analysis are presented in Fig.9.

From Fig.9 one can note that if after every disturbance we modify the dispatching time of the following trip only ($\mathbf{N}_m = 1$), the improvement in service regularity compared to the actual operations (do-nothing case) is $\approx 10.5\%$. If we modify the dispatching times of more upstream trips, $\mathbf{N}_m$, our performance improvement in terms of regularity is monotonically increasing. If we re-set the dispatching times of 5 upstream trips, $\mathbf{N}_m = \{1, 2, 3, 4, 5\}$, the potential benefit rises to $\approx 30\%$. After that, further regularity improvements are marginal.

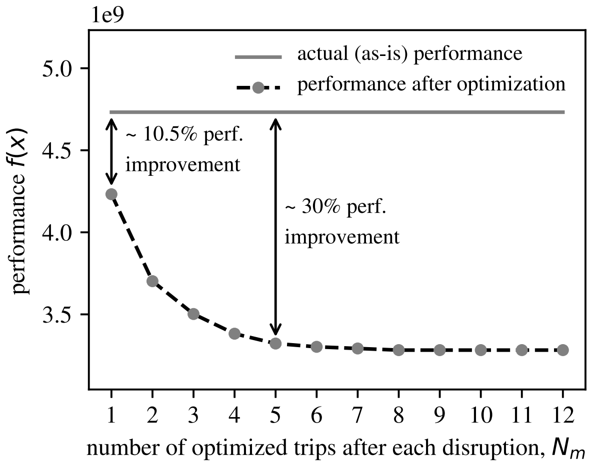

Fig. 9. Improvement of the service regularity expressed by $f(x)$ between 3pm and 7pm on the 11th of March when applying our model after each disturbance to upstream trips $\mathbf{N}_m$, ranging from 1 to 12.

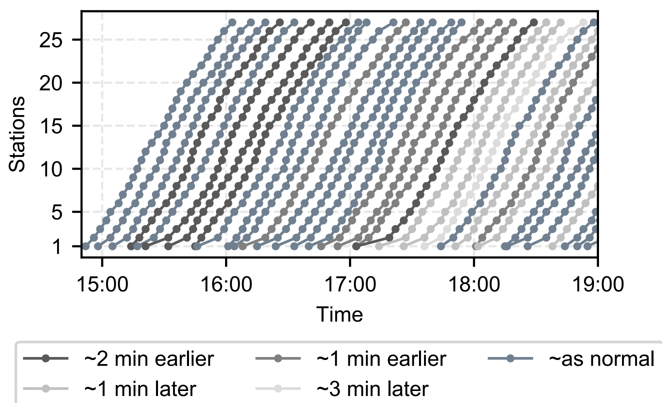

Fig. 10. Dispatching time changes after the reschedulings and resulting train trajectories for trains operating from 3pm until 7pm.

The updated trajectories after applying the reschedulings in the time period 3pm-7pm are presented in Fig.10. Fig.10 indicates also the performed dispatching time changes for every trip compared to the planned dispatching times indicated in Fig.8.

C. Comparison With State-of-the-Art Disturbance Management Methods

In the previous sub-section we observed that we can improve significantly the regularity of metro services when rescheduling the 5 trips proceeding a disturbed train. In this sub-section, we compare the performance of our model when rescheduling a set of $\mathbf{N}_m = \{1, 2, 3, 4, 5\}$ trips after each disturbance against the performance of the most closely related method described in [22]. Similar to the approach adopted in this study, Krasemann [22] also re-set the dispatching times of trips following each disturbance, without allowing for additional extra vehicles or re-routing. Unlike the method we propose in this study, Krasemann [22] did not provide a model for solving the rescheduling problem to global optimality. Instead, she introduced a heuristic solution method without formulating the rescheduling problem as a mathematical program.

TABLE V

IMPROVEMENT IN SERVICE REGULARITY EXPRESSED BY $f(x)$ BETWEEN 3PM AND 7PM WHEN APPLYING DIFFERENT DISRUPTION MANAGEMENT METHODS

	performance (sec^2)	improvement compared to the as-is performance
As-is (do-nothing)	4.73E+09	-
Proposed approach	3.32E+09	29.8%
Krasemann [22]	3.79E+09	19.9%

Similarly to our work, [22] defines a disturbance as the situation where the established timetable has become invalid because one train (or several) is deviating from its schedule (e.g. due to a signal malfunction that increases the travel time of a disturbed train at a particular section). Krasemann [22] extended the model formulation of [52] which divides a railway line into *line sections* and *station sections*. Line sections are discretized further into blocks, where each block can by occupied by no more than one vehicle at a time. Occupying a block is considered an event, and the trajectory of a train j is a sequence of consecutive events (e.g., an event list K_j).

In [22], consecutive trains are required to be always separated by a minimum headway of 3 minutes. Then, a greedy algorithm iteratively searches for the best train events to execute next and builds up a tree of consecutive active or terminated events (tree nodes). The phases of the greedy algorithm in [22] are:

- Phase 1: activate the events that already started when a disturbance occurred;
- Phase 2: perform a depth-first search to quickly find a feasible solution by building up a first complete branch of the tree;
- Phase 3: improve the existing solution by backtracking to potential nodes.

The greedy algorithm of [22] can treat cases where multiple tracks connect two stations of a line. However, in the case study of our metro line this is not needed because there is only one track available for traversing from one station to another station. To compare the disturbance management method of [22] with our proposed approach, we apply them to the operated trips in our case study from 3pm until 7pm on the 11th of March, 2018. Each time a disturbance occurs, the dispatching times of the following trips are re-computed by employing the two aforementioned approaches and the results are summarized in Table V.

From Table V one can note that the greedy algorithm of [22] improves the service regularity by ∼20% compared to the actual (as-is) performance where no rescheduling is applied. Our rescheduling approach improves further the service regularity to ∼30%. This additional improvement can be explained because our mathematical program converges to a globally optimal solution - unlike the greedy heuristic of [22]. This reaffirms the observation of [22] who noted that the greedy algorithm finds very fast a first feasible solution, but it is not always effective in branching and finding an improved solution.

VII. Concluding Remarks

In this study, we proposed a timetable recovery model that modifies the dispatching times of the subsequent trips after a disturbance (e.g. an unexpected increase in trip's travel time). In pursuit of a model that can be solved exactly yet be applied in real-time, we studied the rescheduling problem and introduced a quadratic model reformulation with penalty terms. This reformulated model was then proved to be solved to global optimality.

With this new model, we investigated how many trips, $\mathbf{N}_m$, should we consider for modifying their dispatching times after a disturbance occurs. This is instrumental in the understanding of the practical use of the model because it might not be prudent to reschedule the dispatching times of all remaining daily trips each time a disturbance occurs. This investigation was performed for a case study using actual data from the red metro line in Washington D.C. Results from our experiments based on train movement data indicate that one can consider the dispatching time modification of up to 5 upstream trips to smoothen the headway of a metro line after a disturbance. The optimal number of trips to be subject to dispatching adjustments is arguably dependent on service headway and the severity of the initial disturbance. Additionally, we showed that our model can provide better results compared to a state-of-the-art heuristic algorithm because it converges to a globally optimal solution.

The proposed method has the following main limitations which can be potentially addressed in future research:

- it is suitable for mitigating the effects of mild disturbances to service regularity by modifying the dispatching times of upstream trips. In the case of severe disruptions, metro operators should consider more substantial changes in the planned service provision including potentially trip cancellation, short-turning and inserting reserve vehicles.
- it is designed for optimizing regularity which is adequate in the context of services that operate in high-frequencies (more than 5 trips per hour). In the case of low frequencies, the objective function of our problem should be revised to reflect punctuality-oriented metrics;
- it is suitable in the context where the dwell times at stations are relatively stable and do not vary significantly as a function of passenger demand. This makes our approach particularly suitable for automated public transport systems where the time allotted for doors' opening/closing is fixed.

In future research, some of the assumptions made in this study can be relaxed. In particular, the formulation may be extended to account for flow-dependent dwell times. Such an extension will make the method applicable also for bus services. In addition, more control measures, such as stop-skipping, speed control and train holding at intermediate stations, can be considered in future research to improve further the recovery of services after a disturbance.

Acknowledgment

The authors would like to thank the Washington Metropolitan Area Transit Authority for the provision of the metro line data.

Appendix

Lemma A.1: For $\beta_1 < h_{min} + \bar{\delta}_0$, $\mathcal{F} = \emptyset$.

Proof: Let assume that for $\beta_1 < h_{min} + \bar{\delta}_0$, $\mathcal{F} \neq \emptyset$. Then, $\neg\ \exists\ x^0 \mid (x^0, a, h)$ satisfy Eq.(6)-(8). To satisfy constraint Eq.(8), $h_{min} \leq \delta_1 + x_1^0 - \bar{\delta}_0 \Rightarrow \delta_1 + x_1^0 \geq h_{min} + \bar{\delta}_0$. In addition, to satisfy constraint Eq.(6), $\delta_1 + x_1^0 \leq \beta_1$. Thus, β_1 should be greater than or equal to $h_{min} + \bar{\delta}_0$ and we reached a contradiction. This proves that $\mathcal{F}$ can be an empty set if the inequality constraints of Eq.(6)-(8) are binding (i.e., must be satisfied in all cases). □

References

[1] Ö. Yalçinnkaya and G. M. Bayhan, "Modelling and optimization of average travel time for a metro line by simulation and response surface methodology," *Eur. J. Oper. Res.*, vol. 196, no. 1, pp. 225–233, Jul. 2009.

[2] A. Schöbel, "Line planning in public transportation: Models and methods," *OR Spectr.*, vol. 34, no. 3, pp. 491–510, Jul. 2012.

[3] K. Gkiotsalitis and E. C. van Berkum, "An analytic solution for real-time bus holding subject to vehicle capacity limits," *Transp. Res. C, Emerg. Technol.*, vol. 121, Dec. 2020, Art. no. 102815.

[4] S. Li, M. M. Dessouky, L. Yang, and Z. Gao, "Joint optimal train regulation and passenger flow control strategy for high-frequency metro lines," *Transp. Res. B, Methodol.*, vol. 99, pp. 113–137, May 2017.

[5] K. Gkiotsalitis and E. C. van Berkum, "An exact method for the bus dispatching problem in rolling horizons," *Transp. Res. C, Emerg. Technol.*, vol. 110, pp. 143–165, Jan. 2020.

[6] A. Carrel, R. G. Mishalani, N. H. Wilson, J. P. Attanucci, and A. B. Rahbee, "Decision factors in service control on high-frequency metro line: Importance in service delivery," *Transp. Res. Rec.*, vol. 2146, no. 1, pp. 52–59, 2010.

[7] Y. Zhu and R. M. P. Goverde, "Railway timetable rescheduling with flexible stopping and flexible short-turning during disruptions," *Transp. Res. B, Methodol.*, vol. 123, pp. 149–181, May 2019.

[8] A. Chakraborty, N. Kaza, G.-J. Knaap, and B. Deal, "Robust plans and contingent plans: Scenario planning for an uncertain world," *J. Amer. Planning Assoc.*, vol. 77, no. 3, pp. 251–266, Jul. 2011.

[9] L. Kroon and D. Huisman, "Algorithmic support for railway disruption management," in *Transitions Towards Sustainable Mobility*. Berlin, Germany: Springer, 2011, pp. 193–210.

[10] L. P. Veelenturf, L. G. Kroon, and G. Maróti, "Passenger oriented railway disruption management by adapting timetables and rolling stock schedules," *Transp. Res. C, Emerg. Technol.*, vol. 80, pp. 133–147, Jul. 2017.

[11] D. Golightly and N. Dadashi, "The characteristics of railway service disruption: Implications for disruption management," *Ergonomics*, vol. 60, no. 3, pp. 307–320, Mar. 2017.

[12] J. Jespersen-Groth *et al.*, "Disruption management in passenger railway transportation," in *Robust and Online Large-Scale Optimization*. Berlin, Germany: Springer, 2009, pp. 399–421.

[13] V. Cacchiani *et al.*, "An overview of recovery models and algorithms for real-time railway rescheduling," *Transp. Res. B, Methodol.*, vol. 63, pp. 15–37, May 2014.

[14] N. Ghaemi, O. Cats, and R. M. P. Goverde, "Railway disruption management challenges and possible solution directions," *Public Transp.*, vol. 9, nos. 1–2, pp. 343–364, Jul. 2017.

[15] A. A. Zilko, D. Kurowicka, and R. M. P. Goverde, "Modeling railway disruption lengths with copula Bayesian networks," *Transp. Res. C, Emerg. Technol.*, vol. 68, pp. 350–368, Jul. 2016.

[16] N. Ghaemi, A. A. Zilko, F. Yan, O. Cats, D. Kurowicka, and R. M. P. Goverde, "Impact of railway disruption predictions and rescheduling on passenger delays," *J. Rail Transp. Planning Manage.*, vol. 8, no. 2, pp. 103–122, Sep. 2018.

[17] F. Chu and A. Oetting, "Modeling capacity consumption considering disruption program characteristics and the transition phase to steady operations during disruptions," *J. Rail Transp. Planning Manage.*, vol. 3, no. 3, pp. 54–67, Aug. 2013.

[18] A. D'Ariano, F. Corman, D. Pacciarelli, and M. Pranzo, "Reordering and local rerouting strategies to manage train traffic in real time," *Transp. Sci.*, vol. 42, no. 4, pp. 405–419, Nov. 2008.

[19] F. Corman, A. D'Ariano, D. Pacciarelli, and M. Pranzo, "A tabu search algorithm for rerouting trains during rail operations," *Transp. Res. B, Methodol.*, vol. 44, no. 1, pp. 175–192, Jan. 2010.

[20] L. Meng and X. Zhou, "Simultaneous train rerouting and rescheduling on an N-track network: A model reformulation with network-based cumulative flow variables," *Transp. Res. B, Methodol.*, vol. 67, pp. 208–234, Sep. 2014.

[21] P. Pellegrini, G. Marlière, and J. Rodriguez, "Optimal train routing and scheduling for managing traffic perturbations in complex junctions," *Transp. Res. B, Methodol.*, vol. 59, pp. 58–80, Jan. 2014.

[22] J. T. Krasemann, "Design of an effective algorithm for fast response to the re-scheduling of railway traffic during disturbances," *Transp. Res. C, Emerg. Technol.*, vol. 20, no. 1, pp. 62–78, Feb. 2012.

[23] K. Gkiotsalitis and O. Cats, "Reliable frequency determination: Incorporating information on service uncertainty when setting dispatching headways," *Transp. Res. C, Emerg. Technol.*, vol. 88, pp. 187–207, Mar. 2018.

[24] M. M. Nesheli, A. Ceder, and V. A. Gonzalez, "Real-time public-transport operational tactics using synchronized transfers to eliminate vehicle bunching," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 11, pp. 3220–3229, Nov. 2016.

[25] A. Sun and M. Hickman, "The real-time stop–skipping problem," *J. Intell. Transp. Syst.*, vol. 9, no. 2, pp. 91–109, Apr. 2005.

[26] Z. Liu, Y. Yan, X. Qu, and Y. Zhang, "Bus stop-skipping scheme with random travel time," *Transp. Res. C, Emerg. Technol.*, vol. 35, pp. 46–56, Oct. 2013.

[27] X. Chen, B. Hellinga, C. Chang, and L. Fu, "Optimization of headways with stop-skipping control: A case study of bus rapid transit system," *J. Adv. Transp.*, vol. 49, no. 3, pp. 385–401, Apr. 2015.

[28] K. Gkiotsalitis, "Robust stop-skipping at the tactical planning stage with evolutionary optimization," *Transp. Res. Rec.*, vol. 2673, no. 3, pp. 611–623, 2019.

[29] G. F. Newell, "Control of pairing of vehicles on a public transportation route, two vehicles, one control point," *Transp. Sci.*, vol. 8, no. 3, pp. 248–264, Aug. 1974.

[30] D. Hernández, J. C. Muñoz, R. Giesen, and F. Delgado, "Analysis of real-time control strategies in a corridor with multiple bus services," *Transp. Res. B, Methodol.*, vol. 78, pp. 83–105, Aug. 2015.

[31] W. Wu, R. Liu, and W. Jin, "Modelling bus bunching and holding control with vehicle overtaking and distributed passenger boarding behaviour," *Transp. Res. B, Methodol.*, vol. 104, pp. 175–197, Oct. 2017.

[32] A. Gavriilidou and O. Cats, "Reconciling transfer synchronization and service regularity: Real-time control strategies using passenger data," *Transportmetrica A, Transp. Sci.*, vol. 15, no. 2, pp. 215–243, 2018.

[33] K. Gkiotsalitis and O. Cats, "Multi-constrained bus holding control in time windows with branch and bound and alternating minimization," *Transportmetrica B, Transp. Dyn.*, vol. 7, no. 1, pp. 1258–1285, Dec. 2019.

[34] A. Adamski and A. Turnau, "Simulation support tool for real-time dispatching control in public transport," *Transp. Res. A, Policy Pract.*, vol. 32, no. 2, pp. 73–87, Feb. 1998.

[35] J. G. Strathman *et al.*, "Automated bus dispatching, operations control, and service reliability: Baseline analysis," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 1666, no. 1, pp. 28–36, Jan. 1999.

[36] Z. Cao, A. A. Ceder, and S. Zhang, "Real-time schedule adjustments for autonomous public transport vehicles," *Transp. Res. C, Emerg. Technol.*, vol. 109, pp. 60–78, Dec. 2019.

[37] K. Gkiotsalitis and F. Alesiani, "Robust timetable optimization for bus lines subject to resource and regulatory constraints," *Transp. Res. E, Logistics Transp. Rev.*, vol. 128, pp. 30–51, Aug. 2019.

[38] K. Yang, S. Guan, and C. Wang, "The passenger door system and its control principle on Shenzhen metro line 2," in *Proc. 2nd Int. Conf. Mech. Automat. Control Eng.*, Jul. 2011, pp. 855–857.

[39] P. Marguier, "Bus route performance evaluation under stochastic conditions," Ph.D. dissertation, Dept. Civil Environ. Eng., MIT, Cambridge, MA, USA, 1985.

[40] T.-M. Lin and N. H. Wilson, "Dwell time relationships for light rail systems," *Transp. Res. Rec.*, no. 1361, pp. 287–295, 1992.

[41] X. J. Eberlein, N. H. M. Wilson, and D. Bernstein, "The holding problem with real-time information available," *Transp. Sci.*, vol. 35, no. 1, pp. 1–18, Feb. 2001.

[42] C. F. Daganzo, "A headway-based approach to eliminate bus bunching: Systematic analysis and comparisons," *Transp. Res. B, Methodol.*, vol. 43, no. 10, pp. 913–921, Dec. 2009.

[43] C. G. Walker, J. N. Snowdon, and D. M. Ryan, "Simultaneous disruption recovery of a train timetable and crew roster in real time," *Comput. Oper. Res.*, vol. 32, no. 8, pp. 2077–2094, Aug. 2005.

[44] A. Ceder, *Public Transit Planning and Operation: Modeling, Practice and Behavior*. Boca Raton, FL, USA: CRC Press, 2016.

[45] K. Gkiotsalitis, Z. Wu, and O. Cats, "A cost-minimization model for bus fleet allocation featuring the tactical generation of short-turning and interlining options," *Transp. Res. C, Emerg. Technol.*, vol. 98, pp. 14–36, Jan. 2019.

[46] M. Trompet, X. Liu, and D. J. Graham, "Development of key performance indicator to compare regularity of service between urban bus operators," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2216, no. 1, pp. 33–41, Jan. 2011.

[47] E. E. Osuna and G. F. Newell, "Control strategies for an idealized public transportation system," *Transp. Sci.*, vol. 6, no. 1, pp. 52–72, Feb. 1972.

[48] C. M. Li and F. Manya, "Maxsat, hard and soft constraints," in *Handbook of Satisfiability*, vol. 185. Amsterdam, The Netherlands: IOS Press, 2009, pp. 613–631.

[49] K. Gkiotsalitis. (2019). *Timetable Recovery Model*. [Online]. Available: https://github.com/KGkiotsalitis/Timetable_recovery

[50] R. Storn and K. Price, "Differential evolution—A simple and efficient heuristic for global optimization over continuous spaces," *J. Global Optim.*, vol. 11, no. 4, pp. 341–359, 1997.

[51] R. Mcgill, J. W. Tukey, and W. A. Larsen, "Variations of box plots," *Amer. Statistician*, vol. 32, no. 1, pp. 12–16, Feb. 1978.

[52] J. Törnquist and J. A. Persson, "N-tracked railway traffic re-scheduling during disturbances," *Transp. Res. B, Methodol.*, vol. 41, no. 3, pp. 342–362, Mar. 2007.

Konstantinos Gkiotsalitis (Member, IEEE) received the M.Sc. degree in transport and sustainable development from Imperial College London and University College London, and the Ph.D. degree from the National Technical University of Athens. From 2012 to 2018, he was conducting transportation research and development with the NEC Laboratories Europe (Heidelberg, Germany) and held the positions of a Research Associate and a Research Scientist. He is currently an Assistant Professor of data science in transportation engineering with the Center for Transport Studies, University of Twente. His research interests include public transport modeling, tactical/operational planning, traffic operations, and data-driven optimization.

Oded Cats (Member, IEEE) received the dual Ph.D. degree from the KTH Royal Institute of Technology, Stockholm, and the Technion–Israel Institute of Technology. He is currently an Associate Professor of passenger transport systems with the Delft University of Technology. His research develops methods and models of multi-modal metropolitan passenger transport systems by combining advancements from behavioural sciences, operations research, and complex network theory. His research interests include the development of dynamic transit assignment models, the optimisation of passenger service operations, network robustness analysis, and real-time control methods. He co-directs the Smart Public Transport Laboratory, TU Delft, leading a research group that works closely with public transport authorities and operators.