Alma Mater Studiorum Università di Bologna
Archivio istituzionale della ricerca

A railway timetable rescheduling approach for handling large-scale disruptions

(Article begins on next page)

15 July 2024

A railway timetable rescheduling approach for handling large scale disruptions

Lucas P. Veelenturf

School of Industrial Engineering, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands, l.p.veelenturf@tue.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

On a daily basis, large scale 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 train services on a macroscopic level in a railway network. An integer linear programming model is formulated for solving the timetable rescheduling problem, which minimizes the number of cancelled and delayed train services while adhering to infrastructure and rolling stock capacity constraints. The possibility of rerouting train services in order to reduce the number of cancelled and delayed train services 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 the model is able to find optimal solutions in short computation times. This makes the approach applicable for use in practice.

1. Introduction

The occurrence of unexpected large-scale disruptions, such as the unavailability of railway track segments due to broken overhead wires or rolling stock breakdowns, causes delays and cancellations of train services 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 train services and even train cancellations. The derived timetable is input for 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 train services. 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. In particular, our approach guarantees that there is a rolling stock composition (a set of coupled rolling stock units) assigned to each train service in the timetable. However, we only ensure that there is *a* composition, we do not consider *which* composition. This means that we neglect all characteristics of the rolling stock compositions (e.g. we neglect of which rolling stock units the composition consist). Since we keep track of rolling stock at the composition level and not at the unit level, this could lead to different than expected inventories of rolling stock at the end of the day. Furthermore, maintenance appointments of rolling stock units are neglected and infinite track capacity at shunting yards is assumed. Therefore, rolling stock rescheduling might be necessary after the execution of our approach for the timetable rescheduling 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 train services can still be operated with the available infrastructure capacity. Thereafter, small conflicts at the signaling level should be detected and solved by slightly delaying some train services.

We consider large-scale 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, such disruptions occur on a daily basis. Therefore 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-scale disruptions on a real-world railway network. The formulation takes into account constraints that allow to partially integrate the timetable rescheduling problem with the rolling stock rescheduling phase. In particular, we consider a railway network with a cyclic timetable, i.e. the schedule of the train services is repeated every given time period (for example every hour).

This approach generates a new timetable from the start of a disruption until 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 train services along alternative geographical paths in the network in order to reduce the number of train services 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 the event-activity network is defined. In Section 5 an ILP formulation is given. Section 6 is devoted to the computational results based on instances of Netherlands Railways, and the conclusions are discussed in Section 7.

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 train services in a time horizon from 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 train services or train series, instead of large-scale disruptions, as is the case in our paper. A *train series*, sometimes called *train line*, is a set of train services that have the same departure station, the same stops and the same arrival station. 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. (2007) and D'Ariano et al. (2008a). 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 train services. 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. In Corman et al. (2011a), an optimization framework, also incorporated in ROMA, is presented to reschedule train services according to different classes of priority, which are used to group train services with similar characteristics. The algorithm iteratively schedules trains for an ordered set of priority classes, from the highest one to the lowest one, while preserving the solution quality of the higher priority classes. The algorithm is tested on instances considering the area around Utrecht Central station in the Netherlands.

Other successful approaches based on mathematical models, different from the alternative graph method used in ROMA, are proposed in the literature. An ILP model is proposed in Meng and Zhou (2014) for train rescheduling on an N-track network. It contains cumulative flow variables and allows for train rerouting. The proposed method is tested on an instance adapted from the INFORMS RAS problem competition in 2012.

In Sato et al. (2013), a Mixed Integer Programming (MIP) model is proposed for minimizing passenger inconvenience, evaluated as traveling time on board, the waiting time at platforms and

the number of transfers. The train rescheduling is obtained by changing the departing order of the train services, the assignment of platforms to trains in stations, the train types and the rolling stock schedules at termini. The method is tested on a weekend timetable of the Chuo Line in Tokyo.

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. (2011a), Acuna-Agost et al. (2011b), Dollevoet et al. (2012), Kecman et al. (2013), Schöbel (2009) and Törnquist and Persson (2007).

Only few and recent works deal with large-scale disruptions. In Albrecht et al. (2013) the authors consider disruptions due to unexpected track maintenance extensions, arising when 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. (2011b), 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 Corman et al. (2014a), a framework based on the ROMA tool is used for the evaluation of timetables under disturbances. In particular, regular planned timetables are compared to timetables designed to be robust against disturbances such as multiple train delays, speed restrictions and blocked tracks. Several performance indicators are proposed for this evaluation, i.e. the total delay (i.e. the difference between the calculated train arrival time and the scheduled time at all station stops), the consecutive delay (i.e. the propagated delay) and the so-called generalized travel time which consists of the weighted sum of the waiting time at a station before the travel, the in-vehicle travel time and the waiting time for transfers. In the case of blocked tracks, the transition from the disrupted situation back to the regular timetable is analyzed on the obtained results. The proposed framework is tested on the South-East part of the Dutch railway network.

In Corman et al. (2014b), centralized and distributed procedures for train rescheduling are compared and heuristic algorithms are proposed for coordinating the different dispatching areas. The algorithms are compared in the case of delays and blocked tracks. In the latter case the authors apply some modifications to the planned timetable, before executing the developed algorithms, to reroute trains in order to avoid the disrupted tracks, and then the algorithms are used to reschedule trains when delays occur. The algorithms are tested on the South-East part of the Dutch railway network.

In Jespersen-Groth et al. (2006), the problem of reinserting a cancelled train line, after a major disruption has occurred, is considered for DSB S-Tog, the suburban rail operator in the Copenhagen area in Denmark. Train service cancellation is an important recovery strategy in case of major disruptions. It is implemented by shunting the rolling stock to the depots. When an adequate level of regularity has been re-established, the cancelled train lines should be reinserted according to the original timetable. A MIP model for generating optimal reinsertion plans of an entire cancelled train line is proposed. Periodicity constraints are imposed in order to respect a given frequency interval.

Narayanaswami and Rangaraj (2013) study the problem of rescheduling train services on a single bi-directional line, when a disruption blocks the line between two stations for some time. The trains can meet and pass each other only in the railway stations. The duration of the disruption is assumed to be known at the time of rescheduling. The authors develop a MIP model for rescheduling the train services, with the goal of minimizing the weighted delay of all train services. The model is tested on a number of artificial instances with a small number of train services.

In Törnquist (2012) a greedy algorithm is proposed in order to quickly reschedule train services after a disruption. The railway network is divided into sections (lines or stations), each composed of a number of parallel tracks. An event corresponds to a train occupying a certain section at a certain time. The greedy algorithm iteratively searches for the best of all waiting events to execute next and builds up a tree of consecutive events. It performs a depth-first search, using an evaluation function to prioritise the events when conflicts arise, to quickly find a feasible solution by building up a complete branch of the tree. Then, it tries to improve the existing solution by backtracking to potential nodes in the existing tree and branching to find better solutions. The algorithm is tested on instances of Trafikverket, the Swedish Rail Administration.

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-scale disruptions. A case study involving a fire at the interlocking system of a station in Stockholm is considered.

As mentioned in the introduction, the rescheduled timetable is input to the rolling stock rescheduling phase, in which a new rolling stock allocation is derived, if the planned one turns out to be infeasible for the new timetable. In the ILP model proposed in this paper, we take into account constraints from the rolling stock rescheduling phase, with the goal of ensuring that for each train service rolling stock is available. There are studies in the literature that focus on the rolling stock rescheduling phase or that consist of integrated approaches. Among them, we mention the seminal work by Adenso-Dıaz et al. (1999). The authors develop an optimization model for real-time timetable and rolling stock rescheduling in case of large delays of one or more train services. The algorithm is tested on a regional railway network in Asturias provided by RENFE, the Spanish national railway company.

Almodóvar and García-Ródenas (2013) propose a real-time optimization method for rescheduling rolling stock when a huge and unexpected passenger demand exceeding the capacity occurs. The proposed method is based on a discrete-event simulation model. The proposed method is tested on a synthetic network as well as on a real-world network provided by RENFE.

Jespersen-Groth (2008) describes the problems existing in the railway planning process, from the strategic level to the real-time one, for DSB S-Tog, and focuses on rolling stock dispatching.

Another important aspect that is studied only in very few works in the literature is the transition from the original timetable to the rescheduled timetable, and then to the original planned timetable, once the disruption has been resolved. As explained in Louwerse and Huisman (2014), the railway disruption process consists of three phases: in the first phase there is a transition from the planned timetable to the rescheduled timetable; in the second phase the rescheduled timetable is operated, while the third phase consists of the transition from the rescheduled timetable to the original one. In Chu and Oetting (2013), the transition process is studied for the case in which disruption programs, i.e. sets of pre-defined dispatching measures, also known as contingency plans, are used. These measures consist of turning trains on both sides of the disrupted area and/or cancelling train services. The use of disruption programs can be seen as complementary to real-time traffic management tools (such as the one proposed in our paper), since disruption programs are determined before the disruption occurs. The authors analyze the transition phase, i.e. the phase from the beginning of the disruption until the situation becomes as foreseen in the disruption program. They propose a method for modeling capacity consumption during the transition phase in a realistic way, especially taking into account turnings within stations. Two large German urban railway networks are considered. We refer the reader to Louwerse and Huisman (2014) and Jespersen-Groth (2008) for further details on the different phases of the transition process.

Our work extends the ILP model presented in Louwerse and Huisman (2014), in which two instances of a double track line were considered and a partial or full track blockage was taken into

account. First of all, Louwerse and Huisman (2014) present a formulation in which a temporary timetable is computed only for the second phase of the disruption management process. They determine which train series can still run and which have to be cancelled or delayed in that phase. 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 ends. In addition, they assume that, at the time the disruption starts, the network is empty.

In our approach, on the contrary, we consider the complete disruption management process, i.e. from the start of the disruption to the time at which the normal situation is restored. In addition, we consider the current status/usage of the network at the time the disruption starts. Therefore we do not remove complete train series from the timetable but individual train services. Especially in the transition phase it is quite important to look at individual train services instead of train series.

Another new feature consists of dealing with a real-world railway network instead of one double track line. Furthermore, the number of tracks within the stations and between the stations is not limited to two. Train services are allowed to be scheduled to other tracks within the stations or between the stations with respect to the tracks they are originally scheduled to, and now all tracks may be used in both directions. Trains can only overtake each other if they run on different tracks. To prevent overtakings of trains running on the same track, additional constraints are considered. The model of Louwerse and Huisman (2014) does not contain constraints preventing this. They assume that the delays are not large enough to get solutions in which this happens. However, we believe that in our case, with a large network, these constraints are very important.

With respect to the approach by Louwerse and Huisman (2014), 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 the proposed model in order to increase the probability of getting a feasible rolling stock schedule.

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 train services as possible can still be operated. A train service is partially cancelled if it is only operated to a subset of the stations of its original route. In most situations this means that the train service ends at a location different than the planned end location or starts from a location different than the planned start location.

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

3.1. The disruption

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. The proposed ILP model does not change in the case of multiple track blockages, except that, if a train service traverses multiple blockages, more variables are needed to model it (see Section 5.4.1). However, for the computational results we only consider disruptions where tracks at one geographical location are blocked.

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

Trains which passed their last stop before the blocked track(s) 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 train services are scheduled on the railway network. Train services are characterized by a type, e.g. they can be regional or Intercity train services. Each train service needs a rolling stock composition, i.e. a set of coupled rolling stock units, of the same type as the train service.

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

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

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 running on the same track at the same time in the same direction, or

between two trains running at the same track consecutively in opposite directions. It is assumed that a train entering a station from an open track section is able to reach each track in the station regardless of the track of the open track section it is entering from.

The capacity of a station is characterised by the number of tracks it has. Each track within the station may only be occupied by one train service at any given time. Furthermore, after a train service has used a track in a station, a certain headway time needs to pass before another train service 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. An 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 operate as many train services as possible (and not store as many rolling stock units as possible).

At each station with a shunting yard, a limited number of rolling stock compositions is available at the start of the day. A train service has a rolling stock composition assigned to it for its entire duration, after which the rolling stock composition is moved to a shunting yard or used by another train service. 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 train services may share rolling stock only if they are of the same type. After a train service has ended, a minimum turnaround time is required before the rolling stock composition of that train service may be used by another train service.

4. Event-activity network

Each train service is represented by a set of events, which are departures or arrivals 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 train services T and an original timetable for these train services.

4.1. Event sets

The set $E = E_{\mathrm{train}} \cup E_{\mathrm{inv}}$ of events consists of a set E_{train} of *train events*, and a set E_{inv} of *inventory events*.

Each train event $e \in E_{\mathrm{train}}$ represents either a departure or an arrival at a certain station. The train service of which train event e is a departure or an arrival, is denoted by t_e. The scheduled time at which train event e takes places 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 an open track section (i.e. the number of tracks in the open track section), a station (i.e. the number of tracks at the station), or the shunting yard at the start of the day (i.e. the number of available rolling stock compositions). The number of resources made available by inventory event e is denoted by i_e. In Sections 4.2.1, 4.2.2 and 4.2.3, we will outline the specific subsets of E_{train} and E_{inv} used for each type of resource.

4.2. Activity sets

Activities, which correspond to arcs of N, represent the sequential use of a resource (track of an open track section, track of a station or rolling stock composition) by two events. More precisely, 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, but 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.

For each event e we divide the activities into two groups: in-activities and out-activities. The *in-activities* of event e are all ingoing activities in event e. They are denoted by the set: $A^-(e) = \{a = (f, e) \in A | f \in E\}$. 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 of the in-activities associated with that resource type only. In other words, we define *i)* one subset of activities $A^-_{\text{track}}(e) \subset A^-(e)$ into event e associated with open track section capacity, *ii)* one subset of activities $A^-_{\text{stat}}(e) \subset A^-(e)$ into event e associated with station capacity, and *iii)* one subset of activities $A^-_{\text{rol}}(e) \subset A^-(e)$ into event e associated with rolling stock availability.

The second group of activities of an event e are the *out-activities* which are all activities out of event 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. As for the in-activities, we define for each event e subsets $A^+_{\text{track}}(e)$, $A^+_{\text{stat}}(e)$ and $A^+_{\text{rol}}(e)$ of out-activities.

In order to keep the model formulation (see Section 5) general and not specialized for each type of resources, the set $\mathcal{A}^-(e) = \{A^-_{\text{track}}(e), A^-_{\text{stat}}(e), A^-_{\text{rol}}(e)\}$ collectively refers to the three in-activity subsets of an event e, and likewise we define $\mathcal{A}^+(e) = \{A^+_{\text{track}}(e), A^+_{\text{stat}}(e), A^+_{\text{rol}}(e)\}$. For some events
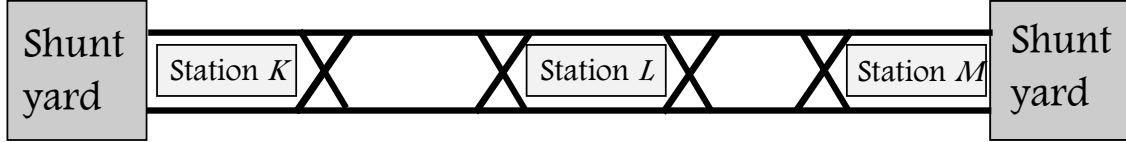
Figure 1   **Example of layout**

Service	Dep. K	Arr. L	Dep. L	Arr. M
Train 1	7:00	7:20	7:23	7:30
Train 2	7:55	8:15	8:18	8:25

Service	Dep. M	Arr. L	Dep. L	Arr. K
Train 3	7:16	7:23	7:26	7:46
Train 4	7:45	7:52	7:55	8:19

Table 1   **Example of timetable**

not all types of resource activities are necessary for modelling the resource constraints. In these cases some of the in-activity or out-activity subsets are empty.

For each train service, the minimum running time between two stations and the minimum dwelling time inside a station are also expressed by activities. Differently from the activities related to resources, in this case, the order of the events is fixed, and allows to use a single set of activities. Let A_{train} be the set of train activities $a = (e, f)$ which represent running or dwelling of a train between consecutive events e and f of the same train service.

The specific sets of events and activities to be used in the ILP model for each type of resources are described in detail in the following sections. As a common example, the layout of Figure 1 is used while showing the activity sets in a graphical way. The departure and arrival times of the train services of this example are given in Table 1.

4.2.1.   Open track events and activities. The open track events and activities are linked to an open track section (k, ℓ), i.e. an open track section between stations k and ℓ. In particular, the class of *open track activities* $a \in A_{\text{track}}$ 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 define $E_{\text{dep}}^k \subset E_{\text{train}}$ as the set of departure events form station k. As a subset of these events we define $E_{\text{dep}}^{k\ell} \subset E_{\text{dep}}^k$ as the set of departure events from station k of train services which visit station l directly after station k. Similarly, we define $E_{\text{arr}}^\ell \subset E_{\text{train}}$ as the set of arrival events at station ℓ. As a subset of these events we define $E_{\text{arr}}^{k\ell} \subset E_{\text{arr}}^\ell$ as the set of arrival events at station ℓ of train services which visit station k directly before station ℓ. Furthermore, let $e_{\text{track}}^{k\ell} \in E_{\text{inv}}$ denote the initial open track section inventory event for the number of tracks in open track section (k, ℓ).

In the following, we describe the activities of the sets A_{track}^- and A_{track}^+ related to station k. At station ℓ the activities are constructed similarly.
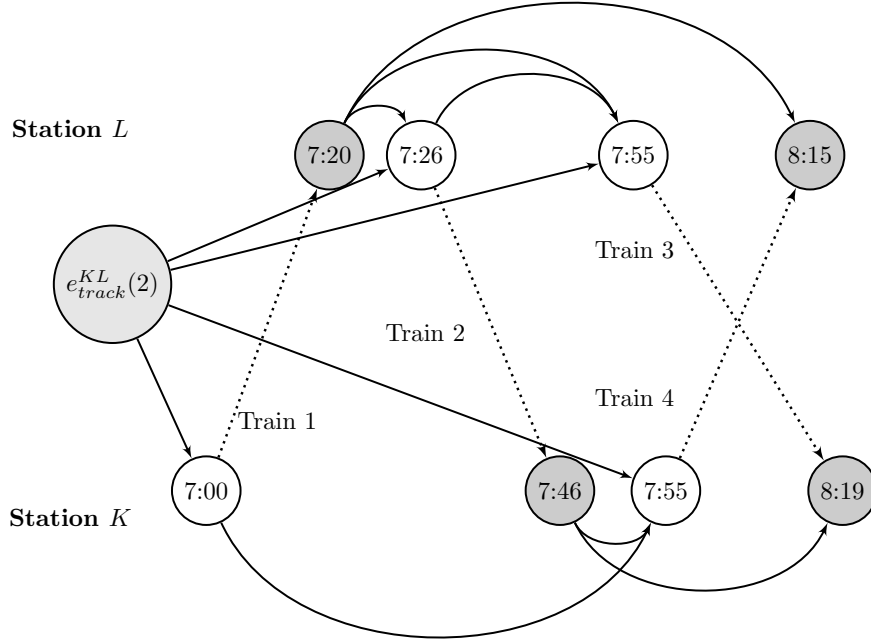
Figure 2   **Example of Event-Activity graph for open track section capacity**

- For a departure event $e \in E_{\mathrm{dep}}^{k\ell}$ from station k onto the open track section (k, ℓ):
  - The set $A_{\mathrm{track}}^-(e)$ contains open track section activities from each event $f \in E_{\mathrm{dep}}^{k\ell} \backslash \{e\}$ to e, from each event $f \in E_{\mathrm{arr}}^{\ell k}$ to e, and from the inventory event $e_{\mathrm{track}}^{k\ell}$ to e.
  - The set $A_{\mathrm{track}}^+(e)$ contains open track section activities to each event $f \in E_{\mathrm{dep}}^{k\ell} \backslash \{e\}$.
- For an arrival event $e \in E_{\mathrm{arr}}^{\ell k}$ at station k from the open track section (ℓ, k):
  - The set $A_{\mathrm{track}}^-(e)$ is empty since there is no capacity restriction for the arrival event, as at the moment of arrival the train is already running on the track.
  - The set $A_{\mathrm{track}}^+(e)$ contains open track section activities to each event $f \in E_{\mathrm{arr}}^{\ell k} \backslash \{e\}$, and to each event $f \in E_{\mathrm{dep}}^{k\ell}$.
- The set $A_{\mathrm{track}}^+(e_{\mathrm{track}}^{k\ell})$ contains open track section activities from the inventory event $e_{\mathrm{track}}^{k\ell}$ to each event $e \in E_{\mathrm{dep}}^{k\ell} \cup E_{\mathrm{dep}}^{\ell k}$.

Figure 2 shows an example of the resulting graph of the open track section activities corresponding to the graphical layout of Figure 1 and the timetable as in Table 1. Here we consider the open track between stations K and L. The departure and arrival events are represented by nodes and placed in a time-space plot where the vertical direction represents space 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 arrival nodes are indicated in gray and the departure nodes are white. The time at which an event takes places is given within the node. Furthermore the inventory event e_{track}^{KL} (which does not have a time) of the open track section between stations K and L is also indicated by a node. The capacity of the inventory event is given between brackets,

which is 2 since there are two tracks between stations K and L. 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 arcs. This means that two nodes linked with a dotted arc represent events of the same train service.

4.2.2. Station events and activities.

The station events and activities are linked to a station k. In particular, a *station activity* $a \in A_{\text{stat}}$ represents the sequential use of the same track in a station or the very first use of a track in a station.

To construct the station activity sets we need the sets E_{dep}^k and E_{arr}^k as defined in Section 4.2.1. Furthermore, we define $e_{\text{stat}}^k \in E_{\text{inv}}$ as the track inventory event at station k.

- For an arrival event $e \in E_{\text{arr}}^k$ at station k:
  — The set $A_{\text{stat}}^-(e)$ contains station activities from each event $f \in E_{\text{dep}}^k$ to e, and from the inventory event e_{stat}^k to e.
  — The set $A_{\text{stat}}^+(e)$ is empty. It is sure that the next event using the same track is the departure event of train service t_e from station k since one train at the time is allowed on one track.
- For a departure event $e \in E_{\text{dep}}^k$ from station k:
  — The set $A_{\text{stat}}^-(e)$ is empty for the same reason as above. It is sure that the previous event using the same track is the arrival event of train service t_e at station k.
  — The set $A_{\text{stat}}^+(e)$ contains station activities to each event $f \in E_{\text{arr}}^k$.
- The set $A_{\text{stat}}^+(e_{\text{stat}}^k)$ contains station activities from the inventory event e_{stat}^k to each event $e \in E_{\text{arr}}^k$.
- Some events $e \in E_{\text{dep}}^k$ represent a departure from station k which is the first departure of a train service. For these events it is not clear if the rolling stock for the train service arrives from the shunting yard or whether the train service 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 service, we construct the set $A_{\text{stat}}^-(e)$ in the same way as if e is an arrival event and the set $A_{\text{stat}}^+(e)$ in the same way as if event e is a departure event. In this way it is imposed that train service t_e needs to have a track available at its departure time.
- Similarly, some events $e \in E_{\text{arr}}^k$ represent an arrival at station k which is the last arrival of a train service. For these events it is not clear whether the rolling stock goes to the shunting yard or whether it stays at the station track to be used by another train service. For an arrival event e which represents a last arrival of a train service, we construct the set $A_{\text{stat}}^-(e)$ in the same way as if event e is an arrival and the set $A_{\text{stat}}^+(e)$ in the same way as if event e is a departure. In this way it is imposed that train service t_e needs an unoccupied track at its arrival time, and that the track is released for the arrival of another train directly (minimum L_a minutes) afterwards.
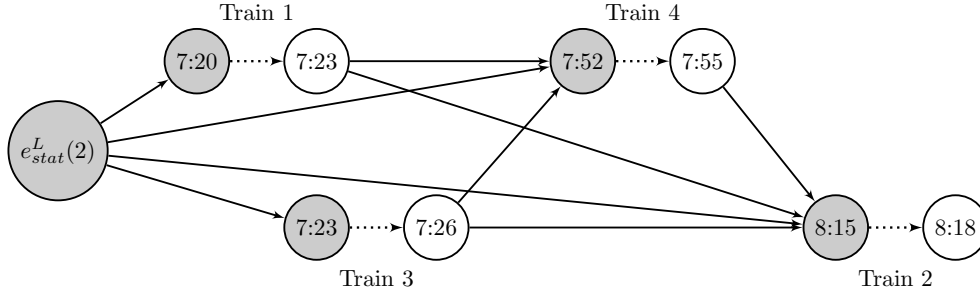
Figure 3 **Example of Event-Activity graph for station capacity**

Figure 3 shows an example of the resulting graph of the station activities, corresponding to the graphical layout of Figure 1 and the timetable as in Table 1. Here we consider station L. The meaning of the arcs and departure and arrival nodes is the same as in Figure 2. We have ordered the departure and arrival events based on time from left to right. The placing amongst the vertical axis does not have a meaning, as all events take place within the same station. The inventory event e^L_{stat} now represents the number of tracks within the station. Since there are 2 tracks, the inventory of 2 is given between brackets. To make the graph more intuitive, the train (in this case dwell) activities are also included and represented by dotted arcs.

4.2.3.   Rolling stock events and activities. The rolling stock events and activities are linked to the rolling stock compositions. In particular, a *rolling stock* activity $a \in A_{\text{rol}}$ represents the transfer of a rolling stock composition from an ending train service to a starting train service, or the very first use of a rolling stock composition.

We define $E^k_{\text{start}} \subset E_{\text{train}}$ as the set of departure events corresponding to the start of a train service from station k, and $E^k_{\text{end}} \subset E_{\text{train}}$ as the set of arrival events corresponding to the end of a train service at station k. Furthermore, we define $e^k_{\text{rol}} \in E_{\text{inv}}$ as the initial rolling stock inventory event at station k.

- For an event $e \in E^k_{\text{start}}$, denoting a start of a train service from station k:
  — The set $A^-_{\text{rol}}(e)$ contains rolling stock activities from each event $f \in E^k_{\text{end}}$ to event e if events e and f are using the same type of rolling stock, and from the inventory event e^k_{rol} to event e.
  — The set $A^+_{\text{rol}}(e)$ is empty since the next event using the same rolling stock is known (the arrival event of train service t_e at the next station).
- For an event $e \in E^k_{\text{end}}$, denoting an end of a train service at station k:
  — The set $A^-_{\text{rol}}(e)$ is empty since the previous event using the same rolling stock is known (the departure event of train service t_e from the previous station).
  — The set $A^+_{\text{rol}}(e)$ contains rolling stock activities to each event $f \in E^k_{\text{start}}$.
- The set $A^+_{\text{rol}}(e^k_{\text{rol}})$ contains rolling stock activities from the inventory event e^k_{rol} to each event $f \in E^k_{\text{start}}$.
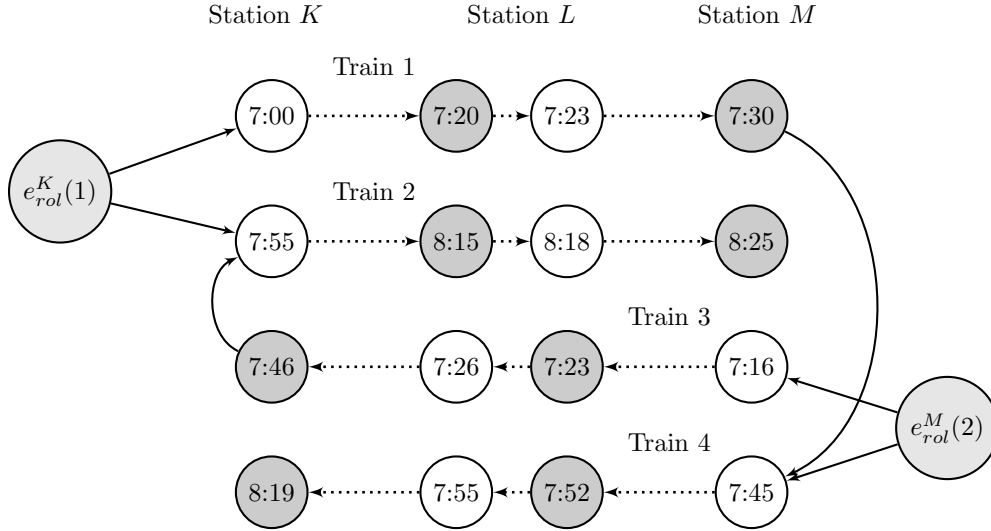
Figure 4 Example of Event-Activity graph for rolling stock capacity

- For arrival and departure events which do not represent the start of a train service nor the end of a train service, no rolling stock activities are constructed. Because we do not allow a train service to switch rolling stock during the complete service, all events of a train service use the same rolling stock.

Figure 4 shows an example of the resulting graph of rolling stock activities, corresponding to the graphical layout of Figure 1 and the timetable as in Table 1. Here we consider the rolling stock available at the shunting yard of stations K and M. The meaning of the arcs and departure and arrival nodes are the same as in Figure 2. We have ordered the departure and arrival events based on the location from left to right. The capacity of the inventory events, representing the available rolling stock at the shunting yard, is given between brackets. Again, to make the graph more intuitive, the train activities are included and represented by dotted arcs.

5. ILP formulation

The Timetable Rescheduling Problem consists of delaying some train services and cancelling some other train services such that maximum delay and capacity constraints are satisfied while minimizing the deviation from the original timetable. 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 service $t \in T$ is cancelled, and $y_t = 0$ otherwise. Moreover, z_a is a binary decision variable such that $z_a = 1$ if activity $a \in A$ is selected, and 0 otherwise. The main part of the proposed ILP formulation is given in (1)-(11), presented in its general form, in which C corresponds to one of the activity sets introduced in the previous sections for the different resources. A summary of the notation used in this model is given in Table 2. Additional constraints are described in the subsequent sections.

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

$$\text{subject to} \qquad\qquad x_e - q_e \geq 0 \qquad\qquad\qquad \forall e \in E_{\text{train}} \tag{2}$$

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

$$\sum_{a \in C} z_a + y_{t_e} = 1 \qquad\qquad \forall e \in E_{\text{train}}, C \in \mathcal{A}^-(e), C \neq \emptyset \tag{4}$$

$$\sum_{a \in C} z_a + y_{t_e} \leq 1 \qquad\qquad \forall e \in E_{\text{train}}, C \in \mathcal{A}^+(e) \tag{5}$$

$$\sum_{a \in C} z_a \leq i_e \qquad\qquad \forall e \in E_{\text{inv}}, C \in \mathcal{A}^+(e) \tag{6}$$

$$x_f - x_e + M(1 - z_a) \geq L_a \qquad\qquad \forall a = (e, f) \in A \backslash A_{\text{train}} \tag{7}$$

$$x_f - x_e \geq L_a \qquad\qquad \forall a = (e, f) \in A_{\text{train}} \tag{8}$$

$$y_t \in \{0, 1\} \qquad\qquad \forall t \in T \tag{9}$$

$$x_e \in \mathbb{N} \qquad\qquad \forall e \in E_{\text{train}} \tag{10}$$

$$z_a \in \{0, 1\} \qquad\qquad \forall a \in A \tag{11}$$

T	set of train services
y_t	whether train service $t \in T$ is cancelled (1 indicates yes)
λ_t	penalty for the cancellation of train service $t \in T$
E	set of events
E_{train}	set of train events
E_{inv}	set of inventory events
t_e	train corresponding to event $e \in E_{\text{train}}$
q_e	scheduled time of event $e \in E_{\text{train}}$
x_e	new time of event $e \in E_{\text{train}}$
d_e	maximum delay of event $e \in E_{\text{train}}$
μ_e	penalty per time unit for delaying event $e \in E_{\text{train}}$
i_e	number of resources made available by inventory event $e \in E_{\text{inv}}$
A	set of activities
z_a	whether activity $a \in A$ is selected (1 indicates yes)
L_a	minimum duration of activity $a \in A$
A_{train}	set of running/dwelling activities
$A_{\text{track}}^-(e)/A_{\text{track}}^+(e)$	track in/out-activities for event $e \in E$
$A_{\text{stat}}^-(e)/A_{\text{stat}}^+(e)$	station in/out-activities for event $e \in E$
$A_{\text{rol}}^-(e)/A_{\text{rol}}^+(e)$	rolling stock in/out-activities for event $e \in E$
$\mathcal{A}^-(e)$	the set $\{A_{\text{track}}^-(e), A_{\text{stat}}^-(e), A_{\text{rol}}^-(e)\}$
$\mathcal{A}^+(e)$	the set $\{A_{\text{track}}^+(e), A_{\text{stat}}^+(e), A_{\text{rol}}^+(e)\}$

Table 2   **Summary of notation used in Model (1)-(11).**

The objective function (1) is a weighted sum of the number of cancelled train services and the sum of the delays of all the events with respect to their original scheduled times. For every train

service t, λ_t describes the penalty for cancelling the train service (which, for example, can depend on the type of the train service or its 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 does not take place before its scheduled time in the original timetable and that its maximum allowed delay is not exceeded. Constraints (3) also ensure that if a train service is cancelled, then it is "virtually" operated at its original time (i.e. no delay penalty is considered).

The capacity constraints are given by constraint sets (4)-(7). Capacity constraints are needed in order to ensure that a resource unit is not occupied by more than one train service at a time, and that train services, which can not get all resources they require, are cancelled. The implications of the capacity constraints are different depending on the type of resource the set C in constraints (4)-(7) refers to. These implications are described in detail in Sections 5.1, 5.2 and 5.3.

The minimum running and dwell times of train services should also be respected. Although this can also be modelled by the capacity constraints, the order of the departure and arrival events is fixed for each train service. Therefore constraints (8) are sufficient to model this. Note that if the train service is cancelled, the corresponding constraint set (8) is satisfied due to constraint set (3).

Constraints (9)-(11) model the binary and integer character of the variables. Note that it is not strictly necessary to model x_e as an integer variable. As long as all paramters (L_a and q_e) are integer, the variable x_e will automatically be integer as well. One could also argue that arrival and departure times do not have to be integer at all, but we believe that from a passengers point of view this is still preferred. Some preliminary tests on the computation time required to solve our model with CPLEX 12.4 showed that sometimes the model with integer x_e variables is faster than the model with continuous x_e variables, and sometimes the opposite occurs. We decided to use the model with constraints (10).

5.1. 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 run at the same time on the same track only if they run in the same direction, but a minimum headway between each pair of trains must be considered.

According to the sets defined in Section 4.2.1, constraint set (4) implies that a train can only depart onto a track of an open track section, if a train has departed on the same track in the same direction, or if a track used in the opposite direction has become empty after an arrival, or if there is an empty track available from the inventory.

Constraint set (5) 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, and 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 the other direction by a departing train.

Constraint set (6) implies that at most a certain number of trains (equal to the number of tracks between the two stations) may depart onto a track which has not been used before.

For these classes of activities, constraint set (7) models the headway time which has to be taken into account between two consecutive train services operated on the same track.

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 $e, f \in E^k_{\text{dep}}$ and $e', f' \in E^\ell_{\text{arr}}$ such that $t_e = t_{e'}$ and $t_f = t_{f'}$, let $a = (e, f)$ be the activity representing the consecutive departures of train services t_e and t_f from station k on a track in (k, ℓ), and let $a' = (e', f')$ be the activity corresponding to the consecutive arrivals of the two train services 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.

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

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 5-8 show, with bold arcs for the selected activities, all feasible solutions for the open track section capacity problem considered in Figure 2 with fixed event times. Remember that the dashed arcs are train activities. These arcs are included so that one can easily see how events use the same resource unit. A selected path, in bold arcs, starting from the inventory event e^{KL}_{track} indicates in which order events take place on a single track in the open track section. For example, in Figure 5 each track is used in only one direction. Train services Train 1 and Train 4 are operated on one track, and train services Train 2 and Train 3 are operated on the other track.

Furthermore, Figures 5 and 8 show how constraints (12) work between train services Train 2 and Train 3. Due to constraint set (12) both the arc between the departures of these train services and the arc between the arrivals of these train services are used.
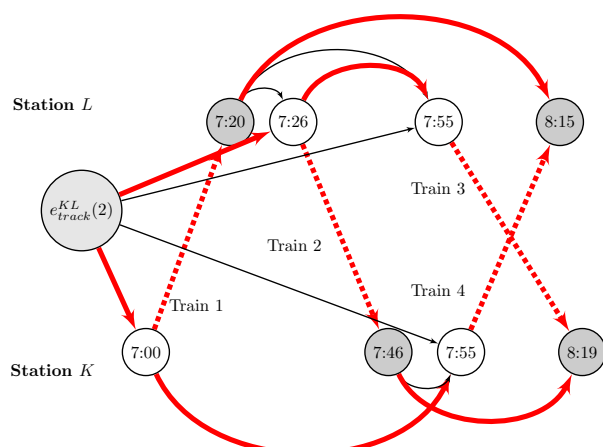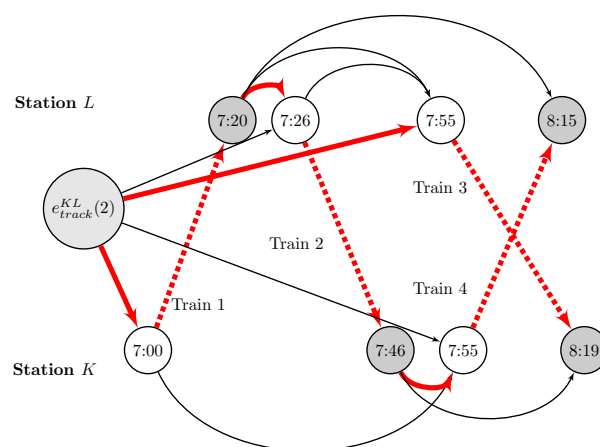
Figure 5   Solution 1
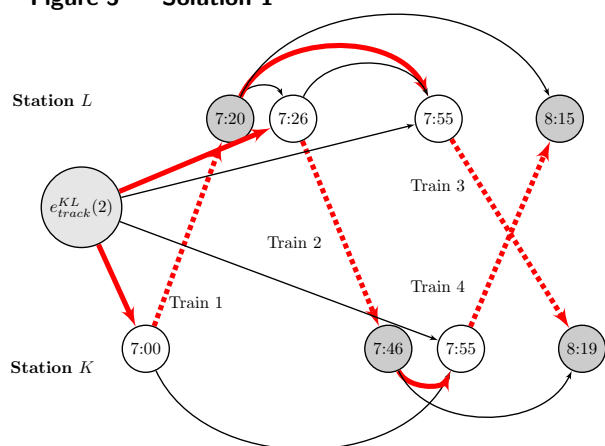

Figure 6   Solution 2

Figure 7   Solution 3
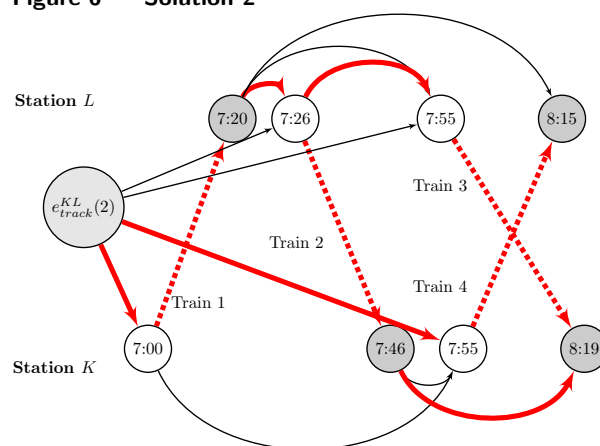
Figure 8   Solution 4

5.2.   Station capacities.

In a station a train service 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 on the same track at the same time. After a train has left the track in the station, another train can arrive on that track.

According to the sets defined in Section 4.2.2, constraint set (4) implies that a train can only arrive at a station 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.

Constraint set (5) implies that, if a train has departed from a station, then the track may be assigned to at most one new arrival.

Constraint set (6) implies that at most a certain number of trains (equal to the number of tracks at that station) may arrive at a track of the station which has not been used before.

For these classes of activities, constraint set (7) 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.

Figures 9-12 show, with bold arcs for the selected activities, all feasible solutions for the station capacity problem shown in Figure 3, given that the times the events take place are fixed. Remember

that the dashed arcs are train activities. A selected path, in bold arcs, 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 9, Train services Trains 1, 4 and 2 are scheduled after each other on one track and train service Train 3 on the other track.

Figure 9   Solution 1

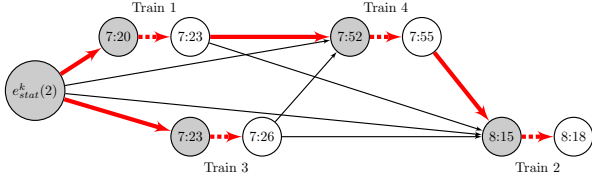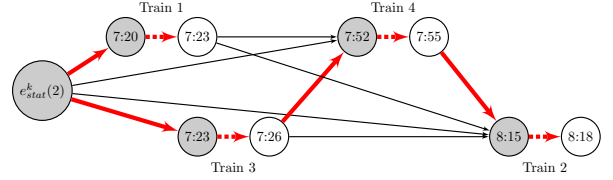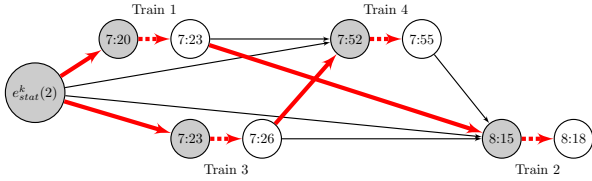**Figure 10   Solution 2**

Figure 11   Solution 3
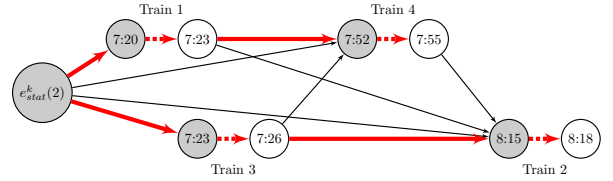
Figure 12   Solution 4

5.3.   Rolling stock capacities.

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

According to the sets defined in Section 4.2.3, constraint set (4) implies that a train service can only start if an earlier train service has ended at that station, or if there is still a rolling stock composition available in the inventory. Constraint set (5) implies that, if a train service has ended at a station, its rolling stock can be assigned to at most one other train service. Constraint set (6) implies that at most $i_{e_{rol}^k}$ train services may start from station k by using rolling stock compositions available at station k. Finally, constraint set (7) models the turnaround time required to transfer rolling stock from an ending train service to a starting train service, and the time required for shunting activities if a starting train service uses rolling stock from the inventory.

Figures 13 and 14 show with bold arcs the two feasible solutions for the rolling stock capacity problem shown in Figure 4. Remember again that the dashed arcs are train activities. A selected path from one of the inventory events (e_{rol}^K or e_{rol}^M) indicates in which order events use a rolling stock composition. For example, in Figure 13, train service Train 1 uses a rolling stock composition from the inventory at the shunting yard of Station K which will be left at the shunting yard of Station M. Train services Trains 3 and 4 both get their rolling stock composition from the inventory at
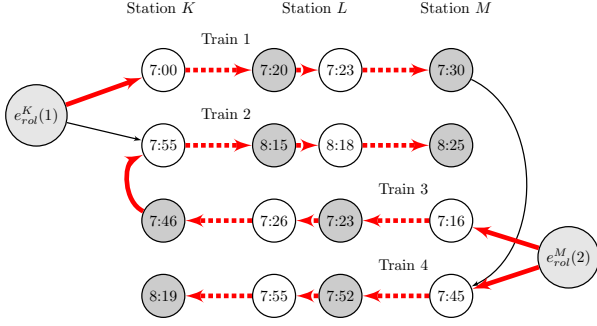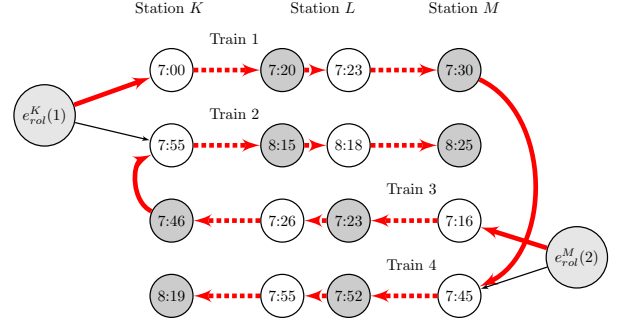
Figure 13 Solution 1

Figure 14 Solution 2

the shunting yard of Station M and train service Train 2 uses the same rolling stock composition as train service Train 3.

If a station has a shunting yard, then the rolling stock composition of an ending train service can be moved to the shunting yard before it is used by a starting train service again. To model this, we add the rolling stock activities which do not include a back-and-forth to the shunting yard to the subset of station activities. Only if the rolling stock composition goes to the shunting yard (with infinite capacity), 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 service has a rolling stock composition assigned to it.

Furthermore, during operations, it is preferred to have a regular turning pattern for the rolling stock compositions, since this requires less communication with the shunting crews. Each train service belongs to a train series. Train services belonging to the same train series traveling in the same direction belong to the same *subseries*. We assume a cyclic timetable, in which in every cycle, one train service of each train subseries is operated. If there is a certain train series with a higher frequency of train services per cycle, we split it into multiple train series of one train service per cycle per direction. 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 service belonging to s to a train service belonging to s'. A rolling stock activity a corresponding to the transfer of rolling stock between two train services therefore belongs to a turning pattern. This turning pattern is denoted by w_a. The set of all rolling stock activities is defined as A_{rol}. 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 binary decision variable such that $u_w = 1$ if only rolling stock activities following turning pattern w may be selected and 0 otherwise. Then constraint sets (13)-(15) ensure a regular turning pattern, in which rolling stock units of ending train services of a certain subseries are used only by starting train services which belong to the same subseries.

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

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

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

5.4. Blockage of an open track section

The disruptions considered in this paper consist of blockages of tracks with a known duration. A *partial blockage* is defined as the temporary unavailability of a subset of tracks between two stations. When all tracks between two stations are blocked, we call it a *full blockage*. Note that, from a modeling point of view, a full blockage is a special case of a partial blockage.

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 service with the time required 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 must be added. The time τ_1 at which the disruption starts and the time τ_2 at which the disruption ends are given as input, together with the 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 time τ_1. Therefore we also take into account what has happened before time τ_1, even though we cannot change it.

5.4.1. Disrupted train services. Train services running over the disrupted area need special attention. For each train service, 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 service $t \in T$ is scheduled. Let station k (ℓ) be the last (first) stopping station of train service t before (after) the blocked section. Then, if train service t has a scheduled departure from station k during the disruption, the events that were associated with train service t (in case of no disruption) are now partitioned and associated with three new partial train services, namely train services α_t, β_t and γ_t. Note that in the case of multiple blockages, if a train service traverses them, then it should be split into more than three services. Let T^{disr} be the set of all train services which have a scheduled departure during the disruption from their last stopping station before the disruption.

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

Constraint sets (16) and (17) are included to ensure that if train service β_t (the part of the train service over the disrupted area) is operated, then both train services α_t and γ_t are also operated. This is equivalent to operating the original train service t. If β_t is cancelled, however, then α_t and γ_t may run or may be cancelled, independently of each other. This is modelled as follows:

$$y_{\beta_t} \geq y_{\alpha_t} \qquad\qquad \forall t \in T^{disr} \qquad\qquad (16)$$

$$y_{\beta_t} \geq y_{\gamma_t} \qquad\qquad \forall t \in T^{disr} \qquad\qquad (17)$$

5.4.2. Additional rolling stock activities. For train services α_t and γ_t, all activities are defined as discussed in Section 4.2. However, in order to ensure that train services α_t, β_t and γ_t use the same rolling stock if they all run, a rolling stock activity (with duration 0) is defined from event e_k^{arr} to event e_k^{dep} and from event e_ℓ^{arr} to event e_ℓ^{dep}. Furthermore, no other rolling stock activities are defined for train service β_t. For event e_ℓ^{arr}, constraint (5) becomes an equality constraint. Hence in the case where none of these three train services are cancelled, they all use the same rolling stock, whereas if train service β_t is cancelled, the rolling stock units of train services α_t and γ_t may turn on other train services at stations k and ℓ, respectively.

5.4.3. Additional open track section activities. If a train service t has a departure time from station k before the disruption, but the disruption starts before the arrival time of the train service at station ℓ, then it is assumed that the train service continues as scheduled along its original route. The reason for this assumption is that decisions to be made for such a specific train service 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 service whose rolling stock is malfunctioning. In these cases, the events associated with train service t are the same as in the case of no disruption. However, the open track section activities of the arrival event of train service t at the end of the disrupted track section are only defined to events which take place after the end of the disruption (after τ_2). This ensures that the blocked tracks are only used again after the disruption has ended.

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

5.4.4. Balancing directions. If the number of tracks is reduced, then more train services can be operated 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 ILP model. Let SS be the set of pairs of subseries (s, s') for which s and s' belong to the same train series, but differ in direction.

$$\sum_{t \in s} y_t \leq \sum_{t \in s'} y_t + 1 \qquad\qquad \forall (s, s') \in SS \qquad (18)$$

$$\sum_{t \in s'} y_t \leq \sum_{t \in s} y_t + 1 \qquad\qquad \forall (s, s') \in SS \qquad (19)$$

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

5.5. 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 train services 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 entire day. Therefore it is assumed that, at some point in time after the disruption has ended, the train services should be operated again according to their original timetable. 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 service 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. If these events are not considered the model assumes that at τ_1 all tracks are empty and directly available. Furthermore, if these events are not considered, the model does not consider that there must be tracks available for the train services starting after τ_3.

In addition, 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, scheduled at the time the disruption is over and connected with the rolling stock activities, is added. The number of selected rolling stock activities to this event must be equal to the number of rolling stock compositions that are normally (without disruption) at that station at that time.

5.6.   Reroutings

One of the contributions of this paper is that we include the possibility for train services 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 service is that passengers wishing to use this train service do not have to reroute themselves (possibly with some transfers). Furthermore, passengers may experience a smaller amount of delay, and the original train services operated on the rerouted area may be less crowded.

To incorporate this functionality, a certain set of train services T^{Rer} is allowed to be rerouted. For each train service $t \in T^{Rer}$, which is scheduled to travel through a disrupted area, an alternative list of stations can be provided as rerouting option. Then, in addition to the new train services α_t, β_t and γ_t (compensating for the blocked area), a fourth train service δ_t is defined which runs on the rerouted path.

In the remainder of this section we assume that the alternative route of train service δ_t starts at k, the start location of either α_t or β_t, and that the alternative route of δ_t is constructed such that it ends at ℓ, the end location of either β_t and γ_t. Let T_t^{Alt} be the set of partial train services for which δ_t is the alternative. The three feasible options for this set are $\{\alpha_t, \beta_t\}$, $\{\beta_t, \gamma_t\}$ and $\{\alpha_t, \beta_t, \gamma_t\}$.

For each station on the rerouted path (apart from stations k and ℓ), an arrival event and a departure event are associated with train service δ_t, while at station k a departure event representing the start of the train service is associated with δ_t, and at station ℓ an arrival event representing the end of the train service is associated with δ_t.

For train service δ_t the scheduled time of the departure event from station k is the same as for the original train service α_t or β_t, while the arrival and departure times for the other events of δ_t 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 train services β_t and δ_t runs, constraint set (20) is added.

$$y_{\beta_t} + y_{\delta_t} \geq 1 \qquad\qquad \forall t \in T^{Rer} \qquad\qquad (20)$$

To ensure that, if train service $t \in T^{Rer}$ is rerouted, the complete train service runs, constraint sets (21)-(22) should be added.

$$y_{\delta_t} \geq y_{\alpha_t} \qquad\qquad \forall t \in T^{Rer} : \alpha_t \notin T_t^{Alt} \qquad\qquad (21)$$

$$y_{\delta_t} \geq y_{\gamma_t} \qquad\qquad \forall t \in T^{Rer} : \gamma_t \notin T_t^{Alt} \qquad\qquad (22)$$

Furthermore, to handle a penalty for rerouting, a variable p_η is introduced for all train services η which are (part) of the original route of a rerouting option δ_t. The binary decision variable p_η is

equal to 1 if the train service is cancelled and not rerouted, and p_η is equal to 0 if the train service is operated (on the planned or alternative route). To model this, constraint set (23) is added.

$$y_{\delta_t} + y_\eta \leq 1 + p_\eta \qquad\qquad \forall t \in T^{Rer}, \forall \eta \in T_t^{Alt} \qquad (23)$$

In the objective function, for all train services $\eta \in T_t^{Alt}$ of each train service $t \in T^{Rer}$, $\lambda_\eta y_\eta$ is replaced by $\theta \cdot \lambda_\eta y_\eta + (1 - \theta) \cdot \lambda_\eta p_\eta$ where θ is a parameter between 0 and 1 indicating the balance between the cost of rerouting and the cost of cancelling a train service. Furthermore $\lambda_{\delta_t} = 0$ for each train service $t \in T^{Rer}$. For example if $T_t^{Alt} = \{\beta_t\}$, these settings assure that if train service β_t is operated as planned ($y_{\beta_t} = 0$, $y_{\delta_t} = 1$ and $p_{\beta_t} = 0$), the costs in the objective function are 0. If train service β_t is rerouted ($y_{\beta_t} = 1$, $y_{\delta_t} = 0$ and $p_{\beta_t} = 0$), the costs are $\theta \cdot \lambda_{\beta_t}$ and if train service β_t is cancelled and not rerouted ($y_{\beta_t} = 1$, $y_{\delta_t} = 1$ and $p_{\beta_t} = 1$), the costs are λ_{β_t}. Note that the variable p_{β_t} is only forced to take the correct value if penalty λ_{β_t} has a positive value. This will always be the case in our application, since cancelling a train service will be penalized and not rewarded.

6. Computational experiments

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

6.1. Case description

For our computational tests we consider a heavily used part of the Dutch railway network which is indicated in Figure 15. 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
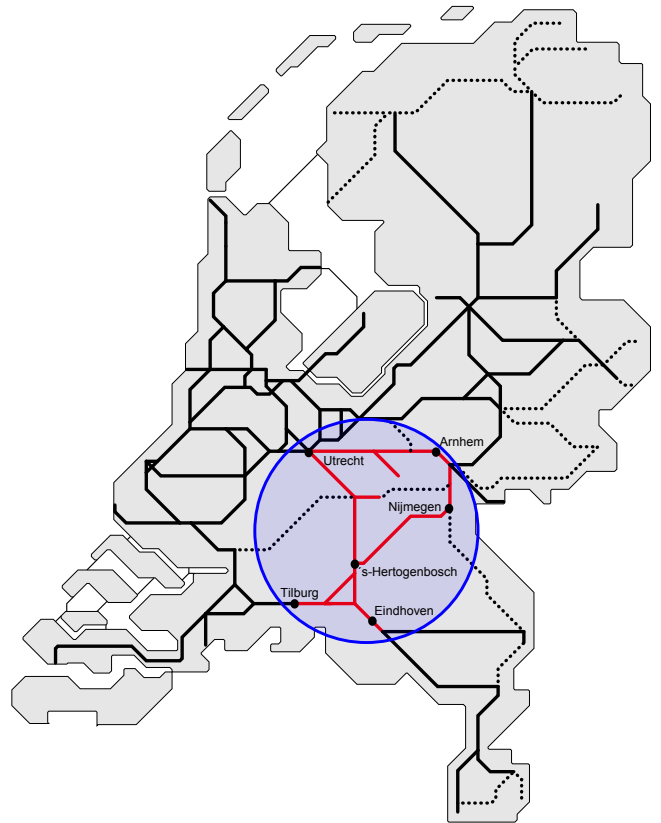

Figure 15: Overview of the Dutch railway network

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 train services 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 are operated (mostly twice per 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 operate all these train services. This means that any spare rolling stock compositions at the shunting yards are not considered. In total 61 rolling stock compositions are necessary to operate the timetable of the complete day.

6.2. 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 to the network described in Section 6.1.

For the maximum allowed delay d_e of a train event e, we make a distinction between three types of events. Train services started before the start of the disruption may be delayed more than a train service which has not started yet. This prevents infeasibilities of the mathematical model, since started train services are not allowed to be cancelled. For events $e \in E_{\text{train}}$ of train services which have been started at the time the disruption starts we have $d_e = 30$ minutes, while for events $e \in E_{\text{train}}$ of rerouted train services 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 L_a of a train service 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 service can be used for a new starting train service is 5 minutes. If the activity between an ending and a starting train service takes longer than 10 minutes at a station with a shunting yard, the rolling stock composition

Considered time after disruption in minutes	10	15	20	30	40	45, 50, 60, 70, 75, 80 & 90
# infeasible instances	46	36	34	24	7	2

Table 3   **Number of infeasible cases resulting from different settings of the considered time after the end of the disruption.**

goes to the shunting yard and releases the station capacity 5 minutes after the train service has ended and requires station capacity from 5 minutes before its next train service starts.

In the objective function we have two penalties. Penalties for delaying train services and penalties for cancelling train services. The aim is to operate as many train services as possible. Therefore cancelling a train service is penalized much more than delaying a train service. Cancelling a train service is penalized by 50 times the running time of the train service. Furthermore for every event there is a penalty of 1 per delayed minute.

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.

6.3.   Disruption scenarios

In order to test the described approach, a large set of disruption scenarios is created. For each of the 27 open track sections we constructed 30 instances 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 instances is constructed where only one track is blocked. For each open track section, the first instance is a 2 hour disruption of that open track section starting at 9:00. Then we increase in every new instance the start time of the 2 hour disruption by one minute. Since the timetable of Netherlands Railways in this region is cyclic with a cycle time of 30 minutes, a disruption starting at 9:30 should be very similar to a disruption starting at 9:00. In total this leads to 810 instances of full blockages and 720 instances where only one track is blocked.

We did some tests on the full blockage instances to determine the value of τ_3 (see Section 5.5), i.e. how much time should be taken into account after the end of the disruption before all train services should be operated again as planned. In Table 3 the results of these tests are provided. For each different value of the considered time after the end of the disruption, the number of the 810 instances for which the model turned out to be infeasible is given. As expected, the more time is available after the end of the disruption the less instances turn out to be infeasible with our model.

In addition to the reduced number of infeasible cases, a larger time available after the disruption also slightly reduces the number of cancelled train services. However this reduction seems not to be significant. For example, in case we allow a maximum delay of 10 minutes for train services

which have not started at the start of the disruption, considering 60 or 90 minutes after the end of the disruption instead of 45 minutes, decreases the total scheduled length of the cancelled train services on average (over the 810 full blockage instances) by 1.1 or 1.3 minutes, respectively. This is a reduction of less than 0.28% or 0.35%, respectively.

These tests induced us to use a time period of 45 minutes after the end of the disruption before the timetable should be able to be operated as originally scheduled. This means that, with a 2 hour blockage, all train services operated in a 2 hour and 45 minute period are taken into consideration in the timetable rescheduling.

6.4. Rerouting of train services

We have tested two different rerouting options. In the first one, if there is a disruption between 's-Hertogenbosch and Eindhoven, then train services can be rerouted via Tilburg (see Figure 15). For this case, 30 instances where all tracks are blocked and 30 instances where one of the two tracks is blocked are constructed in a similar way as described in Section 6.3. In these instances the train services 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 train services can be rerouted. This Intercity line has no scheduled stop between 's-Hertogenbosch and Eindhoven and therefore the rerouting does not skip stops.

If a train service is rerouted, the costs are 20% of the costs for cancelling that train service. This means that rerouting is preferred over cancelling the train service, but also that the original route is preferred over the alternative route.

The second rerouting option considers the train services between Nijmegen and Utrecht when there is a disruption at one of the 6 open track sections between Nijmegen, Arnhem and Utrecht. In the original timetable, all train services between Nijmegen and Utrecht run via Arnhem, but due to the disruption we allow one of the Intercity lines to be rerouted via 's-Hertogenbosch. Note that, due to this rerouting, many stops between Nijmegen and Utrecht are skipped and therefore this rerouting has also a large influence on the capacity between intermediate stops in between Nijmegen and Utrecht. We wish to mention that more information about passenger demand would be helpful to decide whether this rerouting is useful in practice, but this goes over the scope of this paper.

For each of the 6 open track sections, we construct 30 full blockage instances as described in Section 6.3. We believe that, in case only one track is blocked, rerouting is not fair given the previously discussed skipped stops due to the rerouting.

In this option, since the original route is quite long (almost 1 hour), if a train service is rerouted the costs are only 5% of the costs for cancelling that train service. The cost for rerouting should

not be too high, since otherwise cancellation of (other) short train services would be preferred over rerouting. With the previous choice only cancelling a train service shorter than 3 minutes is preferred over rerouting.

These two rerouting options together result in 210 rerouting instances in which all tracks are blocked, and 30 rerouting instances in which only one track is blocked.

To include the rerouted train services 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 in which 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.

6.5. General results

In this section we discuss the results of all the 1530 instances without the option to reroute train services, and in Section 6.6 we describe the results of the 240 instances, as described in Section 6.4, in which we included the option of rerouting train services.

Tables 4 and 5 contain the minimum, average and maximum values of the performance measures over all feasible instances for the full blockages and the blockages of one track. Note that the minimum (maximum) value represents the lowest (highest) value found for that measure. This means that the minimum (maximum) values for the different measures do not necessarily originate from the same instance.

There are two settings: (i) whether or not to include balancing constraints ((18)-(19)) in the case of a partial blockage, and (ii) how much delay (0, 3, 5 or 10 minutes) is allowed for train services that did not start before the start of the disruption. From now on, we refer with the term *allowed delays* only to the delays allowed for train services that did not start before the start of the disruption.

It turned out that for only 2 of the 1530 instances it was not possible, as discussed earlier, to find a feasible solution matching all constraints in our model.

Table 4 gives the results on the number of cancelled train services, the number of partially cancelled train services, the total cancellation time (expressed in minutes) and the computation times (expressed in seconds). In order to make the results maximally comparable, we count the number of the cancelled and partially cancelled train services, and the total cancellation time as follows. A train service t is partially cancelled if the disruption was on its route, and if only one of the train services α_t or γ_t (the parts before and after the disruption) is operated, while the other one is cancelled. If only the part on the disruption (train service β_t) is cancelled, this is not considered as a (partially) cancelled train service, since this part has to be cancelled inevitably

	Minutes of allowed delay	Totally cancelled train services			Partially cancelled train services			Total cancellation time (m)			Computation time (s)		
		Min	Avg	Max	Min	Avg	Max	Min	Avg	Max	Min	Avg	Max
Full	*0*	0	2.0	10	0	4.6	18	28	405	947	2	2.9	8
blockage	*3*	0	1.6	10	0	3.4	15	28	389	908	3	4.2	8
(808 instances)	*5*	0	1.5	11	0	3.4	16	28	382	905	3	4.9	10
	10	0	1.3	9	0	3.5	17	28	373	881	4	11.8	89
One track blocked	*0*	0	1.0	6	0	2.1	10	0	156	465	2	3.8	12
no balancing	*3*	0	0.6	6	0	1.6	8	0	115	348	3	5.8	18
(720 instances)	*5*	0	0.6	5	0	1.0	7	0	100	335	4	7.8	41
	10	0	0.5	4	0	0.7	6	0	69	286	5	36.8	433
One track blocked	*0*	0	0.9	6	0	1.9	11	0	166	465	2	4.0	10
with balancing	*3*	0	0.7	6	0	1.3	8	0	127	420	3	6.5	39
(720 instances)	*5*	0	0.8	5	0	1.0	6	0	108	420	4	8.5	54
	10	0	0.6	4	0	0.6	6	0	75	300	5	37.5	2042

Table 4   **Results on cancellations and computation time**

in case of a full blockage. The total cancellation time is the sum of the scheduled durations in minutes of the cancelled parts of all cancelled, partially cancelled, and inevitably cancelled train services, which is the main part in the objective function. Since the aim is to minimize the total cancellation time and not to minimize the number of cancelled train services, it may happen that there are solutions where more train services (with a short duration) need to be cancelled to have a smaller cancellation time.

For each instance, the total number of train services is approximately 165, since we compute a schedule for 2 hours and 45 minutes in which more than 60 train services per hour run.

In case of a full blockage, as shown in Table 4, if no delays are allowed, on average an amount of 405 minutes (almost 7 hours) of train service is cancelled. By allowing up to 10 minutes of delay, the total cancellation time decreases by approximately 8%.

If only one track is blocked, we see that the effect of allowing more delay is much higher. By allowing up to 10 minutes of delay, the total cancellation time decreases on average by more than 50% (from 156 to 69 minutes). The results also show that, in some parts of the network, it is still possible to operate all train services if only one track is out of service. This may be deduced from the fact that there are solutions with no cancellation time.

The price of forcing the new schedule to be a regular one by imposing balancing constraints, in which the number of trains services per train series in one direction differs at most one from the number of train services in the other direction, gives an increase of around 10% in the total cancellation time.

	Minutes of allowed delay	Delayed train services			Delayed events			Total maximum delay (m)			Total delay (m)		
		Min	Avg	Max	Min	Avg	Max	Min	Avg	Max	Min	Avg	Max
Full	*0*	0	0.1	2	0	0.2	15	0	1	35	0	2	226
blockage	*3*	0	1.4	18	0	6	76	0	3	44	0	10	157
(808 instances)	*5*	0	2.0	18	0	9	76	0	5	48	0	21	187
	10	0	2.7	22	0	12	122	0	11	99	0	47	780
One track blocked	*0*	0	0.1	2	0	0.7	19	0	1	27	0	10	387
no balancing	*3*	0	3.4	20	0	19	100	0	7	38	0	38	266
(720 instances)	*5*	0	5.2	25	0	32	168	0	13	83	0	81	720
	10	0	9.0	30	0	52	215	0	38	159	0	217	1105
One track blocked	*0*	0	0.1	2	0	0.7	17	0	1	27	0	8	387
with balancing	*3*	0	3.0	20	0	17	91	0	6	35	0	37	368
(720 instances)	*5*	0	4.8	20	0	29	115	0	12	67	0	74	396
	10	0	8.1	27	0	48	179	0	33	147	0	195	980

Table 5   Results on delays experienced

The largest computation time is less than 1 minute when a maximum delay up to 5 minutes is allowed. This means that our approach is able to find efficient 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 full blockage case, the computation times remain low (below 1.5 minutes). In case of a single track blockage, by allowing up to 10 minutes of delay, the computation time can increase to values which are not usable in practice. However, on average, the computation time is still relatively low and in both variants (with and without balancing constraints) less than 4 cases (out of 720) 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, 3 of the 4 instances had a solution with the same total cancellation time. For the fourth instance the total cancellation time was 25% higher. For the 4 instances with balancing constraints, 2 instances had no increase in total cancellation time and for the other two cases the total cancellation times were 25% and 28% higher.

The decision on which allowed delay should be preferred can be taken by the dispatchers, also considering the computation time available for obtaining the new timetable. A good trade-off between solution quality and computation 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 this 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.

Table 5 gives the results on the number of delayed train services, the number of delayed events, the sum of the maximum delay experienced by each train service on its route (*Total maximum delay*), as well as the sum of the delays of all events. As expected, by allowing larger delays the average total delay of the train services and events increases, thereby, as can be seen from Table 4, the number of cancelled train services decreases.

In case only one track is blocked, the delays are slightly lower if we impose the constraints which enforce a balance between the number of train services in each direction. Most likely this happens since, in that case, more train services are cancelled, and so smaller delays are necessary.

6.6. Results with the option to reroute trains

6.6.1. Rerouting train services between 's-Hertogenbosch and Eindhoven. Table 6
presents the results of the 60 instances for which the rerouting option for certain Intercity trains has been considered in case of a blockage between 's-Hertogenbosch and Eindhoven.

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 train services which can be rerouted is equal to $2 \times 2 \times 2 = 8$.

In Table 6 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 full blockages, the proposed approach is able to reduce considerably the total cancellation time (approximately 28%) if we allow reroutings. Computation times stay below 1 minute. In addition, having the option to reroute train services reduces the total cancellation time much more than simply allowing larger maximum delays.

As can be seen from the average number of rerouted train services, in almost every instance all 8 train services which are allowed to be rerouted are rerouted. As three train series run over the disrupted area, and now one complete series is rerouted, the reduction of the total cancellation time by 28% matches the expected reduction of cancellation time by one third.

If all tracks are blocked, we originally did not have to include the balancing constraints since, in that case, no train services at all run over the blocked area. However, in case we allow reroutings, it is worthwhile to include balancing constraints for the rerouted train services to ensure that train services are operated evenly in each direction. The results show that, in the case of a full blockage with reroutings, these balancing constraints have no influence on the solution and on the computation time. This is due to the fact that it is possible to reroute all trains in many cases.

Also in case only one of the tracks is blocked, allowing reroutings can considerably reduce the total cancellation time (up to 49%). In case of rerouting and a blockage of only one track, the

	Minutes of allowed delay	Rerouted train services			Total cancellation time (m)			Computation time (s)		
		Min	Avg	Max	Min	Avg	Max	Min	Avg	Max
Full blockage	0				481	520	544	3	3.0	3
no rerouting	3				469	511	536	4	4.2	7
(30 instances)	5				453	505	536	4	4.8	8
	10				451	487	525	6	8.6	12
Full blockage	0	6	7.6	8	328	374	410	5	5.6	7
with rerouting no balancing	3	6	7.6	8	316	366	410	8	9.3	14
(30 instances)	5	6	7.5	8	300	361	410	10	10.9	16
	10	6	7.3	8	300	347	390	15	24.8	55
Full blockage	0	6	7.6	8	328	374	410	5	5.8	6
with rerouting with balancing	3	6	7.5	8	316	366	410	8	9.4	13
(30 instances)	5	6	7.5	8	300	361	410	10	11.0	15
	10	6	7.3	8	300	347	390	16	24.0	60
One track blocked	0				240	273	304	3	5.0	9
no rerouting no balancing	3				185	219	251	4	6.1	15
(30 instances)	5				151	173	217	6	14.3	36
	10				60	92	115	20	52.7	260
One track blocked	0	3	6.3	8	142	172	206	6	11.6	20
with rerouting no balancing	3	3	3.9	4	105	141	172	10	15.7	49
(30 instances)	5	3	4.8	8	76	99	124	12	38.6	86
	10	0	1.9	4	57	81	96	41	115.2	226
One track blocked	0				288	309	334	5	6.4	8
no rerouting with balancing	3				243	280	326	6	13.5	39
(30 instances)	5				236	257	307	12	24.6	54
	10				94	134	172	25	96.9	235
One track blocked	0	6	7.4	8	137	170	201	9	13.4	20
with rerouting with balancing	3	3	3.9	4	105	142	172	12	21.2	44
(30 instances)	5	3	4.2	7	97	133	172	25	61.0	122
	10	3	3.9	4	58	83	114	55	183	436

Table 6   Results with reroutings 's-Hertogenbosch-Eindhoven

balancing constraints still have a small effect. Only if we allow delays of 5 minutes, adding the balancing constraints has a large increase (about 35%) in the total cancellation time. Without rerouting, there were much larger increases in total cancellation time by adding the balancing constraints.

One remark should be made about the instances in which the allowed delay was 0 minutes. In this case, adding the balancing constraints reduces the average total cancellation time. This is caused by the larger number of train services rerouted to get a more balanced timetable. Note that

	Minutes of allowed delay	Rerouted train services			Total cancellation time (m)			Computation time (s)		
		Min	Avg	Max	Min	Avg	Max	Min	Avg	Max
Full blockage	*0*				244	398	565	2	3.0	4
no rerouting	*3*				244	383	528	3	4.3	6
(180 instances)	*5*				244	377	524	4	4.9	7
	10				244	369	511	5	7.7	15
Full blockage	*0*	0	0.6	3	244	392	565	4	29.7	214
with rerouting no balancing	*3*	0	4.9	7	169	291	456	6	37.9	64
(180 instances)	*5*	0	4.8	7	169	285	443	7	48.0	101
	10	0	5.0	6	169	272	443	10	221	1224
Full blockage	*0*	0	0.3	1	244	395	565	4	27.5	145
with rerouting with balancing	*3*	0	4.9	7	169	291	456	6	37.9	64
(180 instances)	*5*	0	4.8	7	169	285	456	7	48.9	78
	10	0	5.0	6	169	272	443	10	198	1375

Table 7   Results with reroutings Nijmegen-Utrecht

in the objective function value a rerouted train service counts for 20% of cancelling the original service, which means, for the specific train services in these instances, that the cost of rerouting the service is equal to 4 minutes of cancellation time. If one wants to prevent this kind of behavior, rerouting should have a lower cost.

6.6.2.   Rerouting train services between Nijmegen and Utrecht. Table 7 presents the results of the 180 instances for which the rerouting option for certain Intercity trains has been considered in case of a full blockage at one of the 6 open track sections between Nijmegen, Arnhem and Utrecht.

Adding the option of rerouting train services between Nijmegen and Utrecht did not have much influence if the allowed delay is 0 minutes. This could have been expected since the detour is quite long, so it is difficult to fit the train service into the timetable if all the other train services are not allowed to be delayed. However, if delays are allowed, the total cancellation time is reduced considerably, by approximately 25%, because of the rerouted train services.

As in the case of rerouting between 's Hertogenbosch and Eindhoven, there is not much difference between the solutions with or without the constraints enforcing a balance between train services in each direction. To be more precise, only in case of an allowed delay of 0 minutes, adding the constraints for a balanced timetable leads to a (very small) increase of the total cancellation time.

Both the rerouting scenario between 's-Hertogenbosch and Eindhoven and the scenario between Nijmegen and Utrecht show that, also in case reroutings are allowed, the proposed approach is able to quickly find an optimal solution, as long as we do not allow more than 5 minutes of delay. If dispatchers want to allow larger delays (10 minutes), the computation times increase.

7.   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 rescheduling problem studied 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 train services, in order to reduce the number of cancelled and delayed train services, 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 service, and an activity refers to the sequential use of a resource unit by two events. The resources considered are the tracks in the open track sections, the tracks in the stations, and the rolling stock compositions at the stations.

Computational tests are performed on a heavily used part of the Dutch railway network. Solutions are obtained in reasonable computation times, which makes the approach suitable for practical usage. Most of the considered scenarios can be solved within 1.5 minutes of computation time using a state-of-the-art commercial ILP solver. The results show that a smaller number of train services needs to be cancelled and the total cancellation time is significantly reduced if a larger delay is allowed or if the rerouting option is considered.

The proposed approach is 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 train services. This makes our approach much more flexible and efficient than the current practice of using contingency plans.

Acknowledgments

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(1):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(1):69–86, 2011b.

B. Adenso-Dıaz, M. Oliva González, and P. González-Torre. On-line timetable re-scheduling in regional train services. *Transportation Research Part B: Methodological*, 33(6):387–398, 1999.

A. Albrecht, D. Panton, and D. Lee. Rescheduling rail networks with maintenance disruptions using Problem Space Search. *Computers & Operations Research*, 40(3):703–712, 2013.

M. Almodóvar and R. García-Ródenas. On-line reschedule optimization for passenger railways in case of emergencies. *Computers & Operations Research*, 40(3):725–736, 2013.

M. Boccia, C. Mannino, and I. Vasilyev. The dispatching problem on multitrack territories: Heuristic approaches based on mixed integer linear programming. *Networks*, 62(4):315–326, 2013.

P. Brucker, S. Heitmann, and S. Knust. Scheduling railway traffic at a construction site. *OR Spectrum*, 24(1):19–30, 2002.

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.

V. Cacchiani and P. Toth. Nominal and Robust Train Timetabling Problems. *European Journal of Operational Research*, 219(3):727–737, 2012.

G. Caimi, M. Fuchsberger, M. Laumanns, and M. Lüthi. A model predictive control approach for discrete-time rescheduling in complex central railway station areas. *Computers & Operations Research*, 39(11):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*, chapter 3, 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. Chu and A. Oetting. Modeling capacity consumption considering disruption program characteristics and the transition phase to steady operations during disruptions. *Journal of Rail Transport Planning & Management*, 3(3):54–67, 2013.

F. Corman, A. D'Ariano, and I. A. Hansen. Evaluating disturbance robustness of railway schedules. *Journal of Intelligent Transportation Systems*, 18(1):106–120, 2014a.

F. Corman, A. D'Ariano, I. A. Hansen, and D. Pacciarelli. Optimal multi-class rescheduling of railway traffic. *Journal of Rail Transport Planning & Management*, 1(1):14–24, 2011a.

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, Delft, The Netherlands, Apr 2011b.

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(6):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(1):175–192, 2010.

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(1):79–94, 2012.

F. Corman, A. D'Ariano, D. Pacciarelli, and M. Pranzo. Dispatching and coordination in multi-area railway traffic management. *Computers & Operations Research*, 44:146–160, 2014b.

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(4):405–419, 2008a.

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(2):643–657, 2007.

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(2):232–245, 2008b.

T. Dollevoet, D. Huisman, M. Schmidt, and A. Schöbel. Delay Management with Rerouting of Passengers. *Transportation Science*, 46(1):74–89, 2012.

J. Jespersen-Groth. *Decision Support for the Rolling Stock Dispatcher*. PhD thesis, DTU, Copenhagen, Denmark, 2008.

J. Jespersen-Groth, J. Clausen, and J. Larsen. Optimal reinsertion of cancelled train lines. Technical report, IMM-DTU, aug 2006. URL `http://www2.imm.dtu.dk/pubdb/p.php?5095`.

P. Kecman, F. Corman, A. D'Ariano, and R. Goverde. Rescheduling models for railway traffic management in large-scale networks. *Public Transport*, 5(1-2):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(3):583–593, 2014.

R. M. Lusby, J. Larsen, M. Ehrgott, and D. Ryan. Railway track allocation: models and methods. *OR Spectrum*, 33(4):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(3):713–724, 2013.

C. Mannino and A. Mascis. Optimal Real-Time Traffic Control in Metro Stations. *Operations Research*, 57(4):1026–1039, 2009.

L. Meng and X. Zhou. Simultaneous train rerouting and rescheduling on an N-track network: A model reformulation with network-based cumulative flow variables. *Transportation Research Part B: Methodological*, 67:208–234, 2014.

S. Narayanaswami and N. Rangaraj. Modelling disruptions and resolving conflicts optimally in a railway schedule. *Computers & Industrial Engineering*, 64(1):469–481, 2013.

K. Sato, K. Tamura, and N. Tomii. A mip-based timetable rescheduling formulation and algorithm minimizing further inconvenience to passengers. *Journal of Rail Transport Planning & Management*, 3(3):38–53, 2013.

A. Schöbel. Capacity constraints in delay management. *Public Transport*, 1(2):135–154, 2009.

J. Törnquist. Design of an effective algorithm for fast response to the re-scheduling of railway traffic during disturbances. *Transportation Research Part C: Emerging Technologies*, 20(1):62–78, 2012.

J. Törnquist and J. A. Persson. N-tracked railway traffic re-scheduling during disturbances. *Transportation Research Part B: Methodological*, 41(3):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.