

A railway timetable rescheduling approach for handling large scale disruptions

Lucas P. Veelenturf

Rotterdam School of Management, Erasmus University,
lveelenturf@rsm.nl

Martin P. Kidd, Valentina Cacchiani

Department of Electrical, Electronic and Information Engineering, University of Bologna,
martin.kidd@unibo.it, valentina.cacchiani@unibo.it

Leo G. Kroon

Rotterdam School of Management, Erasmus University,
The Netherlands and Process quality & Innovation, Netherlands Railways, The Netherlands,
lkroon@rsm.nl

Paolo Toth

Department of Electrical, Electronic and Information Engineering, University of Bologna,
paolo.toth@unibo.it

July 25, 2014

Abstract

On a daily basis, relatively large disruptions require infrastructure managers and railway operators to reschedule their railway timetables together with their rolling stock and crew schedules. This research focuses on timetable rescheduling for passenger trains at a macroscopic level in a railway network. An integer programming model is formulated for solving the timetable rescheduling problem, which minimizes the number of cancelled and delayed trains while adhering to infrastructure and rolling stock capacity constraints. The possibility of rerouting trains in order to reduce the number of cancelled and delayed trains is also considered. In addition, all stages of the disruption management process (from the start of the disruption to the time the normal situation is restored) are taken into account. Computational tests of the described model on a heavily used part of the Dutch railway network show that we are able to find optimal solutions in short computation times. This makes the approach applicable for use in practice.¹

¹The research leading to this paper has received funding from the European Union's Seventh Framework Programme (FP7/2007-2013) in the ON-TIME project under Grant Agreement SCP1-GA-2011-285243.

Furthermore, we want to thank *Vereniging Trustfonds Erasmus Universiteit Rotterdam* for their support for a research visit of Lucas Veelenturf to the University of Bologna.

1 Introduction

The occurrence of unexpected large disruptions, such as the unavailability of railway track segments due to broken overhead wires or rolling stock breakdowns, causes train delays and train cancellations with a consequent reduction of the quality of service to the passengers. Therefore, it is crucial to recover from such situations as quickly as possible in order to reduce passenger dissatisfaction and to restore the service of the railway system.

Due to its complexity, the recovery problem is usually decomposed into phases that are solved in sequence. The main phases consist of timetable rescheduling, rolling stock rescheduling and crew rescheduling. Timetable rescheduling calls for determining a feasible timetable by applying reordering, retiming and rerouting of trains and even train cancellations. The derived timetable is input to the second phase, in which it may be necessary to determine a new rolling stock allocation, due to the changes applied in the previous phase. Similarly, the new timetable and rolling stock allocation are input to the last phase that aims at obtaining a feasible crew schedule. Obviously, a feedback loop is sometimes necessary if no feasible rolling stock or crew plan can be obtained, possibly requiring the cancellation of additional trains. Solving the three phases separately may lead to sub-optimal solutions. However, solving them all in an integrated way would lead to unacceptably long computing times, as we are facing real-time problems. An overview of current models for solving these three steps is presented in Cacchiani et al. [2014].

In this paper, we focus on the timetable rescheduling phase, thereby taking into account constraints from the rolling stock rescheduling phase in order to increase the probability of obtaining a feasible rolling stock schedule during the second phase. Constraints of the crew rescheduling phase are more difficult to include, since there are much more complicated rules about (meal) breaks and durations of crew duties. Therefore these are not considered in our model.

We study timetable rescheduling at a macroscopic level, i.e. with high level constraints disregarding detailed information on signals and routes inside stations or junctions. The reason is that we want to deal with a complex real-world railway network and, at the same time, to solve the problem in very short computing times. First, at a macroscopic level it has to be determined which trains can still run with the available infrastructure capacity. Thereafter, small conflicts at the signaling level should be detected and solved by slightly delaying some trains.

We consider large disruptions related to blockages of one or more railway tracks between stations for a certain period of time (e.g. two hours). Indeed, disruptions of this kind are very hard to manage by railway operators and infrastructure managers, as they cause many changes in the system and decisions need to be taken very quickly. In addition, they are not uncommon events and the support of an automated tool for solving them is highly desirable. Currently such disruptions are handled in practice by selecting an appropriate contingency plan from a large set of such plans.

The main contribution of this paper consists of proposing an Integer Linear Programming (ILP) formulation for the timetable rescheduling problem to deal with large disruptions on a real-world railway network. The formulation takes into account constraints that allow to partially integrate it with the rolling stock rescheduling phase. In particular, we consider a railway network with a cyclic timetable, i.e. the schedule of the trains is repeated every given time period (for example every hour).

This approach generates a new timetable from the start of a disruption till the moment at which the timetable must be back again to the normal state. This means that the approach does not only make a new reduced cyclic timetable for the steady disrupted state, but it also produces a timetable to make the transition from the original timetable to this new reduced cyclic timetable and back to the original timetable after the disruption has ended.

A useful feature of the proposed model is that it takes into account the possibility of rerouting trains along alternative geographical paths in the network in order to reduce the number of trains that are cancelled or delayed.

The model is solved to optimality by a general purpose solver on a set of real-world instances of Netherlands Railways (the major railway operator in the Netherlands) in short computing times.

This paper is organized as follows. Section 2 presents an overview of related research. In Section 3, the problem is presented and in Section 4 an ILP formulation is given. Section 5

is devoted to the computational results based on instances of Netherlands Railways, and the conclusions are discussed in Section 6.

2 Literature overview

Many works study the Train Timetabling Problem during the planning phase, i.e., when an optimal timetable is derived for a set of trains in a time horizon of six months to one year. We refer the reader to the following recent surveys on this topic: Cacchiani and Toth [2012], Caprara et al. [2007], Caprara et al. [2011] and Lusby et al. [2011]. In recent years, many studies have been developed dealing with real-time timetable rescheduling. The majority of them concern train rescheduling when relatively small disturbances affect a subset of trains, instead of large disruptions as is the case in our paper. We refer to Cacchiani et al. [2014] for an overview of real-time rescheduling problems and solution approaches.

For example, the real-time traffic management system ROMA (Railway traffic Optimization by Means of Alternative graphs) is presented in D’Ariano et al. [2008a] and D’Ariano et al. [2007]. ROMA considers the infrastructure at a detailed level and uses a branch-and-bound algorithm for sequencing train movements combined with a local search algorithm for rerouting trains. The experiments described in these papers concern the line between Utrecht and ’s Hertogenbosch, and the congested areas around Utrecht Central Station and around Schiphol Amsterdam Airport in the Netherlands.

Extensions of ROMA are presented in Corman et al. [2009], Corman et al. [2010], Corman et al. [2012], and D’Ariano et al. [2008b], taking into account different objectives (minimization of train delays and preservation of train connections), or building flexible timetables that postpone certain decisions to the operational phase. These works consider a set of instances provided by Netherlands Railways.

Other microscopic approaches for small disturbances are presented in Boccia et al. [2013], Caimi et al. [2012], Lusby et al. [2013] and Mannino and Mascis [2009].

A macroscopic level of detail of the railway network to handle disturbances is considered in Acuna-Agost et al. [2011b], Acuna-Agost et al. [2011a], Dollevoet et al. [2012], Kecman et al. [2013], Schöbel [2009] and Törnquist and Persson [2007].

Only very few works deal with large disruptions. In Albrecht et al. [2013], disruptions due to unexpected track maintenance extensions are considered, i.e. longer maintenance operations are required than were planned. A meta-heuristic is used to construct an integrated timetable which includes track maintenance, and an operational tool is used to generate a new feasible schedule for the disrupted system. A case study for a single track rail network in Queensland Australia is carried out.

In Brucker et al. [2002], train rescheduling is considered in the case of a partial track blockage due to construction works. A local search algorithm is presented with the goal of minimizing lateness. This algorithm is tested on real-life instances of the German railways.

In Corman et al. [2011], the authors consider the case of double track railway lines where some block sections of one track are unavailable. Centralized and distributed approaches are presented: in the centralized approach the entire rescheduling problem is solved, while in the distributed approach a coordinator sets constraints between areas and delegates the scheduling decisions to local schedulers. Computational experiments on a large railway network in the Netherlands show that both approaches face increasing difficulty in finding feasible schedules in a short computation time for increasing time horizons of traffic prediction.

In Wiklund [2007], the author describes a simulation procedure for simulating train traffic at a microscopic level in order to determine the effectiveness of various recovery strategies in case of large disruptions. A case study involving a fire at the interlocking system of a station in Stockholm is considered.

Our work extends the ILP model presented in Louwerse and Huisman [2014], in which two double track lines were considered and a partial or full track blockage was taken into account. The new feature consists of dealing with a real-world railway network where the number of tracks within the stations and between the stations is not limited to two. Furthermore, trains are allowed to take other tracks within the stations or between the stations in comparison with the tracks they are originally scheduled to, and tracks may now be used in both directions. To prevent overtakings of trains running on the same track, additional constraints are considered.

In addition, train reroutings along different paths are allowed in order to avoid the disrupted area. An advantage of rerouting trains is that passengers do not need to reroute themselves (possibly with some transfers) and that they can experience smaller delays.

As in Louwerse and Huisman [2014], we include rolling stock constraints in order to increase the probability of getting a feasible rolling stock schedule.

We also focus on all stages of the disruption management process, i.e. from the start of the disruption to the time at which the normal situation is restored. In Louwerse and Huisman [2014] only a new cyclic timetable is made for execution during the disruption. The authors do not consider the transition from the original timetable to the new temporary timetable, nor the transition back to the original timetable when the disruption has ended. In addition, they assume that, at the time the disruption starts, the network is empty. In our research we do not have these assumptions, and so we also determine how the timetable must be modified during the transition phases.

3 Problem description

In this paper we consider a real-time timetable rescheduling approach for railway networks. In case of a major disruption (i.e. temporarily blocked tracks) this approach is able to determine, by taking into account the available infrastructure capacity, which train services (or parts thereof) should be cancelled and which should be delayed such that as many trains as possible can still be operated. A train service is partially cancelled if it runs only to a subset of the stations it normally runs to. In most situations this means that the train ends at a location different than the planned end location or starts at a location different than the planned start location.

The use of the available infrastructure capacity in this approach is considered from a macroscopic point of view, and by taking into account rolling stock capacity there is a high probability that the new timetable has a feasible rolling stock schedule as well.

3.1 The disruption

For this approach we consider disruptions where a number of tracks between stations are blocked. There can be multiple track blockages at the same time at different locations. However, for the computational results we only consider disruptions where tracks at the same geographical location are blocked.

The duration of the disruption is assumed to be known and there is a fixed limited time available after the disruption before all trains must run again according to their original schedule.

Trains which passed their last stop before the blocked segment at the moment the disruption occurs need special attention. It is not clear whether these trains did or did not pass the critical point which caused the disruption. Therefore we assume in this research that these trains just continue as planned.

3.2 Resource restrictions and assumptions

We consider a railway network that consists of a set of stations (each one with a given capacity) and of a set of open track sections (the parts of the railway network between two consecutive stations), that can be single tracked, double tracked or with even more parallel tracks. A given set of trains runs on the railway network, each one according to its original timetable. Trains are characterized by a type, e.g. they can be regional or intercity trains. Each train uses a rolling stock composition, i.e. a set of coupled rolling stock units, of the same type as the train.

We distinguish three types of resources which a train may occupy at any given moment, namely tracks in open track sections, tracks in stations, and rolling stock compositions.

The capacity of a station is characterised by the number of tracks it has: A track within a station may only be occupied by one train at any given time. Furthermore, after a train has used a track in a station, a certain headway time needs to pass before another train can use the same track. All tracks in a station are assumed to have a platform next to them and some stations have a shunting yard with an infinite capacity of tracks. A simple extension of

the model could handle shunting yards with finite capacity. However the shunting yards are not considered to be the bottlenecks since we want to run (and not store) as many trains as possible.

The capacity of an open track section between two stations is also characterised by the number of tracks it has. The tracks can be used in both directions. It is assumed that a train cannot switch tracks while running on an open track section, and a track can only be used by multiple trains at the same time if the trains run in the same direction. A certain headway time should be taken into account between two trains using a track at the same time in the same direction, or between two trains using a track consecutively in opposite directions. It is assumed that a train entering a station from an open track section is able to reach every track in the station regardless of the track of the open track section it is entering from.

At each station with a shunting yard, a limited number of rolling stock compositions is available at the start of the day. A train uses a rolling stock composition for its entire duration, after which the rolling stock composition is moved to a shunting yard or used by another train. Hence rolling stock compositions may only end their duties at stations with shunting yards, and the compositions are not split during the day. Furthermore, two trains may share rolling stock only if they are of the same type. After a train has ended, a minimum turnaround time is required before the rolling stock composition of that train may be used by another train.

Finally, the minimum running time between two stations and the minimum dwell time inside a station should be respected by all trains. The arrival or departure of a train at a station may further be delayed by only a maximum amount of time, and trains may only end at their final destination or at their last stop before the disrupted tracks.

4 Mathematical formulation

Each train service is represented by a set of events, which are arrivals or departures at certain stations. The aim of the rescheduling approach is to determine the times at which these events take place or to decide to cancel some events.

To do this, the timetable rescheduling approach is based on an *event-activity network* represented by a directed graph $N = (E, A)$, where E is the set of vertices (*events*) and A the set of arcs (*activities*). The graph N is associated with a set of trains T and an original timetable for these trains. The set $E = E_{\text{train}} \cup E_{\text{inv}}$ of events consists of a set E_{train} of *train events*, and a set E_{inv} of *inventory events*.

Each train event $e \in E_{\text{train}}$ represents either a departure or an arrival at a certain station. The train of which train event e is a departure or arrival of, is denoted by t_e . The scheduled time at which train event e takes place in the original timetable is given by q_e . For each train event e , d_e denotes the maximum allowed delay for the event.

Each event $e \in E_{\text{train}}$ is associated with a set of resources (tracks in open track sections, tracks in stations, and rolling stock compositions) which it uses at the moment the event takes place.

An inventory event $e \in E_{\text{inv}}$ has different characteristics. It represents the resource inventory of a certain station, open track section or shunting yard at the start of the day. The number of resources made available by inventory event e is denoted by i_e .

The use of an activity $a = (e, f) \in A$ directed from event $e \in E$ to event $f \in E$ denotes the fact that event f uses one of the resources occupied by e after e has taken place. Between two events there can be multiple activities, and for every resource type there can only be one activity between the same pair of events. Each activity $a = (e, f) \in A$ has an associated minimum duration L_a which is necessary for the specific resource used by e to become available for use by f . In summary, the activities determine the possible orders in which the resource units are used for the events.

The Timetable Rescheduling Problem consists of delaying some trains and cancelling some other trains such that maximum delay and capacity constraints are satisfied while minimising the deviation from the original timetable. The first part of the proposed ILP formulation is given in (1)-(5). The further constraints are described in subsequent sections. In the model, x_e is a decision variable denoting the time at which event $e \in E_{\text{train}}$ takes place in the new timetable and y_t is a binary decision variable such that $y_t = 1$ if train $t \in T$ is cancelled, and $y_t = 0$ otherwise.

The objective function (1) is a weighted sum of the number of cancelled trains and the sum of the delays of all the events from their original scheduled times. For every train t , λ_t describes the penalty for cancelling the train (which, for example, can depend on the train type or the running time), and for each event e , μ_e is the penalty per time unit for delaying the event. Constraint sets (2) and (3) ensure that an event takes place at least at its scheduled time in the original timetable and that its maximum allowed delay is not exceeded. Constraints (3) also ensures that if a train is cancelled, then it “virtually” runs at its original time (i.e. no delay penalty is considered).

$$\text{Minimise} \quad \sum_{t \in T} \lambda_t y_t + \sum_{e \in E_{\text{train}}} \mu_e (x_e - q_e) \quad (1)$$

$$\text{subject to} \quad x_e - q_e \geq 0 \quad \forall e \in E_{\text{train}} \quad (2)$$

$$x_e - q_e \leq (1 - y_{t_e}) d_e \quad \forall e \in E_{\text{train}} \quad (3)$$

$$y_t \in \{0, 1\} \quad \forall t \in T \quad (4)$$

$$x_e \in \mathbb{N} \quad \forall e \in E_{\text{train}} \quad (5)$$

4.1 Capacity constraints

Capacity constraints are needed in order to ensure that a resource unit is not used by more than one train at a time, and that trains which are not able to get the resources they require are cancelled.

The capacity is handled by the activities $a \in A$. For each event e we divide the activities into two groups: in-activities and out-activities. The *in-activities* of event e are all activities into event e and are denoted by the set: $A^-(e) = \{a = (f, e) \in A \mid f \in E\}$. An in-activity $a = (f, e)$ of event e means that event e can use the same resource as event f at least L_a minutes after event f has taken place.

For each type of resource (tracks in open track sections, tracks in stations and rolling stock compositions) and for each event e we can define a subset $C \subset A^-(e)$ of the in-activities associated with that resource type only. Then, the set $\mathcal{A}^-(e)$ denotes the collection of these in-activity subsets for event $e \in E$. This means that for each event e , $|\mathcal{A}^-(e)|$ is smaller than or equal to three (the number of resource types). So for example, the set $\mathcal{A}^-(e)$ can contain three of these in-activity subsets: *i*) one subset of activities into event e associated with open track section capacity, *ii*) one subset of activities into event e associated with station capacity, and *iii*) one subset of activities into event e associated with rolling stock availability. The details of these types of activities are discussed in Sections (4.1.2)-(4.1.4).

The second group of activities of an event e are the *out-activities* which are all activities out of event e and are denoted by the set $A^+(e) = \{a = (e, f) \in A \mid f \in E\}$. An out-activity $a = (e, f)$ of event e means that event f can use the same resource as event e at least L_a minutes after event e has taken place. We have for each type of resource a subset $C \subset A^+(e)$ of the out-activities. These subsets are grouped in the set $\mathcal{A}^+(e)$ denoting the collection of out-activity subsets for event e .

Not every event has all three subsets of in-activities or out-activities. This is because for some events it is fixed which event will be the predecessor or successor event using the same resource and there is no minimum time involved before the resource becomes available. Therefore, we left out decision variables for these activities for which we can easily determine the values.

For example, considering the station capacity, it is sure that after an arrival of a train at a certain track in a station, the next event on that track must be the departure of that train (unless the train has reached its last station). Furthermore, since these events are related to a single train, there is no headway time involved between them.

The capacity constraints are given by constraint sets (6)-(10), where z_a is a binary decision

variable such that $z_a = 1$ if activity $a \in A$ is selected, and 0 otherwise.

$$\sum_{a \in C} z_a + y_{t_e} = 1 \quad \forall e \in E_{\text{train}}, C \in \mathcal{A}^-(e) \quad (6)$$

$$\sum_{a \in C} z_a + y_{t_e} \leq 1 \quad \forall e \in E_{\text{train}}, C \in \mathcal{A}^+(e) \quad (7)$$

$$\sum_{a \in C} z_a \leq i_e \quad \forall e \in E_{\text{inv}}, C \in \mathcal{A}^+(e) \quad (8)$$

$$x_f - x_e + M(1 - z_a) \geq L_a \quad \forall a = (e, f) \in A \quad (9)$$

$$z_a \in \{0, 1\} \quad \forall a \in A \quad (10)$$

For a given event $e \in E_{\text{train}}$ and a subset of activities $C \in \mathcal{A}^-(e)$, constraint set (6) ensures that exactly one unit of the resource associated with subset C must be made available to event e or that the corresponding train must be cancelled. In other words, these constraints ensure that the train is cancelled if at least one resource is not available for an event, and that it is not cancelled if one unit of all resource types is available.

Similarly, for a given event $e \in E_{\text{train}}$ and subset of activities $C \in \mathcal{A}^+(e)$, constraint set (7) ensures that at most one unit of the associated resource is made available by event e to a successor event using the same resource unit. Finally constraint set (8) ensures that the available inventory for a specific type of resources is not exceeded, and constraint set (9) (where M is a large positive value) ensures that the minimum duration of an activity, if selected, is maintained.

4.1.1 Train service

The minimum running and dwell times of trains should be respected. This can be modelled by the capacity constraints of Section 4.1. However, in contrast with the rolling stock and track usage, for each train the order of the running and dwell events is fixed. Therefore we do not need constraint sets (6), (7) and (8).

Let A_{train} be the set of all train activities $a = (e, f)$ which represent running or dwelling of a train between consecutive events e and f of the same train. Then for train activities $a \in A_{\text{train}}$ constraints (9) can be modified into:

$$x_f - x_e \geq L_a \quad \forall a = (e, f) \in A_{\text{train}} \quad (11)$$

4.1.2 Open track section capacities

An open track section between two stations has a limited number of tracks which may be used in both directions. Multiple trains may use a track at the same time only if they run into the same direction, but a minimum headway between the two trains must be considered.

To model this, let $e_{\text{track}}^{k\ell}$ denote the initial open track section inventory event for open track section (k, ℓ) (the open track section between stations k and ℓ), let $E_{\text{arr}}^{k\ell} \subset E_{\text{train}}$ denote the set of events corresponding to an arrival at station ℓ of a train that departed from station k , and let $E_{\text{dep}}^{k\ell} \subset E_{\text{train}}$ denote the set of events corresponding to a departure from station k of a train in the direction of station ℓ .

In what follows the *open track section activities* for the open track section (k, ℓ) are described. This class of activities represents the sequential use of a track of the open track section or the very first use of a track of the open track section. We only describe the events at the side of station k . At the side of station ℓ the activities are constructed similarly.

- For an event $e \in E_{\text{dep}}^{k\ell}$ representing the departure of train t_e from station k onto the open track section (k, ℓ) :
 - The set $\mathcal{A}^-(e)$ contains a subset of open track section activities from each event $f \in E_{\text{dep}}^{k\ell} \setminus \{e\}$ to e , from each event $f \in E_{\text{arr}}^{\ell k}$ to e , and from the inventory event $e_{\text{track}}^{k\ell}$ to e . Thus for these activities, constraint set (6) implies that a train can only depart on a track of an open track section if a train has departed on the same track in the same direction, if a track used in the opposite direction became empty after an arrival, or if there is an empty track available from the inventory.

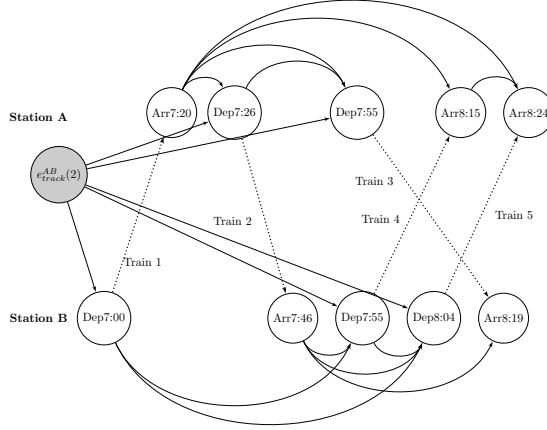


Figure 1: Example Event-Activity graph for open track section capacity

- The set $\mathcal{A}^+(e)$ contains a subset of open track section activities to each event $f \in E_{\text{dep}}^{k\ell} \setminus \{e\}$. Here constraint set (7) implies that, if a train departs onto an open track section, then at most one other train can depart directly after this train on the same track in the same direction.
- For an event $e \in E_{\text{arr}}^{\ell k}$ representing the arrival of train t_e at station k from the open track section (ℓ, k) :
 - There is no set $\mathcal{A}^-(e)$, since there is no capacity restriction for the arrival event as the train is already running on the track. We only have to take care that trains do not overtake each other which will be discussed in constraint set (12).
 - The set $\mathcal{A}^+(e)$ contains a subset of open track section activities to each event $f \in E_{\text{arr}}^{\ell k} \setminus \{e\}$, and to each event $f \in E_{\text{dep}}^{k\ell}$. Here constraint set (7) implies that, if a train arrives from a track of the open track section, then either another train arrives directly after it from the same track, or the track will be used in another direction by a train departing from station k .
- The set $\mathcal{A}^+(e_{\text{track}}^{k\ell})$ contains a subset of open track section activities from the inventory event $e_{\text{track}}^{k\ell}$ to each event $e \in E_{\text{dep}}^{k\ell} \cup E_{\text{arr}}^{\ell k}$. Here constraint set (8) implies that at most $i_{e_{\text{track}}^{k\ell}}$ trains (equal to the number of tracks between stations k and ℓ) may depart onto a track which has not been used before.

For these classes of activities, constraint set (9) models the headway time which has to be taken into account between two consecutive trains using the same track.

Figure 1 shows an example of a graph of open track section activities. Here we consider an open track between stations A and B . The events are represented by nodes and placed in a time-space plot where the vertical direction represents place and the horizontal direction represents time. For a better understanding of the activities, we assume in this example the time at which an event will take place as fixed. The capacity of an inventory event is given between brackets. The open track section activities are represented by arcs. To make the graph more intuitive, the train activities are also included and represented by dotted lines.

Additional constraints are needed to prevent overtaking of trains on the same track. Therefore, *track activity pairs* are introduced. For open track section (k, ℓ) and events $e, f \in E_{\text{dep}}^k$ and $e', f' \in E_{\text{arr}}^\ell$ such that $t_e = t_{e'}$ and $t_f = t_{f'}$, let $a = (e, f)$ be the activity presenting the consecutive departures of trains t_e and t_f from station k on a track in (k, ℓ) , while $a' = (e', f')$ is the activity corresponding to the consecutive arrivals of the two trains at station ℓ (as defined above). Then the pair (a, a') is a *track activity pair*, and B is defined as the set of all track activity pairs. In order to ensure that no overtaking on tracks takes place, constraint set (12) is required.

$$z_a = z_{a'} \quad \forall (a, a') \in B \quad (12)$$

Hence for a track activity pair (a, a') , if a is selected then so must be a' . This ensures that the order in which two trains arrive from a track is the same as the order in which they departed.

Note that forcing tracks to be used in only one direction can easily be achieved by not including activities from arrival events to departure events. Then the only in-activities for departure events come from other departure events. Also, we should construct an inventory event for each direction. The result will be that there are two disjoint graphs: one for each direction.

Figures 2-5 show with bold arcs all feasible solutions for the open track section capacity problem considered in Figure 1 with fixed event times. Note that the dashed arcs are train activities instead of open track section activities. These arcs are included such that one can easily see how events use the same resource unit. A selected path from the inventory event e_{track}^{kl} indicates in which order events take place on a single track in the open track section. For example in Figure 2 each track is used in only one direction. Trains 1, 4 and 5 use one track and Trains 2 and 3 use the other track.

Furthermore, Figures 2 and 5 show how constraints (12) work between Trains 2 and 3. Due to constraint set (12) both the arc between the departures of these trains and the arc between the arrivals of these trains are used. The same holds for Trains 4 and 5 in all Figures 2-5.

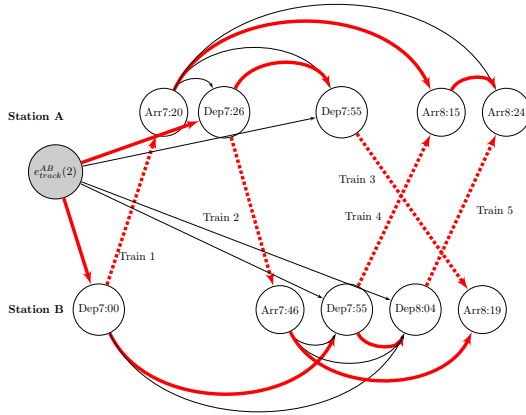


Figure 2: Solution 1

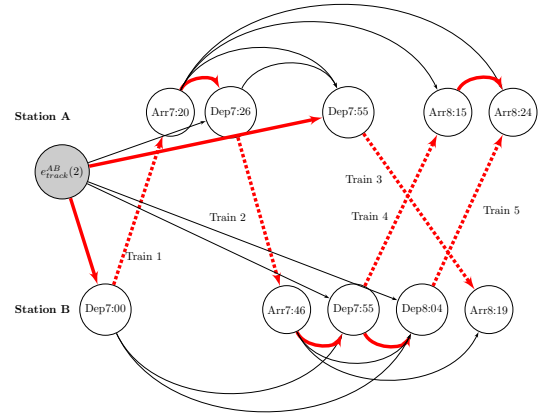


Figure 3: Solution 2

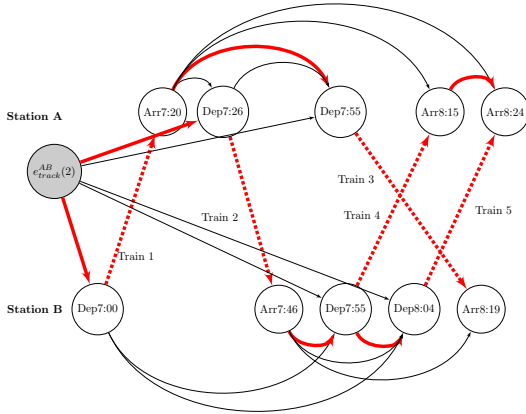


Figure 4: Solution 3

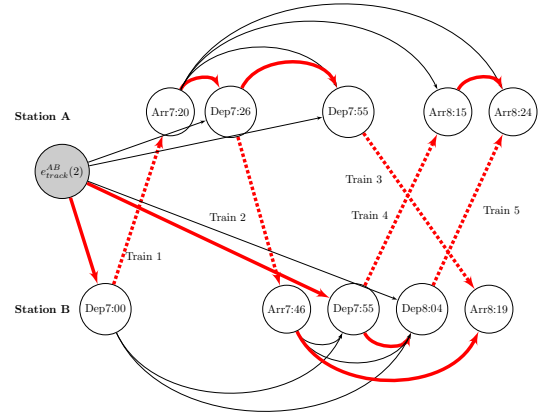


Figure 5: Solution 4

4.1.3 Station capacities

In a station a train needs to be assigned to a track with a platform to dwell or to pass if it does not have a scheduled stop. There cannot be two trains at the same track at the same time. After a train has left the track in the station, another train can arrive on that track.

To model this, we need the sets E_{arr}^k and E_{dep}^k which are all arrival and departure events at station k , respectively. Furthermore, let e_{stat}^k denote the track inventory event at station k .

In the following, the *station activities* are defined. These activities represent the sequential use of the same track in a station or the very first use of a track in a station.

- For an event $e \in E_{\text{arr}}^k$, representing the arrival of train t_e at station k :
 - The set $\mathcal{A}^-(e)$ contains a subset of station activities from each event $f \in E_{\text{dep}}^k$ to e , and from the inventory event e_{stat}^k to e . For this class of activities, constraint set (6) implies that a train can only arrive at station k if there is a track available. This means that a previous train has departed from a track in the station, or that a track has not been used before.
 - The set $\mathcal{A}^+(e)$ is not considered. It is sure that the next event using the same track is the departure of train t_e at station k since one train at the time is allowed on one track. We make use of this information and do not make it a decision variable in the model.
- For an event $e \in E_{\text{dep}}^k$, representing the departure of train t_e from station k :
 - The set $\mathcal{A}^-(e)$ is not considered for the same reason as above. It is sure that the previous event using the same track is the arrival of train t_e at station k .
 - The set $\mathcal{A}^+(e)$ contains a subset of station activities to each event $f \in E_{\text{arr}}^k$. For this class of activities, constraint set (7) implies that, if a train has departed from a station, then the track may be assigned to at most one new arrival.
- The set $\mathcal{A}^+(e_{\text{stat}}^k)$ contains a subset of station activities from the inventory event e_{stat}^k to each event $e \in E_{\text{arr}}^k$. Here constraint set (8) implies that at most $i_{e_{\text{stat}}^k}$ trains (equal to the number of tracks at that station) may arrive at a track of station k which has not been used before.
- Some events in $e \in E_{\text{dep}}^k$ represent a departure from station k which is the first departure of a train. For these events it is not clear if the rolling stock for the train arrives from the shunting yard or whether the train uses rolling stock which was waiting at a station track after it arrived servicing a train which ended in station k . For a departure event e which represents a first departure of a train, we construct the set $\mathcal{A}^-(e)$ in the same way as if e is an arrival event and the set $\mathcal{A}^+(e)$ in the same way as if event e is a departure event. This way it is modeled that train t_e needs to have a track available at the time it departs. The occupation of the tracks by rolling stock assigned to train t_e is handled correctly by the rolling stock constraints which are discussed in Section 4.1.4.
- Some events in $e \in E_{\text{arr}}^k$ represent an arrival at station k which is the last arrival of a train. For these events it is not clear whether the rolling stock goes the shunting yard or whether it stays at the station track to be used by another train. For an arrival event e which represents a first arrival of a train, we construct the set $\mathcal{A}^-(e)$ in the same way as if event e is an arrival and the set $\mathcal{A}^+(e)$ in the same way as if event e is a departure. This way it is modeled that train t_e needs an unoccupied track to arrive, and that the track is released for the arrival of another train directly (minimum L_a minutes) afterwards.

For these classes of activities, constraint set (9) models the headway time which has to be taken into account between a departure from a track in a station and an arrival on the same track.

Figure 6 shows an example of a graph of station activities. The events are represented by nodes and are considered to have a fixed time. In this figure, the horizontal direction represents time. Furthermore, the capacity of an inventory event is given between brackets. The station activities are represented by arcs and the train activities are also included and represented by dotted lines.

Figures 7-10 show with bold arcs all feasible solutions for the station capacity problem shown in Figure 6, given that the times the events take place are fixed. Note that the dashed arcs are train activities instead of station activities. A selected path from the inventory event e_{stat}^k indicates in which order events take place on a single track within the station. For example in Figure 7, Trains 1, 3 and 4 use one track and Train 2 uses the other track.

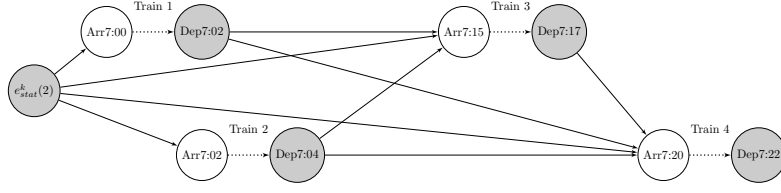


Figure 6: Example Event-Activity graph for station capacity

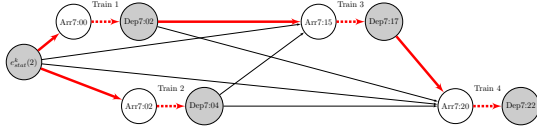


Figure 7: Solution 1

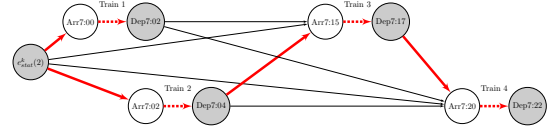


Figure 8: Solution 2

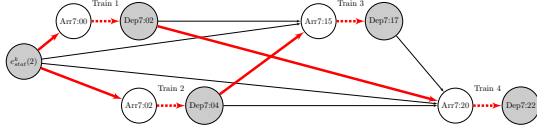


Figure 9: Solution 3

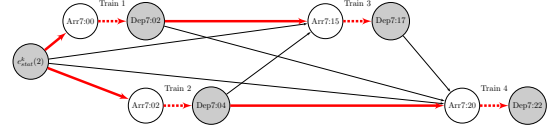


Figure 10: Solution 4

4.1.4 Rolling stock capacities

Every train needs rolling stock units to run. In our model we can have different types of rolling stock: rolling stock for intercity trains and rolling stock for regional trains. We assume that starting trains can use rolling stock of ending trains which are of the same type, or they can use rolling stock from the shunting yard, if available.

To model the assumptions in the structure of Section 4.1, let $E_{\text{start}}^k \subset E_{\text{train}}$ denote the set of departure events corresponding to the start of a train from station k , let $E_{\text{end}}^k \subset E_{\text{train}}$ denote the set of arrival events corresponding to the end of a train at station k , and let e_{rol}^k denote the initial rolling stock inventory event at station k .

In the following, the *rolling stock activities* are defined. These activities represent the transfer of rolling stock from an ending train to a starting train, or the very first use of a rolling stock composition.

- For an event $e \in E_{\text{start}}^k$, denoting a start of a train from station k :
 - The set $\mathcal{A}^-(e)$ contains a subset of rolling stock activities from each event $f \in E_{\text{end}}^k$ to event e if events e and f are using the same type of rolling stock, and from the inventory event e_{rol}^k to event e . For this class of activities, constraint set (6) implies that a train can only depart if an earlier train has ended at that station, or if there is still a rolling stock composition available in the inventory.
 - The set $\mathcal{A}^+(e)$ is not constructed since the next event using the same rolling stock is known (the arrival event of train t_e at the next station) and so no decision variables are needed.
- For an event $e \in E_{\text{end}}^k$,
 - The set $\mathcal{A}^-(e)$ is not constructed since the previous event using the same rolling stock is known (the departure event of train t_e from the previous station) and so no decision variables are needed.
 - The set $\mathcal{A}^+(e)$ contains a subset of rolling stock activities to each event $f \in E_{\text{start}}^k$. Here constraint set (7) implies that, if a train has ended at a station, its rolling stock can be assigned to at most one other train.
- The set $\mathcal{A}^+(e_{\text{rol}}^k)$ contains a subset of rolling stock activities from the inventory event e_{rol}^k to each event $f \in E_{\text{start}}^k$. Here constraint set (8) implies that at most $i_{e_{\text{rol}}^k}$ trains may start from station k by using rolling stock compositions available at station k .

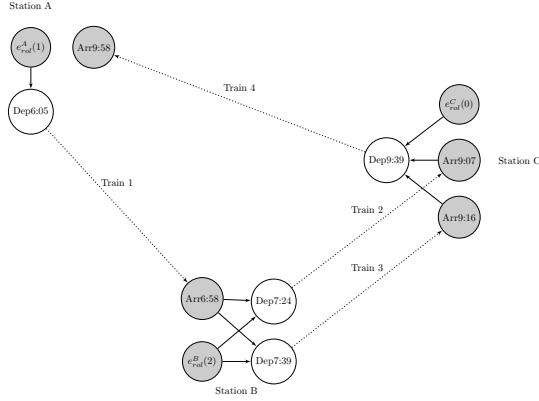


Figure 11: Example Event-Activity graph for rolling stock capacity

- For arrival and departure events which do not represent the start of a train nor the end of a train, no rolling stock activities are constructed. Because we do not allow trains to switch rolling stock during one train service, all events of a train service use the same rolling stock. We use this information and do not need to model it by decision variables.

Furthermore, constraint set (9) models the turnaround time required to transfer rolling stock from an ending train to a starting train, and the time required for shunting activities if a starting train uses rolling stock from the inventory.

Figure 11 shows an example of a graph of rolling stock activities. The events are represented by nodes and ordered by location. The time at which an event takes place is assumed to be fixed. The capacity of an inventory event is given between brackets. The rolling stock activities are represented by arcs and the train activities are also included and represented by dotted lines.

Figures 12 and 13 show with bold arcs two of the feasible solutions for the rolling stock capacity problem shown in Figure 11. Note that again the dashed arcs are train activities. A selected path from one of the inventory events (e_{rol}^A , e_{rol}^B and e_{rol}^C) indicates in which order events use a rolling stock composition. For example in Figure 12, Train 1 uses a rolling stock composition from the inventory of Station A which will be left in Station B. Trains 2 and 3 both get their rolling stock composition from the inventory at Station B and Train 4 uses the same rolling stock composition as Train 2.

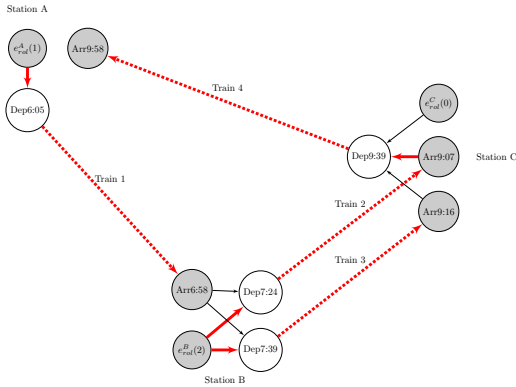


Figure 12: Solution 1

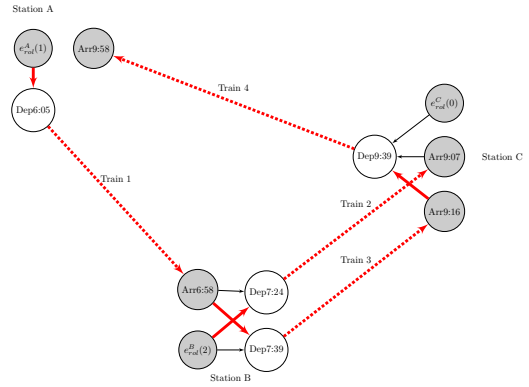


Figure 13: Solution 2

If a station has a shunting yard, then the rolling stock composition of an ending train can be moved to the shunting yard before it is used by a starting train again. To model this, we add the rolling stock activities which include a back-and-forth to the shunting yard to

the subset of station activities. If the rolling stock composition goes to the shunting yard (with infinite capacity), then a track in the station becomes available, and a track needs to be available at the moment the rolling stock composition comes back from the shunting yard. The subset of rolling stock activities does not change to ensure that still every train has a rolling stock composition.

Furthermore, in the operations, it is preferred to have a regular turning pattern for the rolling stock units, since this requires less communication with the shunting crews. Each train belongs to a *series*, which is a set of train services that have the same departure station, the same stops and the same arrival station. In a cyclic timetable in every cycle one train of each train series runs. Trains belonging to the same train series traveling in the same direction belong to the same *subseries*. A *turning pattern* is a pair of subseries $(s, s') \in S \times S$, where S denotes the set of all subseries. A turning pattern represents the transfer of rolling stock from a train belonging to s to a train belonging to s' . A rolling stock activity a corresponding to the transfer of rolling stock between two trains therefore belongs to a turning pattern. This turning pattern is denoted by w_a . Furthermore, W is the set of all possible turning patterns, W_s is the set of all turning patterns containing subseries s , and u_w is a decision variable such that $u_w = 1$ if the activities must correspond to turning pattern w is selected and 0 otherwise. Then constraint sets (13)-(15) ensure a regular turning pattern in which rolling stock units of ending trains of a certain subseries are used only by starting trains which are from the same subseries.

$$\sum_{w \in W_s} u_w \leq 1 \quad \forall s \in S \quad (13)$$

$$z_a \leq u_{w_a} \quad \forall a = (e, f) \in A_{\text{rol}} \quad (14)$$

$$u_w \in \{0, 1\} \quad \forall w \in W \quad (15)$$

4.2 Blockage of an open track section

The disruptions considered in this paper consist of blockages of tracks for a known duration, where a *partial blockage* results in the temporary unavailability of a subset of tracks between two stations. A special case of a partial blockage is a *full blockage* where all tracks between two stations are blocked.

In our test instances we only consider blockages of tracks within an open track section. However, blockages of tracks within a station can be handled in the same way. Another type of disruption is a defective rolling stock unit. Although this also results in a track blockage, another way to handle this kind of disruptions is by increasing the arrival time of this specific train with the time it will take to fix the rolling stock.

Here we limit ourselves to blockages of tracks. Due to the blockage of the tracks, some additional events and activities should be added. The time τ_1 at which the disruption starts and the time τ_2 at which the disruption ends is given as input, together with a list of open track sections which are partially blocked. If there is more than one track in the open track section connecting two stations, information about which tracks are blocked is given as well.

We do not assume that the network is empty at the moment τ_1 the disruption starts. Therefore we should also take into account what has happened before time τ_1 , but we cannot change this.

4.2.1 Disrupted trains

Trains running over the disrupted area need special attention. For each train, a station is classified as a stopping station if a stop has to be made, or as a pass-through station otherwise.

Assume there is a partial blockage of an open track section on which train $t \in T$ is scheduled. Let station k (ℓ) be the last (first) stopping station of train t before (after) the blocked section. Then, if a train t has an arrival time at station k during the disruption, the events that would have been associated with train t (in case of no disruption) are now partitioned and associated with three new trains, namely trains t_α , t_β and t_γ .

Let e^{start} and e^{end} denote the first departure event and the last arrival event of train t , respectively, let e_k^{arr} and e_ℓ^{arr} denote the arrival events at k and ℓ , respectively, and let e_k^{dep} and e_ℓ^{dep} denote the departure events from k and ℓ , respectively. Train t_α has events e^{start} to e_k^{arr} , train t_β has events e_k^{dep} to e_ℓ^{arr} and train t_γ has events e_ℓ^{dep} to e^{end} .

Furthermore, constraint sets (16) and (17) are included to ensure that if train t_β (the train over the disrupted area) runs, then both trains t_α and t_γ also run. This is equivalent to running the original train t . If t_β is cancelled, however, then t_α and t_γ may run or be cancelled, independently of each other. This is modelled as follows:

$$y_{t_\beta} \geq y_{t_\alpha} \quad (16)$$

$$y_{t_\beta} \geq y_{t_\gamma} \quad (17)$$

4.2.2 Additional rolling stock activities

For trains t_α and t_γ , all activities are defined as discussed in Section 4.1. However, in order to ensure that trains t_α , t_β and t_γ use the same rolling stock if they all run, a rolling stock activity (with a duration 0) is defined from events e_k^{arr} to e_k^{dep} and from events e_ℓ^{arr} to e_ℓ^{dep} . Furthermore, no other rolling stock activities are defined for train t_β . For event e_ℓ^{arr} , constraint (7) becomes an equality constraint. Hence in the case where none of these three trains is cancelled, they all use the same rolling stock, whereas if train t_β is cancelled, the rolling stock units of trains t_α and t_γ may turn on other trains at stations k and ℓ respectively.

4.2.3 Additional open track section activities

If train t leaves from station k before the disruption, but the disruption starts before train t reaches station ℓ , then it is assumed that train t continues along its original route. The reason for this assumption is that decisions to be made for this kind of trains depend on microscopic details of the disruption, such as whether the train is before or after the broken switches or overhead wires, or whether this is the train whose rolling stock is malfunctioning. In these cases, the events associated with train t are the same as in the case of no disruption. The open track section activities for such a train are, however, defined differently, in that open track section activities of the arrival event of t at the end of the disrupted track section are only defined for events which take place after the disruption (after τ_2). This ensures that the blocked tracks are only used again after the disruption has ended.

The assumption that these trains continue as planned is a choice made by the authors. It does not affect the model itself. In specific situations also another choice could be made, for example that the trains in the disrupted area are cancelled, or return to station k .

4.2.4 Balancing directions

If the number of tracks is reduced, then more trains can run if all trains run in the same direction. To avoid such unbalanced timetables, which are not preferred in practice, the following constraints can be added to the formulation for each pair of subseries $(s, s') \in S \times S$ for which s and s' belong to the same train series, but differ in direction.

$$\sum_{t \in s} y_t \leq \sum_{t \in s'} y_t + 1 \quad (18)$$

$$\sum_{t \in s'} y_t \leq \sum_{t \in s} y_t + 1 \quad (19)$$

These constraints ensure that for every train series the number of trains in one direction cannot exceed the number of trains in the other direction by more than one. Note that such constraints are also (partially) enforced already by the rolling stock circulation.

4.3 Managing all stages of the disruption process

To correctly handle the disruption, the time of all events that took place before the start of the disruption is fixed, and trains associated with these events cannot be cancelled, since they are already running. Moreover, it is preferable that the disruption does not affect the timetable for the complete day. Therefore it is assumed that at some point in time after the disruption has ended the trains should run according to their original timetable again. For this purpose a time $\tau_3 > \tau_2$ is specified such that any event that takes place after τ_3 cannot be delayed and such that a train starting after τ_3 cannot be cancelled.

The set of events E therefore only needs to contain events that are scheduled to take place after τ_1 and before τ_3 , together with some events outside this range in order to correctly model the availability of capacities at time τ_1 , and to ensure a smooth recovery to the original timetable at time τ_3 . Without going into too much detail, especially events which took place after τ_1 minus the minimum headway or turn around time, as well as events which have to take place before τ_3 plus the minimum headway or turn around time should be considered.

Secondly, after the disruption there must be enough rolling stock at every station to run the timetable for the remainder of the day. Therefore, at every station an inventory event is added which is scheduled at the time the disruption is over, connected with the rolling stock activities. The number of selected rolling stock activities to this event must equal the number of rolling stock compositions there are normally (without disruption) at that station at that time.

4.4 Reroutings

One of the contributions of this paper is that we include the possibility for trains to be rerouted in order to avoid the disrupted area if the network under consideration allows this. The advantage of having an option to reroute a train is that passengers wishing to use this train do not have to reroute themselves (possibly with some transfers). Furthermore, passengers may experience a smaller amount of delay, and the normal trains running on the rerouted area may be less crowded.

To incorporate this functionality, for each train t which is scheduled to travel through a disrupted area an alternative list of stations between station k (the last stop before the blockage) and station ℓ (the first stop after the blockage) can be provided as rerouting option. Then, in addition to the new trains t_α , t_β and t_γ (compensating for the blocked area), a fourth train t_δ is defined which runs on the rerouted path. For each station on the rerouted path (apart from stations k and ℓ), an arrival event and a departure event is associated with train t_δ , while at station k a departure event representing the start of the train is associated with t_δ , and at station ℓ an arrival event representing the end of the train is associated with t_δ .

For train t_δ the scheduled time of the departure event at stations k is the same as for train t_β , while the arrival and departure times for the other events of t_δ are determined using information on the minimum running times necessary between the stations on the rerouted path. In order to ensure that at most one of the two trains t_β and t_δ runs, and that, if one of them runs, then both t_α and t_γ also run, constraint sets (20)-(21) replace constraint sets (16)-(17) for train t .

$$y_{t_\beta} + y_{t_\delta} \geq 1 + y_{t_\alpha} \quad (20)$$

$$y_{t_\beta} + y_{t_\delta} \geq 1 + y_{t_\gamma} \quad (21)$$

Furthermore, to handle a penalty for rerouting, a variable p_t is introduced for all trains t_β which have a rerouting option t_δ . The decision variable $p_t = 1$ if the train is cancelled and not rerouted, and $p_t = 0$ if the train runs as planned or is rerouted. To model this, constraints (22) are added.

$$y_{t_\beta} + y_{t_\delta} \leq 1 + p_{t_\beta} \quad (22)$$

In the objective function $\lambda_{t_\beta} y_{t_\beta}$ is replaced by $\theta \cdot \lambda_{t_\beta} y_{t_\beta} + (1 - \theta) \cdot \lambda_{t_\beta} p_{t_\beta}$ where θ is a parameter between 0 and 1 indicating the balance between the cost of rerouting and cancelling a train. Furthermore $\lambda_{t_\delta} = 0$ for the rerouted copy of the train. These settings assure that if train t_β runs as planned ($y_{t_\beta} = 0$, $y_{t_\delta} = 1$ and $p_{t_\beta} = 0$), the costs in the objective function are 0. If train t_β is rerouted ($y_{t_\beta} = 1$, $y_{t_\delta} = 0$ and $p_{t_\beta} = 0$), the costs are $\theta \cdot \lambda_{t_\beta}$ and if train t_β is cancelled and not rerouted ($y_{t_\beta} = 1$, $y_{t_\delta} = 1$ and $p_{t_\beta} = 1$), the costs are λ_{t_β} .

5 Computational experiments

In order to test our approach, we carried out computational experiments on part of the Dutch railway network. The trains considered in this region are trains of Netherlands Railways, which is the major railway operator of the Netherlands. The mathematical model is solved to

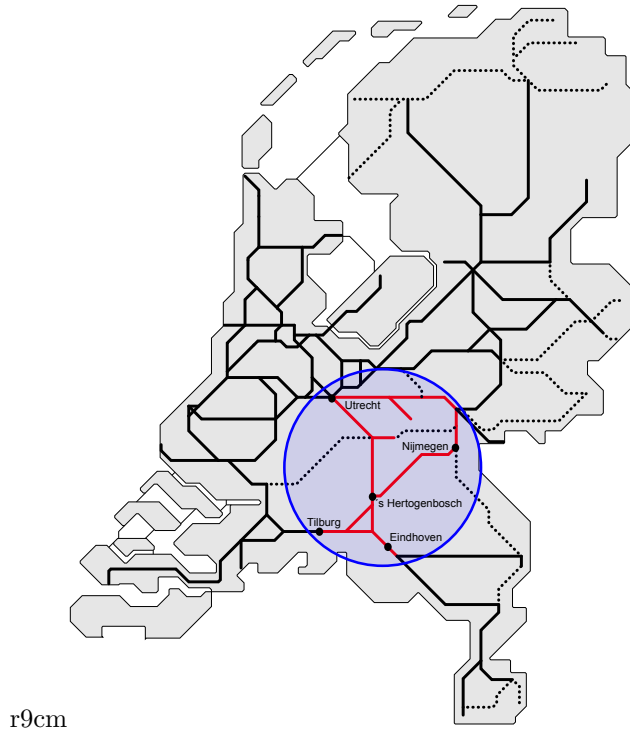


Figure 14: Overview of the Dutch railway network

optimality (with a gap of 0.01%) by CPLEX 12.4 on a PC with an Intel Xeon with 3.1 GHz and 16 GB RAM.

5.1 Case description

For our computational tests we consider a heavily used part of the Dutch railway network which is indicated in Figure 14. This network consists of 39 stations. At some stations, mostly located at a double tracked section, trains cannot switch tracks. Therefore, trains cannot overtake each other in those stations. This means that those stations can be considered as part of an open track section and do not have to be included as a station. In our computational tests, we consider 26 stations/junctions where trains can switch tracks. Note that a junction has the same characteristics as a station, with the difference that trains do not have a scheduled stop there. Therefore, departures are allowed to take place there earlier than scheduled.

Furthermore, our network consists of 27 open track sections between the considered stations. Of these sections, 3 are single tracked, 21 are double tracked, 1 has three parallel tracks and 2 have four parallel tracks.

In total 6 intercity and 10 regional train series run (mostly twice an hour in each direction) on this network which results in more than 60 trains per hour. We only consider the minimum number of rolling stock compositions which is required to run all these trains. This means that any spare rolling stock compositions at the shunting yards are not considered. In total 61 rolling stock compositions are necessary to run the trains of the complete day.

5.2 Parameter settings

In the mathematical model there are parameters for events and parameters for activities. First the parameters for the events are considered. The scheduled event times q_e are copied from the timetable of Netherlands Railways and the capacities of the stations, open track sections and rolling stock inventories i_e are set conform the described network in Section 5.1.

For the maximum allowed delay d_e of an train event e we make a distinction between three types of events. Trains running at the start of the disruption may be delayed more than

a train which has not started yet. This prevents infeasibilities of the mathematical model, since running trains are not allowed to be cancelled. For events $e \in E_{\text{train}}$ of trains which are already running at the time the disruption starts we have $d_e = 30$ minutes, for events $e \in E_{\text{train}}$ of rerouted trains the maximum delay is equal to $d_e = 15$ minutes. For all other events $e \in E_{\text{train}}$ we have $d_e = \underline{d}$, where the value of \underline{d} varies between 0, 3, 5 and 10 minutes over the different experiments.

For an activity, the minimum duration of the activity L_a can have multiple meanings. The minimum running time of a train L_a is set equal to the scheduled running time in the timetable, and the minimum dwell time is equal to the scheduled dwell time with a maximum of 2 minutes. The minimum headway time on the open track sections L_a is equal to 2 minutes if the trains run in the same direction and 0 minutes if the trains run in opposite directions. Within the stations the minimum headway time L_a is equal to 2 minutes.

The time L_a required before a rolling stock composition of an ended train can be used for a new starting train is 5 minutes. If the activity between an ending and a starting train takes longer than 10 minutes at a station with a shunting yard, the rolling stock composition goes to the shunting yard and releases the station capacity 5 minutes after the train has ended and requires station capacity from 5 minutes before the next train starts.

In the objective function we have two penalties. Penalties for delaying trains and penalties for cancelling trains. The aim is to run as many trains as possible. Therefore cancelling a train is penalized much more than delaying one. Cancelling a train is penalized by 50 times the running time of the train. Furthermore for every event there is a penalty of 1 per delayed minute. If a train is rerouted the costs are 20% of the costs for cancelling that train such that rerouting is preferred over cancelling the train, but also that the original route is preferred over the alternative route.

This research does not focus on *how* the available capacities are used. Just the utilization of the capacity is maximized, and if multiple capacity allocations lead to the same utilization, then just one possible allocation is provided, since the objective function does not contain penalties on activities.

5.3 Disruption scenarios

In order to test the described approach, a large set of disruption scenarios is created. For all of the 27 open track sections we constructed 30 scenarios of full blockages where all tracks of that section are blocked, and for the 24 open track sections with more than one track, an additional set of 30 scenarios is constructed where only one track is blocked. The first scenario is a 2 hour disruption of that open track section starting at 9:00. Then we increase in every new scenario the start time of the 2 hour disruption by one minute. Since the timetable of Netherlands Railways in this region is a half an hour cyclic timetable, a disruption starting at 9:30 should be very similar to a disruption starting at 9:00. In total this leads to 810 scenarios of full blockages and 720 scenarios where only one track is blocked.

We take a buffer time of 1 hour into account before all trains should be able to run as planned again after the disruption is over. This means that with a 2 hour blockage all trains in a 3 hour period are taken into consideration in the timetable rescheduling.

5.4 Rerouting of trains

If there is a disruption between 's Hertogenbosch and Eindhoven, then trains can be rerouted via Tilburg (see Figure 14). For this case, 30 scenarios where all tracks are blocked and 30 scenarios where one of the two tracks is blocked are constructed in a similar way as described in Section 5.3. In these scenarios the trains of one of the Intercity lines were allowed to be rerouted. This intercity line runs twice an hour in each direction, which means that in a disruption of 2 hours 8 trains can be rerouted.

To include the rerouted trains in the timetable, more events need to get a new time x_e which differs from the scheduled time q_e . This means that it is harder to find the optimal solution. Therefore, for these cases, we first find the solution for the case where rerouting is not allowed, and then, in a second run, that solution is used as start solution for the case with rerouting. In the results, the presented computation times include the computation times of both runs.

	<i>Minutes of delay allowed</i>	<i>Cancelled trains</i>			<i>Partially cancelled trains</i>			<i>Cancelled minutes</i>			<i>Computation time (s)</i>		
		<i>Min</i>	<i>Avg</i>	<i>Max</i>	<i>Min</i>	<i>Avg</i>	<i>Max</i>	<i>Min</i>	<i>Avg</i>	<i>Max</i>	<i>Min</i>	<i>Avg</i>	<i>Max</i>
<i>Complete blockage (708 instances)</i>	<i>0</i>	0	1.9	12	0	4.4	18	28	385	947	3	3.2	12
	<i>3</i>	0	1.5	9	0	3.4	15	28	368	908	4	5.1	8
	<i>5</i>	0	1.4	10	0	3.2	15	28	361	887	4	5.9	11
	<i>10</i>	0	1.2	8	0	3.3	17	28	352	881	5	14.1	89
<i>One track blocked no balancing (720 instances)</i>	<i>0</i>	0	1.0	6	0	2.1	11	0	156	487	2	4.5	18
	<i>3</i>	0	0.6	6	0	1.5	8	0	115	351	4	7.1	33
	<i>5</i>	0	0.6	5	0	1.0	7	0	99	335	4	9.3	95
	<i>10</i>	0	0.5	4	0	0.7	6	0	68	286	6	45.2	971
<i>One track blocked with balancing (720 instances)</i>	<i>0</i>	0	0.9	6	0	1.9	11	0	166	487	3	4.8	14
	<i>3</i>	0	0.7	7	0	1.3	8	0	128	420	4	8.1	93
	<i>5</i>	0	0.8	5	0	1.0	6	0	108	420	5	10.7	119
	<i>10</i>	0	0.5	4	0	0.6	6	0	74	300	6	50.7	1992

Table 1: Results on cancellations and computation time

5.5 General results

In this section we discuss the results of all 1530 scenarios and in Section 5.6 we describe the results in which we included the option to reroute trains in 60 of the scenarios as described in Section 5.4.

Tables 1 and 2 contain the minimum, average and maximum values of the performance measures over all feasible scenarios for the complete blockages and the blockages of one track. There are two settings: (i) whether or not balancing constraints ((18)-(19)) were used in the case of a partial blockage, and (ii) how much delay (0, 3, 5 or 10 minutes) is allowed for trains that are not running at the start of the disruption. From now on, we refer with the term *allowed delays* only to the delays allowed for trains that are not running at the start of the disruption.

It turned out that 102 of the 1530 scenarios were infeasible in our model. Most of the time these infeasibilities are caused by the constraint that already running trains are not allowed to be cancelled anymore. If there is no capacity for those trains within the maximum allowed delay of 30 minutes, there does not exist a solution matching the restrictions. Another tight constraint is that the rolling stock inventory at each station must be equal to its planned inventory within one hour after the disruption.

Table 1 gives the results on the number of cancelled trains, the number of partially cancelled trains, the number of cancelled minutes and the computation times. In order to make the results maximally comparable, we count the number of the cancelled and partially cancelled trains, and the number of cancelled minutes as follows. A train t is partially cancelled if the disruption was on its route, and if only one of the trains t_α or t_γ (the parts before and after the disruption) is operated, while the other one is cancelled. If only the part on the disruption (train t_β) is cancelled, this is not considered as a (partially) cancelled train, since this part has to be cancelled inevitably in case of a complete blockage. The number of cancelled minutes is the sum of the scheduled lengths in minutes of the cancelled parts of all cancelled, partially cancelled, and inevitably cancelled trains, which is the main part in the objective function. Since the aim is to minimize the cancelled minutes and not to minimize the number of cancelled trains, it may happen that there are solutions where more trains (with a short duration) need to be cancelled to have less cancelled minutes.

For each instance, the total number of trains is approximately 180, since we compute a schedule for 3 hours in which more than 60 trains per hour run.

In case of a complete blockage, as shown in Table 1, if no delays are allowed, on average

	<i>Minutes of delay allowed</i>	<i>Delayed trains</i>			<i>Delayed events</i>			<i>Total maximum delay</i>			<i>Total delay</i>		
		<i>Min</i>	<i>Avg</i>	<i>Max</i>	<i>Min</i>	<i>Avg</i>	<i>Max</i>	<i>Min</i>	<i>Avg</i>	<i>Max</i>	<i>Min</i>	<i>Avg</i>	<i>Max</i>
<i>Complete blockage (708 instances)</i>	0	0	0.1	3	0	0.2	15	0	1.0	35	0	2	226
	3	0	1.4	19	0	6.0	78	0	3.1	45	0	11	169
	5	0	2.2	23	0	10.7	129	0	6.0	61	0	25.8	331
	10	0	2.8	20	0	12.9	131	0	11.1	106	0	49.2	838
<i>One track blocked no balancing (720 instances)</i>	0	0	0.1	2	0	1	19	0	2	30	0	13	388
	3	0	3.3	20	0	19	91	0	7	42	0	42	688
	5	0	5.4	27	0	33	188	0	15	83	0	87	720
	10	0	9.1	27	0	53	179	0	39	145	0	226	1000
<i>One track blocked with balancing (720 instances)</i>	0	0	0.1	2	0	1	17	0	1	30	0	11	387
	3	0	3.0	20	0	18	91	0	7	57	0	40	686
	5	0	4.9	20	0	30	141	0	13	72	0	78	453
	10	0	8.3	33	0	50	233	0	35	170	0	208	1190

Table 2: Results on delays experienced

385 minutes (6.5 hours) of trains are cancelled. Allowing up to 10 minutes of delay, reduces the amount of cancelled minutes by more than 8%.

If only one track is blocked, we see that the effect of allowing more delay is much higher. Allowing 10 minutes of delay decreases the number of cancelled minutes on average with more than 50% (from 156 to 68 minutes). The results also demonstrate that in some parts of the network it is still possible to run all trains if only one track is out of service. This may be deduced from the fact that there are solutions with 0 cancelled minutes.

The price of forcing the new schedule to be a regular one by including balancing constraints, in which the number of trains per train series in one direction differs at most one from the number of trains in the other direction, is an increase of around 10% in the number of cancelled minutes.

The largest computation time is less than 2 minutes in case a maximum delay up to 5 minutes is allowed. This means that our approach is able to find advanced timetables with a high probability of having a feasible rolling stock schedule within a computation time which is reasonable in practice.

We can even find better solutions by allowing more delays (up to 10 minutes). For the complete blockage case, the computation times are still below 2 minutes then. However, in case of a single track blockage, allowing delays for trains up to 10 minutes can increase the computation time to values which are not usable for practice. However, on average the computation time is still relatively low and in both variants (with and without balancing constraints) less than 13 out of the 720 cases had a computation time larger than 5 minutes.

For the cases with a computation time larger than 5 minutes we have computed the results with a maximum CPLEX CPU time of 5 minutes. In case of no balancing constraints, on average the number of cancelled minutes was 13% higher than in the optimal solution computed without the time limit. The maximum increase in cancelled minutes was 50%. In cases with balancing constraints the average increase in cancelled minutes was 27% and the maximum increase was 65%.

The decision on which allowed delay should be preferred can be taken by the dispatchers, also based on the available computing time for obtaining the new timetable. A good trade-off between solution quality and computing time corresponds to an allowed delay of 5 minutes.

Another strategy could be to first find the solution for the problem in which a delay of 5 minutes is allowed and then use the found solution as a feasible start solution for the problem in which 10 minutes of delay is allowed. In this strategy the dispatchers can give the approach (a maximum) time to improve the solution found with the setting of a maximum delay of 5 minutes.

	Minutes of delay allowed	Rerouted trains			Cancelled minutes			Computation time (s)		
		Min	Avg	Max	Min	Avg	Max	Min	Avg	Max
<i>Complete blockage no rerouting (29 instances)</i>	0				497	521	544	3	3.0	4
	3				469	512	536	4	4.9	6
	5				469	506	536	5	5.6	8
	10				451	488	525	7	10.1	16
<i>Complete blockage with rerouting no balancing (29 instances)</i>	0	4	7.4	8	379	419	489	6	6.8	7
	3	4	7.7	8	375	410	462	10	11.5	20
	5	4	7.7	8	351	405	462	11	14.6	25
	10	4	7.7	8	335	391	443	19	37.2	75
<i>Complete blockage with rerouting with balancing (29 instances)</i>	0	1	7.5	8	379	423	543	6	7.0	11
	3	1	7.5	8	375	414	517	10	11.6	20
	5	1	7.5	8	351	409	517	12	15.7	43
	10	8	8	8	335	393	468	20	39.9	81
<i>One track blocked no rerouting no balancing (30 instances)</i>	0				256	274	304	3	5.4	10
	3				193	221	251	5	7.8	29
	5				151	168	189	7	14.6	39
	10				60	88	96	19	71.8	284
<i>One track blocked with rerouting no balancing (30 instances)</i>	0	6	6.8	7	196	217	243	8	15.1	25
	3	4	4.0	4	178	193	218	17	23.8	40
	5	3	4.0	4	96	103	124	15	42.9	109
	10	5	6.8	8	60	88	96	48	170.9	407
<i>One track blocked no rerouting with balancing (30 instances)</i>	0				288	311	334	5	7.2	11
	3				243	283	326	8	16.8	49
	5				243	257	307	13	27.8	55
	10				94	129	153	30	144.4	396
<i>One track blocked with rerouting with balancing (30 instances)</i>	0	1	6.6	7	206	238	326	9	17.8	55
	3	4	4.0	4	180	193	218	20	32.1	60
	5	4	5.1	8	163	182	218	30	74.0	141
	10	5	5.1	6	94	121	137	165	284.7	830

Table 3: Results with reroutings

Table 2 gives the results on the number of delayed trains, the number of delayed events, the sum of the maximum delay experienced by each train on its route (*Total maximum delay*), as well as the sum of the delays of all events. As expected, allowing delays increases the total delay of the trains and events, but, as can be seen in Table 1, it also reduces the number of cancelled trains.

5.6 Results with the option to reroute trains

Table 3 presents the results of the 60 scenarios in which there is a rerouting option for intercity trains. In the case of a complete disruption there was one instance which was not feasible for all settings. Therefore, we removed this instance from the results. This leaves us with 29 scenarios in case all tracks are blocked and 30 scenarios in case only one track is blocked.

One train series is allowed to be rerouted. Since this train series runs twice an hour in each direction and since we are dealing with a 2 hour disruption, the total number of trains which can be rerouted is equal to $2 \times 2 \times 2 = 8$.

In Table 3 we compare the results in which rerouting is not allowed with the results of the same instances in which rerouting is allowed. We can see that, in case of complete blockages,

our approach is able to reduce the number of cancelled minutes considerably (approximately 20%) if we allow reroutings. Computation times stay below 2 minutes. In addition, having the option to reroute trains reduces the number of cancelled minutes much more than simply allowing larger maximum delays.

If all tracks are blocked, we originally did not have to include the balancing constraints since, in that case, no trains at all run over the blocked area. However, in case we allow reroutings, it is worthwhile to include balancing constraints for the rerouted trains to ensure that trains are rerouted evenly in each direction. The results demonstrate that in the case of a complete blockage with reroutings these balancing constraints do not have much influence on the solution and on the computation time. Therefore we recommend to include these balancing constraints. This is due to the fact that it seems to be possible to reroute all trains in many cases. If all trains are rerouted, then the result is automatically a balanced solution with 4 rerouted trains in each direction.

Also in case only one of the tracks is blocked, allowing reroutings can reduce the number of cancelled minutes considerably (up to 38%). However, if delays of 10 minutes are allowed, we discover that our approach without reroutings is able to find solutions with the same number of cancelled minutes. In case rerouting is allowed, the reroutings are used to reduce the delays of the trains.

In case of a blockage of only one track, the balancing constraints have a larger effect than in case of a complete blockage. Especially if we allow delays of 5 or 10 minutes, adding the balancing constraints leads to an increase of about 50% in the number of cancelled minutes, independently of whether or not reroutings were allowed.

Also in case reroutings are allowed, our approach is able to quickly find an optimal solution, as long as we do not allow more than 5 minutes delay. If dispatchers want to allow larger delays (10 minutes) the computation times increase, especially in the case with balancing constraints.

6 Conclusions

In this paper we introduced an Integer Linear Program (ILP) to solve the real-time railway timetable rescheduling problem for a railway network. The railway timetable rescheduling problem considered in this paper has a macroscopic view on the infrastructure network which consists of stations and open track sections with certain numbers of tracks. Furthermore, constraints on the available rolling stock are also considered in order to have a high probability that there is a feasible rolling stock schedule for the new timetable. The possibility of rerouting trains in order to reduce the number of cancelled and delayed trains is also considered. In addition, all stages of the disruption management process (from the start of the disruption to the time the normal situation is restored) are taken into account.

The ILP is modeled as an event activity network in which each event represents an arrival or a departure of a train, and in which an activity refers to passing on a resource unit from one event to another event. The resources considered are the tracks in the open track sections, the tracks in the stations, and the rolling stock compositions.

Computational tests are performed on a heavily used part of the Dutch railway network. Solutions are provided within computation times which are very well suitable for use in practice. Most of our cases can be solved within 2 minutes of computation time. The results show that a smaller number of trains needs to be cancelled and the number of cancelled minutes is significantly reduced if we allow to slightly delay or to reroute some trains.

Our approach turns out to be able to handle, in short computing time, every state of the network at the time the disruption starts, and to decrease cancellations and delays of trains. This makes our approach much more flexible and efficient than the current practice of using contingency plans.

References

- R. Acuna-Agost, P. Michelon, D. Feillet, and S. Gueye. Sapi: Statistical analysis of propagation of incidents. a new approach for rescheduling trains after disruptions. *European Journal of Operational Research*, 215:227 – 243, 2011a.
- R. Acuna-Agost, P. Michelon, D. Feillet, and S. Gueye. A mip-based local search method for the railway rescheduling problem. *Networks*, 57:69–86, 2011b.
- A. Albrecht, D. Panton, and D. Lee. Rescheduling rail networks with maintenance disruptions using problem space search. *Computers & Operations Research*, 40:703 – 712, 2013.
- M. Boccia, C. Mannino, and I. Vasilyev. The dispatching problem on multitrack territories: Heuristic approaches based on mixed integer linear programming. *Networks*, 62:315–326, 2013.
- P. Brucker, S. Heitmann, and S. Knust. Scheduling railway traffic at a construction site. *OR Spectrum*, 24:19–30, 2002.
- V. Cacchiani and P. Toth. Nominal and robust train timetabling problems. *European Journal of Operational Research*, 219:727–737, 2012.
- V. Cacchiani, D. Huisman, M. Kidd, L. Kroon, P. Toth, L. Veelenturf, and J. Wagenaar. An overview of recovery models and algorithms for real-time railway rescheduling. *Transportation Research Part B: Methodological*, 63:15 – 37, 2014.
- G. Caimi, M. Fuchsberger, M. Laumanns, and M. Lthi. A model predictive control approach for discrete-time rescheduling in complex central railway station areas. *Computers & Operations Research*, 39:2578 – 2593, 2012.
- A. Caprara, L. Kroon, M. Monaci, M. Peeters, and P. Toth. Passenger railway optimization. In C. Barnhart and G. Laporte, editors, *Transportation*, volume 14 of *Handbooks in Operations Research and Management Science*, pages 129 – 187. Elsevier, 2007.
- A. Caprara, L. Kroon, and P. Toth. Optimization problems in passenger railway systems. In J. Cochran, L. Cox, P. Keskinocak, J. Kharoufed, and J. Cole Smith, editors, *Wiley Encyclopedia of Operations Research and Management Science*, volume 6, pages 3896–3905. John Wiley & Sons, Inc., 2011.
- F. Corman, A. D’Ariano, D. Pacciarelli, and M. Pranzo. Evaluation of a green wave policy in real-time railway traffic management. *Transportation Research Part C: Emerging Technologies*, 17:607–616, 2009.
- F. Corman, A. D’Ariano, D. Pacciarelli, and M. Pranzo. A tabu search algorithm for rerouting trains during rail operations. *Transportation Research Part B: Methodological*, 44:175–192, 2010.
- F. Corman, A. D’Ariano, I. Hansen, D. Pacciarelli, and M. Pranzo. Dispatching trains during seriously disrupted traffic situations. In *Proceedings of the IEEE International Conference on Networking, Sensing and Control*, pages 323–328, 2011.
- F. Corman, A. D’Ariano, D. Pacciarelli, and M. Pranzo. Bi-objective conflict detection and resolution in railway traffic management. *Transportation Research Part C: Emerging Technologies*, 20:79–94, 2012.
- A. D’Ariano, D. Pacciarelli, and M. Pranzo. A branch and bound algorithm for scheduling trains on a railway network. *European Journal of Operational Research*, 183:643–657, 2007.
- A. D’Ariano, F. Corman, D. Pacciarelli, and M. Pranzo. Reordering and local rerouting strategies to manage train traffic in real time. *Transportation Science*, 42:405–419, 2008a.
- A. D’Ariano, D. Pacciarelli, and M. Pranzo. Assessment of flexible timetables in real-time traffic management of a railway bottleneck. *Transportation Research Part C: Emerging Technologies*, 16:232–245, 2008b.

- T. Dollevoet, D. Huisman, M. Schmidt, and A. Schöbel. Delay management with rerouting of passengers. *Transportation Science*, 46:74–89, 2012.
- P. Kecman, F. Corman, A. D'Ariano, and R. Goverde. Rescheduling models for railway traffic management in large-scale networks. *Public Transport*, 5:95–123, 2013.
- I. Louwerse and D. Huisman. Adjusting a railway timetable in case of partial or complete blockades. *European Journal of Operational Research*, 235:583 – 593, 2014.
- R. Lusby, J. Larsen, M. Ehrgott, and D. Ryan. Railway track allocation: models and methods. *OR Spectrum*, 33:843–883, 2011.
- R. M. Lusby, J. Larsen, M. Ehrgott, and D. M. Ryan. A set packing inspired method for real-time junction train routing. *Computers & Operations Research*, 40:713 – 724, 2013.
- C. Mannino and A. Mascis. Optimal real-time traffic control in metro stations. *Operations Research*, 57:1026–1039, 2009.
- A. Schöbel. Capacity constraints in delay management. *Public Transport*, 1:135–154, 2009.
- J. Törnquist and J. A. Persson. N-tracked railway traffic re-scheduling during disturbances. *Transportation Research Part B: Methodological*, 41:342 – 362, 2007.
- M. Wiklund. Serious breakdowns in the track infrastructure - calculation of the effects on rail traffic. In *Proceedings of the 2nd International Seminar on Railway Operations Modelling and Analysis - RailHannover 2007*, Hannover, Germany, Mar 2007.

ERIM Report Series <i>Research in Management</i>	
ERIM Report Series reference number	ERS-2014-010-LIS
Date of publication	2014-07-25
Version	25-07-2014
Number of pages	24
Persistent URL for paper	http://hdl.handle.net/1765/51678
Email address corresponding author	lveelenturf@rsm.nl
Address	Erasmus Research Institute of Management (ERIM) RSM Erasmus University / Erasmus School of Economics Erasmus University Rotterdam PO Box 1738 3000 DR Rotterdam, The Netherlands Phone: +31104081182 Fax: +31104089640 Email: info@erim.eur.nl Internet: http://www.erim.eur.nl
Availability	The ERIM Report Series is distributed through the following platforms: RePub, the EUR institutional repository Social Science Research Network (SSRN) Research Papers in Economics (RePEc)
Classifications	The electronic versions of the papers in the ERIM Report Series contain bibliographic metadata from the following classification systems: Library of Congress Classification (LCC) Journal of Economic Literature (JEL) ACM Computing Classification System Inspec Classification Scheme (ICS)