



---

## Generation and evaluation of passenger-oriented railway disposition timetables

Stefan Binder

Jianghang Chen

Michel Bierlaire

Transport and Mobility Laboratory, EPFL

April 2014

**STRC**

**14th Swiss Transport Research Conference**

Monte Verità / Ascona, May 14-16, 2014

Transport and Mobility Laboratory, EPFL

## Generation and evaluation of passenger-oriented railway disposition timetables

Stefan Binder  
Transport and Mobility  
Laboratory  
Ecole Polytechnique Fédérale  
de Lausanne  
Station 18, 1015 - Lausanne  
phone: +41-21-693 93 29  
fax: +41-21-693 80 60  
s.binder@epfl.ch

Jianghang Chen  
Transport and Mobility  
Laboratory  
Ecole Polytechnique Fédérale  
de Lausanne  
Station 18, 1015 - Lausanne  
phone: +41-21-693 25 32  
fax: +41-21-693 80 60  
jianghang.chen@epfl.ch

Michel Bierlaire  
Transport and Mobility  
Laboratory  
Ecole Polytechnique Fédérale  
de Lausanne  
Station 18, 1015 - Lausanne  
phone: +41-21-693 25 37  
fax: +41-21-693 80 60  
michel.bierlaire@epfl.ch

April 2014

### Abstract

As delays are one of the main sources of passenger dissatisfaction in public transportation networks, railway companies put considerable effort in trying to avoid them. However, on a daily basis, delays occur for a number of reasons, e.g., a jammed door at a station or a temporarily unavailable track. Once an initial delay has occurred, the original timetable needs to be updated to a so-called disposition timetable. This new timetable has to be conflict-free in terms of operational constraints (e.g., trains cannot use the same track section at the same time) and as convenient as possible for the passengers. The problem of finding a conflict-free disposition timetable has received a huge attention in the scientific literature in recent years, both at a macroscopic and at a microscopic level. The literature focusing on minimizing the negative effects of delays for passengers has also been developing but is much sparser. We propose a model that describes the railway recovery problem in a way that takes passengers explicitly into account. Our model focuses mainly on severe disruptions (as opposed to minor disturbances that can be handled by slight modifications to the timetable) and proposes and evaluates several recovery strategies. The evaluation of the recovery strategies (e.g., partial train cancellation, complete train cancellation, train addition, train replacement) provided by the disposition timetable is based on a number of passenger satisfaction indicators, such as total travel time, number of connections and train saturation. This model will assist train operating companies when evaluating the trade-off between economic and infrastructural feasibility of recovery schemes on the one hand side and passenger satisfaction on the other.

### Keywords

Railway rescheduling, severe disruptions, recovery schemes, passenger satisfaction indicators, transit assignment model, timetable evaluation.

# 1 Introduction

As delays are one of the main sources of passenger dissatisfaction in public transportation networks, railway companies put considerable efforts in trying to avoid them. However, on a daily basis, delays occur for a number of reasons, e.g., a jammed door at a station or a temporarily unavailable track due to maintenance work. Once an initial delay has occurred, the original timetable needs to be updated to a so-called disposition timetable.

This new timetable has to be conflict-free in terms of operational constraints (e.g., trains cannot use the same track section at the same time) and as convenient as possible for the passengers. The current practice in the field still heavily relies on predetermined “what-if” scenarios and personal experience of the train traffic controllers. However, due to the high utilization rate of modern railway networks, a decision made at one location in the network can have domino effects on the whole network. Up to date, no comprehensive tool assessing all consequences of train controllers’ decisions exists to assist in their decisions, hence possibly leading to unforeseen additional conflicts as well as sub-optimality in the network.

The problem of finding a conflict-free disposition timetable has received a huge attention in the scientific literature in recent years, both at a macroscopic and at a microscopic level. The literature focusing on minimizing the negative effects of delays for passengers has also been developing but is much sparser (see Section 2 for an overview of the current state of research in the domain). Motivated by the need for a passenger-centric representation of the problem, this work proposes a model that describes the railway recovery problem in a way that takes passengers explicitly into account.

As a first step, this paper presents a timetable evaluation tool that computes, for a given timetable, several passenger satisfaction indicators, such as total passenger travel time, number of connections or train saturation. The evaluation tool takes as input (a) a timetable, and (b) time-dependent passenger origin-destination demand. It then assigns the passenger demand on the trains provided in the timetable according to a generalized travel time function (made of in-vehicle time, waiting time and penalties for connections, early departures or late arrivals) that considers passengers at a disaggregate level. This evaluation tool is then used in order to assess the performance of several recovery strategies in case of severe disruptions in railway networks (as opposed to minor disturbances that can be handled by slight modifications to the timetable). The model can evaluate several recovery strategies that are relevant in case of severe disruptions: train cancellations, partial train cancellations, train re-routings and additional train or bus services.

The remainder of this paper is structured as follows. Section 2 reviews the current state of research in the train timetable rescheduling area, as well as in the domain of passenger demand assignment models for schedule-based transportation systems. Section 3 then presents the proposed timetable evaluation tool in detail and Section 4 describes how a disposition timetable can be generated. Finally, Section 5 concludes the paper and provides directions for future research.

## 2 Literature review

The literature review presented in this section focuses on two main topics. First, recent contributions to the train timetable rescheduling (TTR) problem are reviewed extensively (Section 2.1). The publications are classified according to three criteria that facilitate the identification of gaps where contributions can be made to the TTR literature, hence justifying the relevance of this work. Second, schedule-based passenger assignment models are reviewed in Section 2.2.

### 2.1 Train timetable rescheduling

The recent TTR literature can be classified according to three main criteria (Cacchiani *et al.*, 2013):

- i) Distinction between disturbances and disruptions. In a railway network, a disturbance is a primary delay (i.e., a process that takes longer than initially scheduled) — or a set of primary delays — that causes secondary delays that can be handled by rescheduling the timetable only, without rescheduling the resource duties (such as crews and rolling stock). On the other hand, a disruption is a (relatively) large external incident strongly influencing the timetable and requiring resource duties to be rescheduled as well.
- ii) Distinction between microscopic and macroscopic representation of the railway infrastructure. In a microscopic approach, the railway infrastructure is modelled very precisely, sometimes at the switch or track section level, in order to compute detailed running times and headways between trains. In a macroscopic approach, the infrastructure is considered at a higher level where stations and tracks are represented by nodes and arcs in a graph, respectively. Details such as signals or track sections are ignored.
- iii) Distinction between operations-centric or passenger-centric models. Operations-centric models focus on minimizing parameters related to the train company operations, such as delays or the number of cancelled trains, whereas passenger-centric models focus on minimizing the negative effects of disruptions and disturbances for passengers.

The thorough review of railway recovery models presented in Cacchiani *et al.* (2013) shows that the major part of the recent scientific literature deals with disturbances rather than disruptions. Further, in most papers, the railway network is represented at the microscopic rather than at the macroscopic level. Most papers also have an operations-centric approach to railway timetable rescheduling, instead of a passenger-centric view.

### 2.1.1 Disturbances

**Microscopic approach** A major part of the recent scientific literature on TTR has been dedicated to disturbances considered at the microscopic level.

One of the most widely used concepts, the Alternative Graph (AG) model, originally introduced by Mascis and Pacciarelli (2002), has been extensively used in this domain: D'Ariano *et al.* (2007) present a branch-and-bound algorithm based on an AG model for scheduling trains in real-time and D'Ariano *et al.* (2008) describe the implementation of the AG model in the Railway traffic Optimization by Means of Alternative graphs (ROMA) tool. In ROMA, the combined problem of train sequencing and routing is approached iteratively: for given train routes, an optimal sequencing is computed using the branch-and-bound algorithm introduced in D'Ariano *et al.* (2007), and then the solution is improved by local rerouting of trains. The ROMA tool was extensively used thereafter: see, e.g., Corman *et al.* (2009, 2010a,b,c, 2011, 2012) and D'Ariano (2008).

There is however a large amount of additional contributions to the field that propose other microscopic models for the TTR problem. As they are only of indirect interest to this work, they will not be reviewed here; refer to Cacchiani *et al.* (2013), and references therein, for a complete overview.

**Macroscopic approach** The literature discussing TTR for disturbances on a macroscopic level can be split in two categories: operations-centric and passenger-centric formulations.

Operations-centric contributions to the literature include Törnquist and Persson (2007), Törnquist Krasemann (2012), Acuna-Agost *et al.* (2011) and Min *et al.* (2011). Törnquist and Persson (2007) introduce a Mixed-Integer Linear Programming (MILP) model for the rescheduling problem on a N-tracked network, with binary decision variables allowing for reordering and local track changes of trains. The objective function minimizes the final arrival delay of all trains. As some instances could not get solved in reasonable computational time, a greedy heuristic is introduced in Törnquist Krasemann (2012). Acuna-Agost *et al.* (2011) extend the

MILP formulation of Törnquist and Persson (2007) by adapting feasible travel times on tracks according to unplanned stops and by allowing more than one train per track section for trains running in the same direction. Min *et al.* (2011) propose a MILP formulation to solve the train-conflict resolution problem. Based on the observation that if the timetable is fixed at the segment level, the remaining problem is again a train-conflict resolution problem, a column generation-based method is used to solve the problem. The objective function minimizes the difference to the original timetable.

The literature on passenger-centric formulations is mainly based on the delay management problem (DMP) introduced by Schöbel (2007), which decides if connecting trains should wait for a delayed feeder train or if they should depart on time. The main decision variables of the proposed MIP model, based on an event-activity network, are binary variables deciding if a passenger connection between two trains is maintained or if it is dropped. The objective is to minimize the total passenger delay. The DMP is extended in Schöbel (2009) and Schachtebeck and Schöbel (2010) to include infrastructure capacity constraints, by introducing disjunctive constraints for conflicting train paths. Dollevoet *et al.* (2012) further extend the DMP to include re-routing of passengers. The routing decisions of the passengers are modelled with binary variables that describe if a connection is used by a passenger on its path from its origin to its destination. Finally, station capacities are also considered in another extension of the DMP in (see, e.g., Dollevoet, 2013). Other contributions to passenger-centric formulations are provided by Kumazawa *et al.* (2010) and Kanai *et al.* (2011). The former deals with passenger dissatisfaction by simulating passenger behavior on a train plan created in a previous stage. Further, passenger overflow (i.e., waiting time on train platform due to insufficient train capacity) is analyzed. The latter defines several passenger disutility functions and uses an algorithm based on simulation and optimization to minimize passengers' disutility.

### 2.1.2 Disruptions

**Microscopic approach** Only very few publications deal with disruptions at the microscopic level. Hirai *et al.* (2009) formulate the train stop deployment problem (i.e., decisions about where disrupted trains should stop in order to let unobstructed trains pass) after a disruption as a MIP. The main idea is to penalize stops outside stations or deviating from the original schedule. Corman *et al.* (2011) consider a severe disruption on a double-track network, where some of the tracks become unavailable. The problem is split in two separate dispatching areas, each of which are modelled through an Alternative Graph formulation. Finally, boundary constraints between the two dispatching areas make sure that the local solutions are globally feasible.

**Macroscopic approach** The following papers describe disruptions in railway networks at the macroscopic level. Louwse and Huisman (2012) consider the case of partial and complete blockades in case of a major disruption. They develop a mixed-integer programming model to generate the disposition timetable. Two disruption measures are applied: train cancelling and train delaying. Schedule regularity constraints (e.g., operating approximately the same number of trains in each direction during a partial blockade) are included in the formulation in order to take the rolling stock problem into account implicitly. In case of a complete blockade, both sides of the disruption are considered independently (i.e., trains will reverse before the disrupted area but no coordination with the other side is considered). Albrecht *et al.* (2013) consider the problem of disruptions due to track maintenance, arising when maintenance operations take longer than scheduled and thus force to cancel additional trains. A disposition timetable including track maintenance is constructed using a problem space search meta-heuristic. This heuristic is also used to generate quickly disposition timetables in case of a disrupted system. Cadarso *et al.* (2013) consider an integrated timetabling and rolling stock problem that accounts for passenger demand by splitting it up into two steps. In the first step, anticipated disrupted demand is computed using a multinomial logit model. As demand figures are estimated before the timetable is adjusted, they are based on line frequencies in an anticipated disposition timetable, rather than on actual arrival and departure times. In the second step, the timetabling and rolling stock rescheduling problem is formulated and solved as a MILP model, subject to the anticipated demand calculated in the first step.

### 2.1.3 Summary and relevance of the present work

As pointed out previously, the overview of the TTR literature showed that passenger-centric models are much less common than operations-centric formulations, and that most of the recent literature deals with disturbances (as opposed to disruptions). Further, one can also notice the lack of network-wide disruption recovery frameworks in this field. This is why the present work focuses on passenger satisfaction indicators to evaluate timetables in case of severe disruptions and proposes to assess the impacts of recovery strategies on a network level.

The formulation coming closest to the present work is the one by Cadarso *et al.* (2013), where the effects of disruptions on the passenger demand is explicitly dealt with. However, the main difference between the two approaches is that, in Cadarso *et al.* (2013), the passenger demand is evaluated on an expected timetable before solving the timetable and rolling stock rescheduling problem, whereas in our case, the passenger demand is assigned on an actual timetable (with determined departure and arrival times for every train) in order to evaluate passenger satisfaction indicators. Furthermore, the interaction between demand and supplied capacity is ignored in Cadarso *et al.* (2013).

## 2.2 Passenger assignment models

The recent literature on passenger assignment models for transit systems is either frequency-based or schedule-based. In the former approach, transit services are represented by lines with travel frequencies and single vehicles are not explicitly considered. Single vehicle loads can therefore only be approximated. As this work is interested in exceptional events such as passenger demand peaks in case of severe disruptions, an explicit modelling of the remaining available capacity of the (presumably irregular, and therefore not frequency-based anymore) vehicles is necessary. The frequency-based approach will therefore not be reviewed here.

In the schedule-based model, each vehicle is considered individually with its capacity, either implicitly or explicitly. The implicit approach is similar to road network modelling, where link costs are related to link flows through non-decreasing functions. This method has the main disadvantage that the effects of congestion (represented by a general discomfort increase) are equal for all users of a train run, whereas they should be different for passengers already onboard and passengers trying to board a congested train at a station. The explicit schedule-based model deals with this issue by introducing vehicle capacity constraints, thus letting waiting passengers board the arriving train according to its residual capacity. The following papers use the explicit schedule-based approach to assign passengers on transit networks.

Tong and Wong (1999) formulate a stochastic transit assignment model on a dynamic schedule-based network, where passengers are assumed to travel on paths of minimal generalized travel cost, consisting of in-vehicle time, waiting time, walking time and transfer penalty time, weighted by a sensitivity coefficient. Stochasticity is included in the formulation by allowing these sensitivity coefficients to be randomly generated instead of being constant for every passenger. For given values of the sensitivity coefficients, the passenger demand is loaded onto the transit network by means of an all-or-nothing assignment process.

Nguyen *et al.* (2001) consider the case where timetables are reasonably reliable, and the number and frequencies of transit vehicles are low. For this kind of networks, departure time and route choice are both equivalently important decisions that passengers face. It is the first paper that is not limited to a single origin-destination pair and that considers both departure time choice and route selection simultaneously. Further, the concept of path available capacity is introduced in order to capture the flow priority aspect (i.e., giving priority to passengers already onboard the transit vehicles with respect to passengers waiting at the station). A traffic equilibrium model of the assignment problem is presented, and a computational procedure based on asymmetric boarding penalty functions is suggested to avoid the explicit enumeration of all paths connecting origins and destinations.



Poon *et al.* (2004) propose a model that explicitly describes the available capacity of every vehicle at each station, as well as the queuing time for every passenger (assuming a First-In-First-Out queue for passengers waiting at the station). The paper focuses on the route choice problem, ignoring other choice dimensions, such as departure time or departure station. In their formulation, route choice for every passenger is modelled by selecting a path that minimizes a generalized cost function consisting of in-vehicle time, waiting time, walking time and line change penalties. The network is loaded (i.e., user equilibrium is achieved) by using a Method of Successive Averages algorithm.

Hamdouch and Lawphongpanich (2008) also propose a user-equilibrium transit assignment model that explicitly considers individual vehicle capacities. For every O-D pair, passengers are divided into groups according to their desired arrival time intervals. It is assumed that every passenger group has a travel strategy resulting, at each station and each point in time, in a list of subsequent travel options that are ordered according to the passenger groups' preferences to continue their trip. Passenger preferences are described by the minimization of expected travel costs, made of in-vehicle time, fare and costs associated with early departures from home and/or arrivals outside the desired arrival time interval. Travel strategies can therefore be adaptive over time. When loading a vehicle at a station, onboard passengers continuing to the next station remain in the vehicle and waiting passengers are loaded according to the available remaining vehicle capacity. If the vehicle is full, passengers unable to board need to wait for the next vehicle. Demand-supply interactions are defined by a user equilibrium approach and a solution method based on successive averages is proposed.

Nuzzolo *et al.* (2012) propose a new schedule-based dynamic assignment problem for congested transit networks, explicitly considering vehicle capacities. It is new in the sense that more complex behavioral choice models are used for passengers (including updating of departure time, access stop and transit vehicle run), especially in the case of failure-to-board experiences. A learning process for the passengers is also included in the model. In this formulation, passengers are characterized by their origins and destinations, as well as desired departure or arrival times. They are also assumed to be flexible (in a certain range) in order to avoid congestion effects.

### **3 Passenger assignment model**

The timetable evaluation tool is constructed by considering three different layers of the problem: the infrastructure level consisting of tracks and stations, the train route level describing the train lines and the passenger level for passenger paths. Section 3.1 describes these three levels and explains how they build upon each other. Section 3.2 then presents the assignment procedure

that is used to “load” the passengers on the train network; some of its elements are inspired by the works presented in the literature review on schedule-based transit assignment models of Section 2.2.

### 3.1 The three representation layers of the passenger assignment problem

The problem of assigning passengers on a railway network consists of three different layers. The two lower levels model the supply side of the railway network, while the top layer represents the demand as passenger paths applied on the supplied network.

**Railway infrastructure graph** The railway infrastructure (tracks and stations) represents the lowest layer of the problem. It can be visualized as a graph with a node for every station and an arc for every track existing in the real world (see Fig. 1 for an example). It is at this level that the disruption is initially represented by removing one or several arcs from the graph (or by imposing different headway requirements on the arcs in the case of a partial track unavailability). The components of the infrastructure graph remain static during the assignment process.

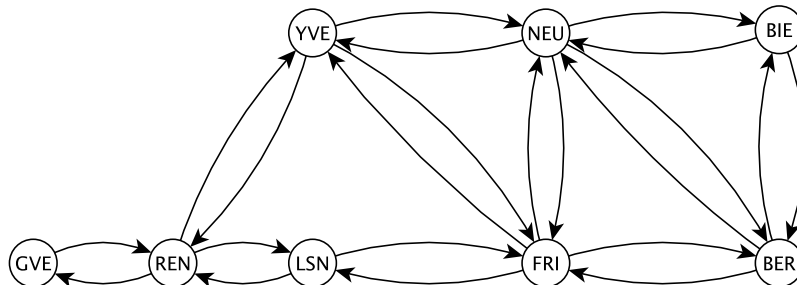


Figure 1: Example of infrastructure graph (based on the Western part of the Swiss railway network).

**Route network** The second layer is the route network which depicts the routes of all train lines in a static manner, taking as an input a timetable that lists the visited stations for every train line. The set of nodes and arcs in this graph will be denoted by  $N$  and  $A$ , respectively. A node  $s \in N$  in the route network stands for a station where a train can either stop to load and unload passengers or pass by without stopping. Arcs correspond to route segments of train lines. Each arc  $(s, s')$  in the route network is weighted by its travel time  $tt_{s,s'}$ , defined as the deterministic time between departure from station  $s$  and arrival at station  $s'$ . The route network builds upon the infrastructure graph in the sense that train routes connect stations represented in

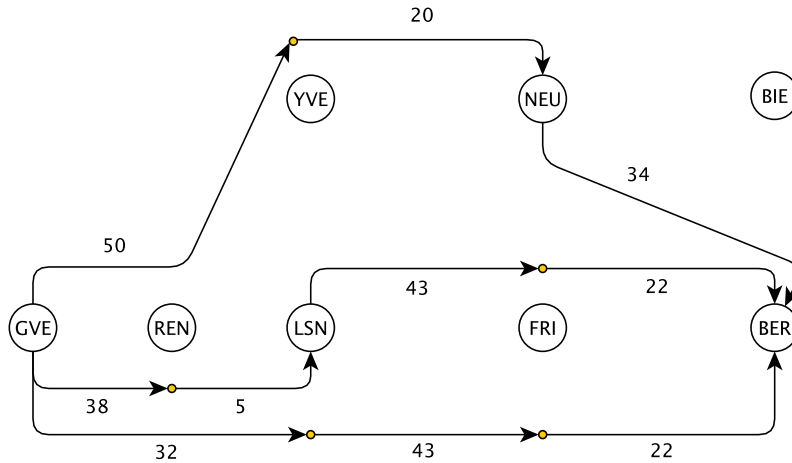


Figure 2: Example route network using the infrastructure graph of Fig. 1, with five train lines. The arcs represent route segment. Travel times are indicated for every arc in minutes.

the infrastructure graph. Further, train routes can only use arcs that exist in the infrastructure graph.

To illustrate, Fig. 2 gives an example of a route network, based on the infrastructure graph of Fig. 1, with five train lines. Nodes GVE, REN, LSN, YVE, NEU, FRI, BER and BIE are station nodes. The arcs correspond to the line segments of the five train lines. For instance, a train line begins its route at node GVE, stops at nodes LSN and FRI, and terminates at node BER. The line thus consists of three route segments (GVE, LSN), (LSN, FRI) and (FRI, BER).

**Time-expanded network** Because of its static nature, a route network cannot display parameters varying with time, such as train occupation rates or passenger flows. It is also not possible to distinguish events taking place at different times, such as passengers leaving their origin or trains starting on a train line. One way to incorporate this kind of temporal information is by using a time-expanded network (TEN) — see, e.g., Ahuja *et al.* (1993), Nguyen *et al.* (2001), Hamdouch and Lawphongpanich (2008). This is the third layer of the problem and passenger demand is represented as paths from origin to destination in this network.

Let  $[0, T]$  represent the daily operating interval of the railway system (i.e.,  $T$  stands for the end time of the last operation of the day). In a time-expanded network, the operating interval is discretized into a set of  $n \in \mathbb{N}$  points, with times  $0, \tau, 2\tau, \dots, n\tau$ .  $\tau = T/n$  is the time step of the system — in practice,  $\tau$  can be equal to 1 or 5 minutes, for instance. It is assumed that all times (departure times, arrival times, dwell times) are integer multiples of the time step.

In a time-expanded network (TEN), each node  $s \in N$  in the route network is expanded into

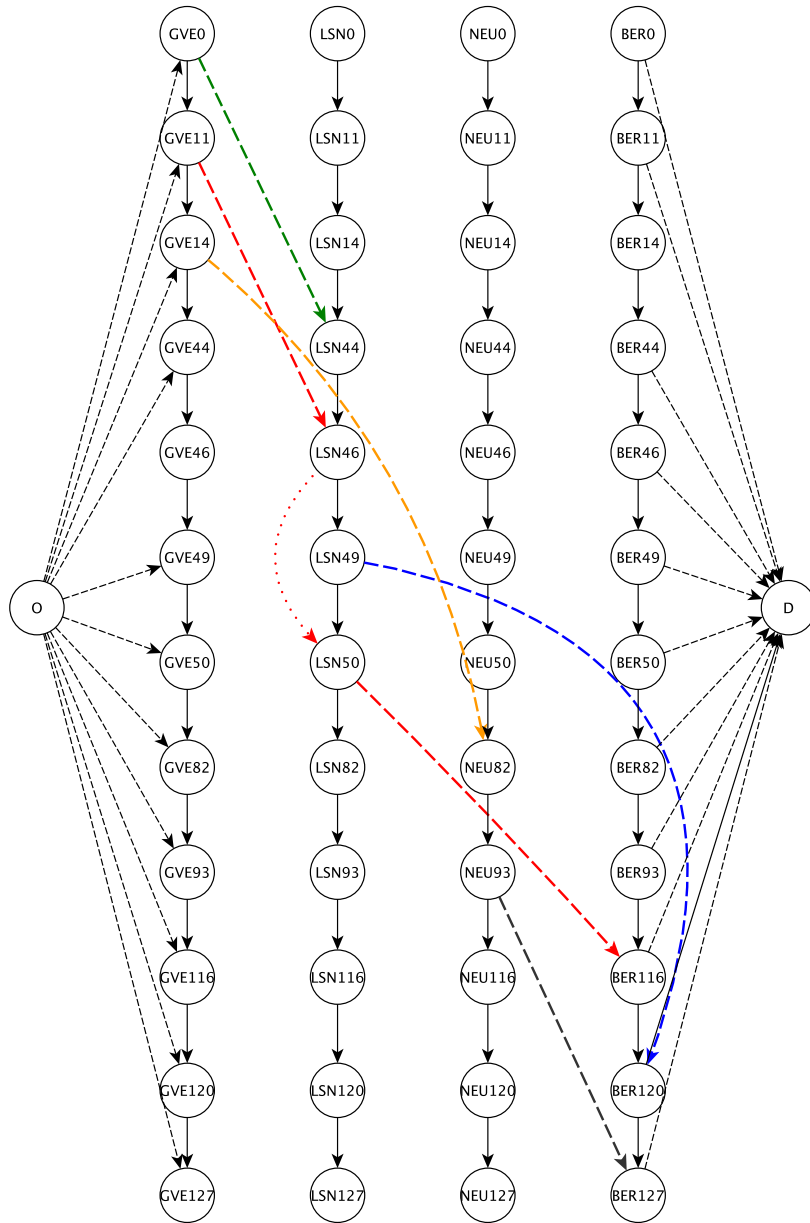


Figure 3: Time-expanded network based on the route network of Fig. 2 and the timetable of Table 1, for origin-destination pair GVE-BER. Coloured dashed arcs represent train driving arcs, coloured dotted arcs train waiting arcs, dashed arcs passenger walking arcs and solid arcs stand for passenger waiting arcs. Note that only time steps where a train arrival or departure takes place are represented to improve the readability of the graph.

$(n + 1)$  nodes  $s_t$ , with  $t = 0, \tau, \dots, n\tau$ . Thus, a node in a TEN has two labels: one for space,  $s$ , representing the station and one representing the time step,  $t$ . Additionally, for every passenger origin-destination pair, two time-invariant nodes are included: one for the origin and one for the destination. The set of all time-expanded nodes will be denoted by  $N^T = N^* \cup O \cup D$ , where  $N^*$  is the set of space-time nodes of the form  $s_t$ , and  $O$  and  $D$  are respectively the sets of time-invariant origins and destinations. Further, there are four types of arcs in the time-expanded

network: train driving arcs, train waiting arcs, passenger waiting arcs and passenger walking arcs. They will be denoted by  $A_1^T, A_2^T, A_3^T$  and  $A_4^T$ , respectively.

- Similarly to the time-expanded nodes, an arc  $(s, s')$  in the route network is expanded into *train driving arcs* of the form  $(s_t, s'_{t+tt_{ss'}})$ , where, as defined earlier,  $tt_{ss'}$  denotes the travel time on arc  $(s, s')$  in the route network. The number of time-expanded arcs  $(s_t, s'_{t+tt_{ss'}})$  depends on the timetable of each train line: for each arc in the route network, there will be as many arcs in the TEN as there are train runs for this particular line during  $[0, T]$ . The travel time on this arc is given by  $tt_{(s_t, s'_{t+tt_{ss'}})} = tt_{ss'}$ .
- *Train waiting arcs* are included in the TEN in order to link the arrival of a train at a station with its subsequent departure. A train waiting arc representing a train arriving at station  $s$  at time  $a$  and leaving at time  $d$  is of the form  $(s_a, s_d)$ . The waiting time on this arc is given by  $wt_{(s_a, s_d)} = t_d - t_a$ .
- *Passenger waiting arcs* of the form  $(s_t, s_{t+\tau})$  are included in the TEN to model passengers waiting (while not onboard a train) at every station  $s$  from time  $t$  to time  $t + \tau$ . The waiting time on this arc is the time step:  $wt_{(s_t, s_{t+\tau})} = \tau$ .
- Finally, *passenger walking arcs* of the form  $(o, s_t)$  or  $(s_t, d)$  model passengers walking from their origin to their first station at time  $t$  and passengers walking from their last station to their destination at time  $t$ . The walking time is assumed to be zero.

Together, the four types of arcs form the set of time-expanded arcs, i.e.  $A^T = \cup_{i=1}^4 A_i^T$ . Every arc in  $A^T \setminus A_4^T$  is weighted by its respective waiting or driving time. Passenger walking arcs are weighted in order to model penalty costs for early or late departures or arrivals (see Section 3.2.3). Further, every arc in  $A^T$  has a passenger capacity, defined by the capacity of the train run it represents (passenger waiting and walking arcs have infinite capacity). The passenger flows on the arcs in  $A^T$  are determined in the assignment process described in Section 3.2, taking these arc capacities into account.

As an illustration, Fig. 3 shows the time-expanded network for the route network of Fig. 2, based on the example timetable presented in Table 1. It displays all possible passenger paths for one passenger origin-destination pair, between origin node GVE and destination node BER, with a start time between 0 and 127.<sup>1</sup> For instance, the path GVE0→LSN44→LSN46→LSN50→BER116 corresponds to passengers whose origin is GVE, boarding train IR1403 at time 0, arriving at station LSN at time 44, waiting at station LSN until time 46, boarding train IR2517 at time 46, leaving station LSN at time 50 on that train and arriving at station BER at time 116.

<sup>1</sup>Note that passengers arriving after time 14 at station GVE will not be able to reach their destination in this simple example network. In reality, there are more trains in order to accommodate the demand, but for ease of representation they are not shown in Fig. 3.

	GVE	LSN	NEU	BER
IR1403	7:00	7:44	-	-
IR2511	-	7:49	-	9:00
IR2517	7:11	7:46	-	-
IR2517	-	7:50	-	8:56
ICN617	7:14	-	8:22	-
RE3029	-	-	8:33	9:07

Table 1: Example timetable for the route network.

## 3.2 Passenger assignment

### 3.2.1 Passenger demand representation

The TEN framework lists all possible passenger paths, for all passenger origin-destination pairs, in the supplied train route network. The question that needs to be answered in the next step is the assignment of passengers to the different arcs in the network, i.e. to model passenger decisions regarding departure time from origin as well as path choice.

In contrast to the road network case, where it is commonly assumed that travel times on a link increase with the number of passengers (vehicles) on the link, the number of passengers aboard a train usually does not have a direct influence on the length of the train trip. Nevertheless, congestion on the train network (i.e., lack of available capacity) may force a passenger to wait for the next available train, thus extending his journey time. This might particularly be the case when a severe disruption happens in the network and trains are cancelled. It is therefore necessary to focus on train capacity constraints in order to model the asymmetric passengers inter-influence between passengers already aboard and boarding passengers.

### 3.2.2 Departure time choice

We assume that there are two types of passengers: passengers with a desired departure time (DDT) from their origin and passengers with a desired arrival time (DAT) at their destination. An example of the first type of passengers would be passengers taking the train system to commute back home in the evening or for leisure-related trips, while passengers commuting to work in the morning would be an example of the second type of passengers.

For every origin-destination pair, the passenger demand is subdivided into groups (indexed by  $g$ ), by desired departure time from origin:  $\mathbb{D}(o, d, g)$  stands for the set of passengers with origin  $o \in O$ , destination  $d \in D$ , and desired departure time  $DDT(g) \in \{0, \tau, \dots, n\tau\}$ . For passengers with a DDT, the separation of passengers into the different groups is straightforward. On the other hand, the departure time choice of passengers having a DAT can be modelled by traversing the TEN in reverse topological order (see, e.g., Nguyen *et al.* (2001)) in order to compute the (free-flow) latest possible time  $L(o, d, DAT)$  a passenger can depart from his origin and still arrive at destination at the latest at his DAT. Passengers in this situation are added to the passenger demand  $\mathbb{D}(o, d, g)$  group with  $DDT(g) = L(o, d, DAT)$ .

To illustrate the departure time choice, refer to Fig. 3. Assume a DDT demand of 100 passengers from GVE to BER, with a DDT from GVE equal to 14. Further, for the same origin-destination pair, assume a DAT demand of 50 passengers with a DAT at BER of 127. The latest possible departure time from GVE is equal to 14 for these 50 passengers (as they can take train ICN617 from GVE to NEU, change in NEU to train RE3029 and arrive at BER at time 127). The total demand on origin-destination pair (GVE, BER) with DDT=14 is thus equal to 150 passengers.

### 3.2.3 Route choice

**Generalized travel time** To model a passenger's route choice in the TEN, a linear disutility function is associated with every path between pairs of origin and destination nodes (see, e.g., Nguyen *et al.* (2001), Poon *et al.* (2004), Tong and Wong (1999)). This disutility function encompasses four components: in-vehicle driving time, in-vehicle waiting time, off-vehicle waiting time and potential penalty costs. For every passenger group  $g$ , a generalized travel cost is thus defined for a path  $p \in P_{o,d}$ , where  $P_{o,d}$  is the set of all paths in the TEN between nodes  $o \in O$  and  $d \in D$ , as a weighted combination of these four components:

$$C_p(g) = \sum_{(i,j) \in A_p \cap A_1^T} tt_{(i,j)} + \beta_2 \cdot \sum_{(i,j) \in A_p \cap A_2^T} wt_{(i,j)} + \beta_3 \cdot |A_p \cap A_3^T| \cdot \tau + \eta_1 \cdot M_p + K_p(g). \quad (1)$$

$A_p \subset A^T$  is the set of time-expanded arcs that path  $p$  uses. Here it is assumed that the weights of the various components of the disutility function are defined relative to the in-vehicle driving time of the path, and  $\beta_2$  and  $\beta_3$  denote the weighting factors for in-vehicle waiting time and off-vehicle waiting time, respectively (usually,  $\beta_3 \geq \beta_2 \geq 1$ ).  $\eta_1 > 0$  is the line change penalty in weighted time units<sup>2</sup> and  $M_p$  is the number of line changes along path  $p$ .  $K_p(g)$  denotes the penalty cost associated with early or late arrivals or departures, for path  $p$  and passenger group  $g$ . Similarly to Nguyen *et al.* (2001), the penalty cost of early or late departures from the

<sup>2</sup>See, e.g., Guo and Wilson (2011), Wardman (2004) for values for  $\eta_1$ .

passenger's origin is modelled by weighting the passenger walking arcs of the form  $(o, s_t)$  using the following arc penalty costs:

$$D_{(o,s_t)}(g) = \begin{cases} 0 & \text{if } t = DDT(g), \\ \delta_1(DDT(g) - t) & \text{if passenger has a DDT and } DDT(g) > t, \\ \delta_2(t - DDT(g)) & \text{if passenger has a DDT and } DDT(g) < t, \\ \max(\delta_3(DDT(g) - t), 0) & \text{if passenger has a DAT,} \end{cases} \quad (2)$$

where  $\delta_1, \delta_2, \delta_3$  are positive scalars. Note that in the case of a passenger with a desired arrival time,  $DDT(g) = L(o, d, DAT(g))$  and no penalty is incurred at the origin whenever the passenger leaves after  $L(o, d, DAT(g))$ . Likewise, late and early arrivals at the passenger's destination also induce a penalty cost that is modelled by weighting the passenger walking arcs of the form  $(s_t, d)$  using the following arc penalty costs<sup>3</sup>:

$$A_{(s_t,d)}(g) = \begin{cases} 0 & \text{if } t = DAT(g), \\ \alpha_1(DAT(g) - t) & \text{if } DAT(g) > t, \\ \alpha_2(t - DAT(g)) & \text{if } DAT(g) < t, \end{cases} \quad (3)$$

where  $\alpha_1, \alpha_2$  are positive scalars. Hence, for a path  $p$  of the form  $\{o, s_t, \dots, s'_t, d\}$ , the total penalty cost is the sum of the two penalty costs:

$$K_p(g) = D_{(o,s_t)}(g) + A_{(s'_t,d)}(g). \quad (4)$$

Note that the weighting factors  $\beta_2, \beta_3, \eta_1, \delta_1, \delta_2, \delta_3, \alpha_1, \alpha_2$  are user-defined parameters and can differ from one passenger to the next. Tong and Wong (1999) for instance introduce stochasticity in the assignment procedure by randomly generating the sensitivity coefficients from known density functions instead of keeping them constant for every passenger.

**Assignment procedure** The aim of the passenger assignment model is to determine the passenger flows on every arc of the TEN, taking train capacity constraints into account. A time-dependent shortest path-type algorithm associated with the above network is proposed. It consists of three main phases:

1. Assign passenger groups on ideal path according to  $C_p(g)$ .
2. If an arc capacity is exceeded, decide which passengers need to be re-assigned. Otherwise, stop.
3. Re-assign unassigned passengers on reduced network, then go to step 2.

<sup>3</sup>Here it is assumed that passengers with a desired departure time do not care about their arrival time at destination.



In the first phase, it is assumed that all passengers have full predictive information about present and future network conditions, and every group of passengers selects the path from origin to destination that minimizes its generalized travel cost function  $C_p(g)$ . A Dijkstra-type algorithm is used to determine the shortest path from origin to destination, then the passenger flows on the arcs of this shortest path are updated accordingly. However, passengers do not know about future train occupancy levels, and passenger flows  $f \in \mathbb{N}^{|A^T|}$  are therefore updated without taking arc capacities into consideration. At the end of this phase of the assignment process, all passengers are assigned on the shortest path (in terms of  $C_p(g)$ ) between their origin and their destination (and thus on the corresponding arcs in  $A^T$ ), but this assignment might not be feasible as arc capacities were not yet considered.

The second phase of the assignment procedure deals with infeasibilities in terms of arc capacities, i.e., for every train waiting and driving arc, the algorithm checks if the passenger flow assigned in the previous phase exceeds the arc capacity. If no arc capacity is exceeded, the assignment is feasible and every passenger reaches his destination by the shortest path (in terms of  $C_p(g)$ ). Otherwise, the surplus passengers are removed from the infeasible arc one by one until its capacity constraint is verified. In that case, priority rules are applied in order to decide which passengers are allowed to remain on the infeasible arc and which are not.

- The passengers with highest priority are those who are already onboard when new passengers try to board the train. In the TEN, these passengers were assigned (during the first phase of the algorithm) on a previous train waiting or driving arc representing the same train run as the infeasible arc. Highest-priority passengers can only be removed from the infeasible arc if the latter is the first train driving or waiting arc of a train run.
- Among passengers that are boarding the train, a first-come-first-served priority rule is applied. That is, for two passengers that want to board the same train, the one that arrived at the station earlier has priority on the one that arrived later.
- Among passengers that are boarding the train and that have arrived at the station at the same time, two priority rules can be applied. The first one is to choose randomly who can board the train and who cannot, but it might not be very realistic. The second option is to calculate the marginal utility loss for every passenger not taking the train and giving priority to passengers with highest marginal utility loss.

Starting from the ones with lowest priority, passengers are thus removed one by one from the passenger flow on the infeasible arc (as well as on following arcs on their path), until the passenger flow does not exceed the arc capacity anymore.

In the third phase of the assignment procedure, passengers that were removed from the passenger flows because of the arc capacity constraints in the second phase are re-assigned on a reduced

TEN. For every unassigned passenger, the reduced TEN is constructed by removing the infeasible arc from  $A^T$ . The assignment procedure of the first phase is then re-run on the reduced TEN for every unassigned passenger.

The assignment procedure iterates between the second and third phases, until all passenger flows are feasible in terms of arc capacities or the set of unassigned passengers remains constant between two iterations, meaning that additional passengers cannot be assigned on the network because of capacity constraints.

### 3.3 Passenger satisfaction indicators

Finally, passenger satisfaction indicators will be derived from the assignment model in order to evaluate the performance of a timetable. The following indicators are proposed:

- **Travel time indicators:** The most important indicator is the time spent by the passengers in the system. It is assumed that every passenger tries to minimize the time between his desired departure time (that is pre-determined) and his actual arrival time at destination. The travel time thus includes time spent waiting at stations. Of interest are total and average travel times for all passengers, as well as maximal travel times for passengers that are worst off. By comparing the values of the indicators in the disrupted case and in the non-disrupted case, the delays incurred by passengers can be assessed.
- **Passenger connection indicators:** It is assumed that passengers want to minimize the number of connections they need to make in order to arrive to destination. Again, total, average and maximal number of connections will be indicated for the evaluated timetable.
- **Train saturation indicators:** The trade-off between the interests of the railway company on the one side and of the passengers on the other is modelled by the inclusion of train saturation (i.e., train capacity utilization) indicators: railway companies try to maximize train utilization while passengers would prefer less saturated trains. Average and maximal train saturation will be evaluated.

## 4 Generation of disposition timetables

**Definition of recovery strategies** In case of a severe disruption, the available infrastructure network is diminished and it is therefore very likely (at least in heavily utilized networks like the Swiss one) that not all trains originally scheduled will be able to run. Recovery strategies are

therefore needed in order to accommodate the passenger demand. Based on the literature and current practice in railway companies, the following recovery strategies may be formulated:

- **Train cancellation:** Cancelling a train that was scheduled to use the unavailable track is the easiest recovery scheme for a train company (if rolling stock considerations are ignored). It resolves the infrastructure conflict and provides additional “space” for other trains in the network. It should however be avoided as much as possible from the passengers’ point of view, as they will have to take another (and presumably later) train to reach their destination.
- **Partial train cancellation:** A partial cancellation (also called turn-around) is a recovery strategy where a train turns around at the last station before the disruption and starts over in the opposite direction, possibly taking over the scheduled service of a train that should have traversed the disrupted area in the opposite direction.
- **Re-ordering of trains:** In case of a partial track unavailability, local train re-ordering may be necessary in order to schedule the trains on the available track capacity.
- **Global re-routing:** In case of a total track unavailability, global rerouting strategies (i.e., routing trains through a different part of the network) for trains need also to be investigated.
- **Additional service:** It might also be necessary to schedule additional service, in the form of extra trains or buses. These services can be scheduled where additional resources are available in depots.

**Generation of disposition timetables** Once the preferred recovery strategies are defined, an actual disposition timetable (i.e., adjusted departure times for every train in the system) is generated and its infrastructural feasibility needs to be ensured. To that end, a macroscopic timetable re-scheduling framework can be used to create the disposition timetable including the recovery strategies.

An example of this kind of macroscopic rescheduling framework is proposed in Kecman *et al.* (2013). In that model, the train rescheduling problem is formulated as a job-shop scheduling problem: the goal is to schedule a finite set of jobs (trains), defined by fixed sequences of operations (train runs and dwellings), that cannot be interrupted, on a finite set of resources (track sections, for instance) that can perform one operation at a time (no-store constraint). The decision variables of the formulation are the scheduled starting times  $x_r^t$  of train  $t$  on resource  $r$  and sequencing variables  $k_r^{t,u}$ , equal to one if  $x_r^t < x_r^u$ , 0 otherwise. The objective function minimizes the start time of the last operation of the day. General precedence constraints, of the form

$$x_r^t \geq x_s^u + p_s^u, \forall r, s \in R, \forall t, u \in T, \quad (5)$$

where  $p_s^u$  is the scheduled processing time of train  $u$  on resource  $s$ ,  $R$  is the set of all resources and  $T$  the set of all trains, are included in order to model the fact that an operation cannot start before the previous one has terminated. Further, sequencing constraints of the form

$$k_r^{t,u} + k_r^{u,t} = 1, \forall r \in R, \forall t, u \in T_r, \quad (6)$$

where  $T_r$  is the set of all trains using resource  $r$ , in order to make sure that trains using the same resource are sequenced properly. This formulation also allows to model headway constraints on resource, in the following way:

$$x_r^t \geq x_r^u + h_r^{u,t} - M \cdot k_r^{t,u}, \forall r \in R_h, \forall t, u \in T_r, \quad (7)$$

where  $h_r^{u,t}$  is the scheduled headway time on resource  $r$  between trains  $u$  and  $t$ ,  $M$  is a large positive number and  $R_h$  is the set of resources where headway constraints need to be enforced.

**Evaluation of disposition timetables** The next step is to compute the passenger satisfaction indicators of the generated disposition timetables with the timetable evaluation tool proposed in Section 3, thus enabling to compare the performance of different timetables. The evaluation of timetables generated according to current practice will also be performed, in order to compare their performance with the ones we present.

## 5 Conclusion

Motivated by the need for a passenger-centric framework for the train timetable rescheduling problem, this paper presented a timetable evaluation tool, that computes, for a given timetable, passenger satisfaction indicators, such as total passenger travel time, number of connections and train saturation. The evaluation tool is used to assess the performance of several recovery strategies in case of severe disruptions in railway networks.

The railway disposition timetables that are used as an input to the evaluation tool are generated using a macroscopic job-shop scheduling formulation. Disposition timetables are generated in this framework by including recovery strategies one at a time (on a trial-and-error basis) to be evaluated thereafter. This aggregation of recovery strategies to form a disposition timetable is based on the experience acquired in the evaluation phase. As this procedure only allows for the evaluation of a limited set of recovery strategies, it is not clear if it yields an optimal solution (in terms of passenger satisfaction), nor how close this solution is to an optimal solution. Directions for further research thus include the introduction of an optimization framework that will build

upon the knowledge gained by the evaluation of single disposition timetables.

## 6 References

- Acuna-Agost, R., P. Michelon, D. Feillet and S. Gueye (2011) A MIP-based Local Search Method for the Railway Rescheduling Problem, *Networks*, **57** (1) 69–86.
- Ahuja, R. K., T. Magnanti and J. Orlin (1993) *Network Flows: Theory, Algorithms, and Applications*, Prentice Hall, NJ.
- Albrecht, A., D. Panton and D. Lee (2013) Rescheduling rail networks with maintenance disruptions using Problem Space Search, *Computers & Operations Research*, **40** (3) 703–712, March 2013, ISSN 03050548.
- Cacchiani, V., D. Huisman, M. Kidd, L. Kroon, P. Toth, L. Veelenturf and J. Wagenaar (2013) An Overview of Recovery Models and Algorithms for Real-time Railway Rescheduling, *Technical Report*, Erasmus University (Institute).
- Cadarso, L., A. Marín and G. Maróti (2013) Recovery of disruptions in rapid transit networks, *Transportation Research Part E: Logistics and Transportation Review*, **53**, 15–33, July 2013, ISSN 13665545.
- Corman, F., A. D’Ariano and I. A. Hansen (2010a) Disruption handling in large railway networks, *Computers in Railways XII*, 629–640.
- Corman, F., A. D’Ariano, I. A. Hansen, D. Pacciarelli and M. Pranzo (2011) Dispatching trains during seriously disrupted traffic situations, paper presented at the *IEEE International Conference on Networking, Sensing and Control (ICNSC)*, 2011, 323–328, April 2011.
- Corman, F., A. D’Ariano, D. Pacciarelli and M. Pranzo (2009) Evaluation of green wave policy in real-time railway traffic management, *Transportation Research Part C: Emerging Technologies*, **17** (6) 607–616, December 2009, ISSN 0968090X.
- Corman, F., A. D’Ariano, D. Pacciarelli and M. Pranzo (2010b) A tabu search algorithm for rerouting trains during rail operations, *Transportation Research Part B: Methodological*, **44** (1) 175–192, January 2010, ISSN 01912615.
- Corman, F., A. D’Ariano, D. Pacciarelli and M. Pranzo (2010c) Centralized versus distributed systems to reschedule trains in two dispatching areas, *Public Transport*, **2** (3) 219–247, November 2010, ISSN 1866-749X.

- Corman, F., A. D'Ariano, D. Pacciarelli and M. Pranzo (2012) Bi-objective conflict detection and resolution in railway traffic management, *Transportation Research Part C: Emerging Technologies*, **20** (1) 79–94, February 2012, ISSN 0968090X.
- D'Ariano, A. (2008) Improving Real-Time Train Dispatching: Models, Algorithms and Applications, Ph.D. Thesis, TRAIL Thesis Series T2008/6, The Netherlands.
- D'Ariano, A., F. Corman, D. Pacciarelli and M. Pranzo (2008) Reordering and Local Rerouting Strategies to Manage Train Traffic in Real Time, *Transportation Science*, **42** (4) 405–419, November 2008, ISSN 0041-1655.
- D'Ariano, A., D. Pacciarelli and M. Pranzo (2007) A branch and bound algorithm for scheduling trains in a railway network, *European Journal of Operational Research*, **183** (2) 643–657, December 2007, ISSN 03772217.
- Dollevet, T. (2013) Delay Management and Dispatching in Railways, Ph.D. Thesis, Erasmus University Rotterdam.
- Dollevet, T., D. Huisman, M. Schmidt and A. Schöbel (2012) Delay Management with Re-Routing of Passengers, *Transportation Science*, **46** (1) 74–89.
- Guo, Z. and N. H. Wilson (2011) Assessing the cost of transfer inconvenience in public transport systems: A case study of the London Underground, *Transportation Research Part A: Policy and Practice*, **45** (2) 91–104, February 2011, ISSN 09658564.
- Hamdouch, Y. and S. Lawphongpanich (2008) Schedule-based transit assignment model with travel strategies and capacity constraints, *Transportation Research Part B: Methodological*, **42** (7-8) 663–684, August 2008, ISSN 01912615.
- Hirai, C., T. Kunimatsu, N. Tomii, S. Kondou and M. Takaba (2009) A Train Stop Deployment Planning Algorithm Using a Petri-net-based Modelling Approach, *Quarterly Report of RTRI*, **50** (1) 8–13.
- Kanai, S., K. Shiina, S. Harada and N. Tomii (2011) An optimal delay management algorithm from passengers' viewpoints considering the whole railway network, *Journal of Rail Transport Planning & Management*, **1** (1) 25–37, November 2011, ISSN 22109706.
- Kecman, P., F. Corman, A. D'Ariano and R. M. Goverde (2013) Rescheduling models for railway traffic management in large-scale networks, *Public Transport*, **5** (1-2) 95–123, March 2013, ISSN 1866-749X.
- Kumazawa, K., K. Hara and T. Koseki (2010) A Novel Train Rescheduling Algorithm for Correcting Disrupted Train Operation in a Dense Urban Environment, in I. A. Hansen (ed.) *Timetable Planning and Information Quality*, 199–208, WIT Press.

- Louwerse, I. and D. Huisman (2012) Adjusting a Railway Timetable in case of Partial or Complete Blockades, *Technical Report*, Technical Report EI-2012-23, Econometric Institute of Erasmus University Rotterdam.
- Mascis, A. and D. Pacciarelli (2002) Job-shop scheduling with blocking and no-wait constraints, *European Journal of Operational Research*, **143** (3) 498–517, December 2002, ISSN 03772217.
- Min, Y.-H., M.-J. Park, S.-P. Hong and S.-H. Hong (2011) An appraisal of a column-generation-based algorithm for centralized train-conflict resolution on a metropolitan railway network, *Transportation Research Part B: Methodological*, **45** (2) 409–429, February 2011, ISSN 01912615.
- Nguyen, S., S. Pallotino and F. Malucelli (2001) A Modeling Framework for Passenger Assignment on a Transport Network with Timetables, *Transportation Science*, **3**, 238–249.
- Nuzzolo, A., U. Crisalli and L. Rosati (2012) A schedule-based assignment model with explicit capacity constraints for congested transit networks, *Transportation Research Part C: Emerging Technologies*, **20** (1) 16–33, February 2012, ISSN 0968090X.
- Poon, M., S. Wong and C. Tong (2004) A dynamic schedule-based model for congested transit networks, *Transportation Research Part B: Methodological*, **38** (4) 343–368, May 2004, ISSN 01912615.
- Schachtebeck, M. and A. Schöbel (2010) To Wait or Not to Wait - And Who Goes First? Delay Management with Priority Decisions, *Transportation Science*, **44** (3) 307–321.
- Schöbel, A. (2007) Integer Programming Approaches for Solving the Delay Management Problem, in F. Geraets, L. Kroon, A. Schoebel, D. Wagner and C. Zaroliagis (eds.) *Algorithmic Methods for Railway Optimization*, vol. 4359 of *Lecture Notes in Computer Science*, 145–170, Springer Berlin Heidelberg, ISBN 978-3-540-74245-6.
- Schöbel, A. (2009) Capacity constraints in delay management, *Public Transport*, **1** (2) 135–154, June 2009, ISSN 1866-749X.
- Tong, C. O. and S. C. Wong (1999) A stochastic transit assignment model using a dynamic schedule-based network, *Transportation Research Part B: Methodological*, **33** (2) 107–121.
- Törnquist, J. and J. A. Persson (2007) N-tracked railway traffic re-scheduling during disturbances, *Transportation Research Part B: Methodological*, **41** (3) 342–362, March 2007, ISSN 01912615.

Törnquist Krasemann, J. (2012) 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, February 2012, ISSN 0968090X.

Wardman, M. (2004) Public transport values of time, *Transport Policy*, **11** (4) 363–377, October 2004, ISSN 0967070X.