

Meeting functional requirements for real-time railway traffic management with mathematical models

Ambra Toletti · Marco Laumanns · Peter Grossenbacher · Ulrich Weidmann

Abstract Railway traffic management is the process of executing production plans, supervising traffic, detecting conflicts, and applying dispatching measures to resolve them. The activity space of traffic managers is constrained by infrastructure, rolling stock, and operations related features and rules. These constraints result in functional requirements that mathematical models should satisfy to be useful to support real-time traffic management. First, this paper identifies and describes these requirements. Second, it presents an overview of the most prominent mathematical formulations for timetabling and rescheduling in the literature. Using sample scheduling problems, it highlights how individual train runs and interactions are represented. The similarities and the differences among the formulations are strengthened to distinguish two main categories of models. Finally, the models are analysed with respect to the functional requirements from real-time traffic management. As expected, off-line scheduling models do not satisfy the requirements related to real-time monitoring and intervention. In contrast, on-line models satisfy most requirements. In particular, there are two rescheduling models that satisfy all functional requirements: the Resource Conflict Graph and an extension of the Alternative Graph.

Keywords Dispatching · functional requirements · mathematical models · railway operations · real-time traffic management · rescheduling

A. Toletti
ETH Zurich - Institute for Transport Planning and Systems, 8093 Zurich, Switzerland
E-mail: ambra.toletti@ivt.baug.ethz.ch

M. Laumanns
IBM Research, 8803 Rüschlikon, Switzerland

P. Grossenbacher
SBB - Infrastructure Division, 3000 Bern, Switzerland

U. Weidmann
ETH Zurich - Institute for Transport Planning and Systems, 8093 Zurich, Switzerland

1 Introduction

Real-time traffic management is the process of executing production plans, supervising the traffic situation, detecting conflicts, and executing dispatching actions to prevent or resolve conflicts. In recent years, several mathematical models of railway operations have been developed to automatically generate timetables or to determine dispatching actions. Infrastructure, rolling stock, and operations define the boundaries for traffic managers' activities and thus result in functional requirements that mathematical models of railway operations must respect so that they can be used in actual real-time traffic management.

Cordeau et al (1998); Törnquist (2006); Lusby et al (2011); Corman and Meng (2013); Cacchiani et al (2014) reviewed approaches for solving several tasks related to railway operations planning and management. Cordeau et al (1998) reviewed approaches for assembling, routing, and scheduling trains at all planning levels: strategic, tactical, and operational. In addition, also approaches for planning rolling stock circulations were reviewed. Törnquist (2006) covered tactical and operational scheduling and rescheduling approaches. Scheduling and rescheduling were distinguished as follows:

"Scheduling (or timetabling) is the process of constructing a schedule from scratch, while rescheduling (or dispatching) indicates that a schedule already exists and will be modified."

In this work, we adopt these definitions. Lusby et al (2011) reviewed strategic, tactical, and operational scheduling and rescheduling approaches. Corman and Meng (2013) and Cacchiani et al (2014) reviewed approaches for rescheduling, and they also covered aspects related to crew and rolling stock rescheduling. These previous reviews aimed at classifying optimization approaches with respect to the task they were designed to solve, the modelling and solution approaches, and the kind of experiments that was applied to assess their performance. This paper suggests a next step towards the development of a decision support system for real-time rescheduling by evaluating the existing mathematical models with respect to the actual functional requirements of railway operations. This analysis will facilitate the choice of appropriate models and their potential further development for implementation into a productive system.

The models are illustrated on the network of the *Railway operations laboratory (EBL)* at the ETH Zurich, which is shown in Figure 1. Two trains are assumed to travel through it. Macroscopically (Figure 2), the first train travels from Testadt (T) to Pewald (P) running through Ypslikon (Y) and Zetthausen (Z), and the second train travels from Utal (U) to Pewald (P) passing through Testadt (T), Ypslikon (Y), and Wedorf (W). In Ypslikon a passenger connection is expected. Station Ypslikon (Figure 3) is used as an instance of a microscopic topology. To make the modelling task more challenging, it is assumed that the first train reaches Ypslikon from signal A105 and the second one from A205.

The next section summarizes the functional requirements for real-time traffic management. Section 3 outlines the main classes of mathematical models that have been proposed for scheduling and rescheduling tasks. Section 4 analyses whether and

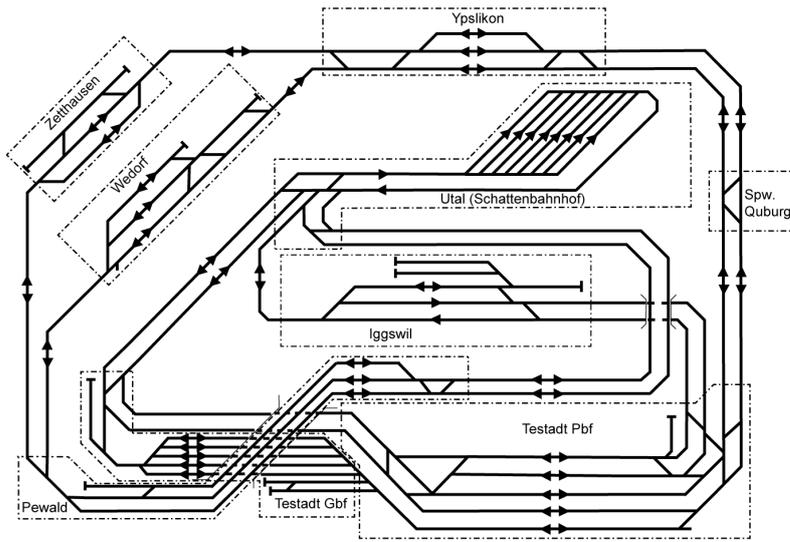


Fig. 1 Railway network of the *Railway operations laboratory (EBL)* at the ETH Zurich (archive of the institute for transport planning and systems (IVT))

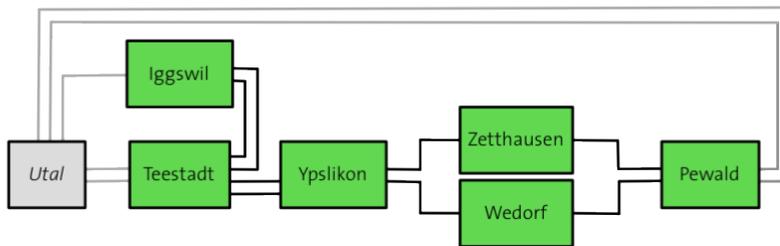


Fig. 2 Macroscopic topology of the network of the *Railway operations laboratory* at the ETH Zurich (IVT archive)

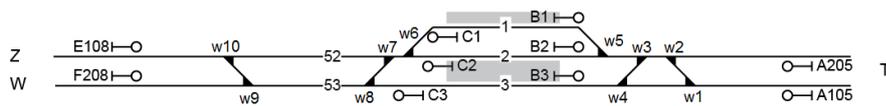


Fig. 3 Microscopic topology of station Ypslikon in the *Railway operations laboratory* at the ETH Zurich (IVT archive)

how the different mathematical models satisfy the functional requirements for real-time traffic management. In Section 5, conclusions and future work are presented.

2 Functional requirements from real-time traffic management

To reflect the actual functioning of railway operations and real-time traffic management, models for rescheduling railway traffic should satisfy a number of functional requirements. This section summarizes these functional requirements based on previous works and on the general features of railway operations. The requirements can be grouped in different categories depending on the type of resources they are linked to: (1) infrastructure; (2) rolling stock; (3) operations.

2.1 Infrastructure

Railway infrastructure prescribes the static physical boundaries of railway operations. Railway infrastructure networks are usually represented using graphs. The granularities of these representations differ considerably depending on the applications. Macroscopic infrastructure models (or macroscopic topologies) usually contain only stations, junctions, and links between them. Information on length of tracks, number of tracks, average running time, and capacity is usually associated with these macroscopic links. Detailed representations of railway infrastructure including information about permissible speeds, gradients, radii, signals, block sections, and release points are usually referred to as microscopic infrastructure models (or microscopic topologies). Models combining characteristics of microscopic and macroscopic topologies are called mesoscopic. Examples of mesoscopic topologies can be obtained by dropping some elements of microscopic models, or using different granularities for representing different areas.

According to Radtke (2008), conflict detection and resolution requires a microscopic infrastructure model. However, if major disturbances occur, it may be necessary to look for a solution in a large partition of the network. To perform this in a microscopic topology may be prohibitive because of the long computational times. Thus, this paper also considers approaches based on topologies with coarser granularities.

Note that, if the energy supply of a network is not uniform, also this element should be considered by the infrastructure model. As this paper aims at identifying suitable models for rescheduling railway traffic on the Swiss railway network, and this network is completely and uniformly electrified, this element is neglected in our analysis.

2.2 Rolling stock

Rolling stock prescribes the dynamic physical boundaries of railway operations. The maximum speed on a track is determined by acceleration and braking capabilities of

rolling stock combined with train length, mass, resistance, and infrastructure properties (curvature, gradient). By considering also the timetable, the feasible dynamics for each train on each track can be found (Brünger and Dahlhaus, 2008). Cordeau et al (1998) highlighted that speed can be represented either with fixed profiles or with variable ones. Considering fixed speed profiles reduces the number of variables, which reduces the size of the model but also limits the degrees of freedom for finding solutions. For instance, energy consumption cannot be optimized using fixed speed profiles because the relevant degrees of freedom are missing.

The length of trains also influences the occupation time of infrastructure resources and may limit overtaking and routing. In order to delimit the range of study of this article and to focus on operations related requirements, this issue has been addressed by another paper that has recently been submitted for publication and is neglected here.

2.3 Operations

Additional requirements are prescribed by safety rules, monitoring and intervention features, and operational interdependencies. Safe railway operations are currently ensured by fixed block signalling. Networks are partitioned into blocks that are delimited by signals. Blocks can host at most one train at a time, and some time is required for route setting and releasing. The sequence of block allocations on a train path is referred to as blocking time stairway. In a long term perspective, this signalling system will probably be replaced by moving block signalling. However, the migration process will last several decades, and this technology will most likely be limited to some corridors only for many years. Thus, this paper considers the current signalling system with fixed blocks.¹

During operations, trains should run as close as possible to the given timetable, and departure from a station before schedule is usually forbidden for passenger trains. Delays are deviations from the timetable. Within this paper, the notation proposed by Cacchiani et al (2014) is used. Disturbances indicate perturbations that can be solved by modifying the timetable, but leaving the duties of rolling stock and crew unchanged. Disruptions indicate perturbations that cannot be resolved by modifying only the timetable. This paper focuses on rescheduling strategies for disturbances. Thus, the following analysis neglects rolling stock and crew rostering.

Corman and Meng (2013) identify the following five actions considered by rescheduling:

- re-timing an event (e.g. the arrival at or the departure from a station);
- re-ordering trains on a shared infrastructure;
- local re-routing (e.g. platform change);
- global re-routing;
- re-servicing.

¹ If the signalling system of the reference network is not uniform, then the representations of infrastructure and rolling stock should include the corresponding pieces of information.

Re-servicing corresponds to invasive dispatching measures such as breaking connections, modifications of stopping patterns, turnarounds before destination, train cancellations and replacements. As most rescheduling models have not been conceived for representing re-servicing features, this paper analyses them only with respect to breaking connections and cancelling trains.

3 Mathematical models for railway operations

Several mathematical models for supporting different tasks of railway traffic planning and operations have been developed during the last decades. Different models have been developed to represent different properties of railway operations, depending on whether they would be applied for macroscopic timetabling, microscopic timetabling², or dispatching. This section presents and analyses the representatives of the main classes of formulations for railway traffic scheduling and rescheduling that are most prominent in the literature, and which were identified thanks to Cacciani et al (2014); Corman and Meng (2013); Törnquist (2006); Lusby et al (2011); Cordeau et al (1998). The presentation follows the outline of Baccelli et al (1992)'s "guidelines for modelling of dynamic systems":

1. describe the evolution of each resource in the system individually;
2. integrate the interactions among the resources;
3. tackle the problem of initialization.

In this framework, the *resources* are the train runs, which are fully defined by sequences of discrete events coinciding with arrivals and departures at relevant infrastructure points. First, the formulations modelling time as a continuous variable are presented. Then, the formulations relying on discretizations of time are introduced. All models are illustrated using the example presented in Section 1.

3.1 Continuous time formulations

Event Scheduling Problems (ESP), Alternative Graphs (AG), and FlexiblePath (FP) are the continuous time formulations that have been applied to railway planning and operations most broadly. The Periodic ESP (PESP) has been extensively used to produce periodic macroscopic timetables for given train routes (Caimi, 2009; Herrigel et al, 2013; Kroon et al, 2009; Peeters and Kroon, 2001; Serafini and Ukovich, 1989). AG has been primarily exploited to generate microscopic schedules both off- and on-line (Corman et al, 2010; D'Ariano et al, 2007a, 2008, 2014; Mascis and Pacciarelli, 2002; Mazzarello and Ottaviani, 2007). FP has been applied to find train routes and schedules in a microscopic topology in real-time (Mu and Dessouky, 2014; Yan and Yang, 2012).

Let z be a train and S be a topology node where discrete events take place. v_S^z denotes a discrete event that is related to z and takes place at S . In all three cases,

² A macroscopic(microscopic) timetable is a schedule on a macroscopic(microscopic) topology.

continuous variables t_v^z represent the times when the discrete events v_S^z occur. The relevant discrete events for ESP are the arrivals and departures at nodes of a macroscopic topology that coincide with stations where services begin or end or connections take place. In contrast, the discrete events of AG and FP are associated with nodes of microscopic topologies. AG considers not only arrivals and departures at all stations, but also passages at signals. The discrete events associated with FP are the entrances in and the exits from sections of infrastructure that correspond to tracks between junctions and can host at most one train at a time (refer to Lu et al (2004) for a complete description of the underlying network partitioning). Figure 4 shows the graph associated with ESP model of the sample macroscopic scheduling problem presented in Section 1 on the macroscopic topology of EBL (Figure 2). Figure 5 shows the sample microscopic scheduling problem in station Ypslikon (Figure 3). In both figures, the white nodes represent the discrete events. The black nodes indicate zero events, which are events take place at time zero independently from any other event.

Given the route of a train z , its run can be fully described as the sequence of discrete events (v_S^z) associated with the topology nodes S on its route. Zero events are considered for fixing time intervals, when necessary. The relations between the events are described by inequalities of the form

$$cType : t_{S_2}^z - t_{S_1}^z \geq f_{(S_1, S_2)}^z \quad (1)$$

These constraints fix the minimum time separation $f_{(S_1, S_2)}^z$ allowed between two events $v_{S_1}^z, v_{S_2}^z$, and they can describe:

- cRun: the minimum running time $f_{(S_1, S_2)}^z$ of train z from departure $v_{S_1}^z$ to arrival $v_{S_2}^z$;
- cRun: the maximum running time $-f_{(S_1, S_2)}^z$ of train z from departure $v_{S_2}^z$ to arrival $v_{S_1}^z$;
- cDwell: the minimum dwell time $f_{(S_1, S_2)}^z$ of train z from arrival $v_{S_1}^z$ to departure $v_{S_2}^z$;
- cDwell: the maximum dwell time $-f_{(S_1, S_2)}^z$ of train z from arrival $v_{S_2}^z$ to departure $v_{S_1}^z$;
- cPass: $v_{S_1}^z$ corresponds to a zero event, and $v_{S_2}^z$ cannot take place before $f_{(S_1, S_2)}^z$ (Mazzarello and Ottaviani (2007) refers to these constraints as passing constraints);
- cPass: $v_{S_2}^z$ corresponds to a zero event, and $v_{S_1}^z$ cannot take place after $-f_{(S_1, S_2)}^z$;
- cOverall: the minimum running time $f_{(S_1, S_2)}^z$ from the departure from the first station $v_{S_1}^z$ to the arrival at destination $v_{S_2}^z$ (usually defined for ESP only);
- cOverall: the maximum running time $-f_{(S_1, S_2)}^z$ from the departure from the first station $v_{S_2}^z$ to the arrival at destination $v_{S_1}^z$ (usually defined for ESP only);

Note that ESP usually contains for each minimum time constraint also the corresponding maximum time constraint. Therefore, the pairs are identified by *cRun*, *cDwell*, *cPass*, and *cOverall* and are associated with intervals $\left[\underline{f}_{(S_1, S_2)}^z, -\overline{f}_{(S_2, S_1)}^z \right]$:

$$cType : \underline{f}_{(S_1, S_2)}^z \leq (t_{S_2}^z - t_{S_1}^z) \leq -\overline{f}_{(S_2, S_1)}^z \quad (2)$$

Figures 4 and 5 show these relations for ESP and AG as solid grey lines respectively.

FP does not fix the routes in advance but includes a binary variable x_i^z for each train z and infrastructure node i indicating whether z passes i . O_z and D_z denote the origin and destination node for z . $\delta_-(J)$ and $\delta_+(J)$ denote sets of nodes connected with junction J from the two travel directions. As each train follows a unique continuous route, the following additional constraints must be satisfied:

$$\begin{cases} x_{O_z}^z = 1 \\ x_{D_z}^z = 1 \end{cases} \quad \forall z \quad (3)$$

$$\sum_{i \in \delta_-(J)} x_i^z = \sum_{v \in \delta_+(J)} x_v^z \quad \forall J, \forall z \quad (4)$$

D'Ariano et al (2014) extend the AG formulation for rescheduling to feature rerouting. This formulation includes a binary variable x_i^z for each train z and for each route i from the entrance of z in the considered area to its exit. As for FP, this variable denotes whether z follows route i . In this case, the route continuity constraint (4) is no longer needed, and the uniqueness constraint (3) becomes

$$\sum_i x_i^z = 1 \quad \forall z \quad (5)$$

In both cases, the constraints (1) become

$$cType : t_{S_2}^z - t_{S_1}^z + M(1 - x_i^z) \geq f_{(S_1, S_2)}^z \quad (6)$$

where M is an large enough constant.

Note that if a train stops unexpectedly, it cannot enter the successive section with maximum speed. D'Ariano et al (2007b) proposed an iterative approach in which AG is combined with a second step that updates the realisable speeds according to the solution found by AG.

The interactions between the different runs can be represented using the time variables t_S^z . In the classical ESP and AG formulations (i.e. no routing), the relation between event $v_{S_1}^z$ of train z and event $v_{S_2}^w$ of another train w is given by

$$cType : t_{S_2}^w - t_{S_1}^z \geq f_{(S_1, S_2)}^{z,w} \quad (7)$$

This can model:

- \underline{cConn} : minimum connection time $f_{(S_1, S_2)}^{z,w}$ from the arrival $v_{S_1}^z$ of z to the departure $v_{S_2}^w$ of the destination train w ;
- \overline{cConn} : maximum connection time $-f_{(S_1, S_2)}^{z,w}$ from the arrival $v_{S_2}^w$ of w to the departure $v_{S_1}^z$ of the destination train z ;
- $\underline{cDep}, \overline{cDep}$: minimum and maximum separation times $f_{(S_1, S_2)}^{z,w}, -f_{(S_2, S_1)}^{w,z}$ between the departures $v_{S_1}^z, v_{S_2}^w$ of trains z, w with similar services.

Couples of interactions within ESP can also be modelled using intervals (crf. (2)). Although the authors do not mention it, interactions could be modelled within FP and AG's extension by modifying (7) as follows:

$$cType : t_{S_2}^w - t_{S_1}^z + M(1 - x_i^z) + M(1 - x_j^w) \geq f_{(S_1, S_2)}^{z,w} \quad (8)$$

where i, j are the either infrastructure resources (FP) or routes (AG) connected with the discrete events $v_{S_1}^z$ and $v_{S_2}^w$ respectively. These relations are shown as solid orange lines in Figures 4 and 5.

To coordinate the movement of all trains and to measure the total time needed to bring all trains to destination, AG contains a "Start Node" v_O corresponding to a "start" command and an "End Node" v_N corresponding to a "finish" command. These additional nodes are represented as orange disks in Figure 5. A constraint, $cStart$, forces the first event $v_{S_1}^z$ of every train to take place after the start command. A constraint, $cExit$, imposes that the finish command takes place after the last event $v_{S_2}^z$ of every train run.

$$cStart : t_{S_1}^z - t_O \geq 0 \quad cExit : t_N - t_{S_2}^z \geq 0 \quad (9)$$

This for D'Ariano et al (2014)'s AG extension becomes

$$cStart : t_{S_1}^z - t_O + M(1 - x_i^z) \geq 0 \quad cExit : t_N - t_{S_2}^z + M(1 - x_i^z) \geq 0 \quad (10)$$

As the ordering of trains in conflict points is not usually known in advance, both possibilities should be modelled. PESP models headway constraints with a pair of constrains of type (7) on the events corresponding to the trains entering and leaving the common section. However, if periodicity is dropped, this yields to infeasibility. AG represents headway constraints as pairs of alternative arcs, which correspond to disjunctive constraints of the form

$$cHead : (t_{S_3}^w - t_{S_2}^z \geq f_{(S_1, S_2, S_3, S_4)}^{z,w}) \vee (t_{S_1}^z - t_{S_4}^w \geq f_{(S_1, S_2, S_3, S_4)}^{w,z}) \quad (11)$$

where $v_{S_1}^z$ corresponds to train z entering the common section; $v_{S_2}^z$ to z leaving it; $v_{S_3}^w$ to train w entering it; $v_{S_4}^w$ to w leaving it. Alternative arcs can also be represented using binary variables $h_i^{z,w}$ indicating whether train z passes a shared infrastructure i before w .

$$\begin{aligned} cHead : t_{S_3}^w - t_{S_2}^z + Mh_i^{z,w} &\geq f_{(S_1, S_2, S_3, S_4)}^{z,w} \\ t_{S_1}^z - t_{S_4}^w + M(1 - h_i^{z,w}) &\geq f_{(S_1, S_2, S_3, S_4)}^{w,z} \end{aligned} \quad (12)$$

Analogously, FP and AG extension model headway constraints as

$$\begin{aligned} cHead : t_{S_3}^w - t_{S_2}^z + Mh_i^{z,w} + M(1 - x_j^z) + M(1 - x_k^w) &\geq f_{(S_1, S_2, S_3, S_4)}^{z,w} \\ t_{S_1}^z - t_{S_4}^w + M(1 - h_i^{z,w}) + M(1 - x_j^z) + M(1 - x_k^w) &\geq f_{(S_1, S_2, S_3, S_4)}^{w,z} \end{aligned} \quad (13)$$

where $i = j = k$ in FP and j, k are routes containing i in AG. The dashed orange dashed lines in Figures 4 and 5 represent this second type of interactions. Corman et al (2012) uses alternative arcs for modelling connections. While connections represented using (7) are forced, the formulation via alternative arcs allows breaking connections. Disjunctive constraints (11) can also model out of Order constraints ($cOut$), which represent the closure of an infrastructure resource (Mazzarello and Otaviani, 2007). In this case, trains are forced to pass the closed section either before the closure or after the reopening time.

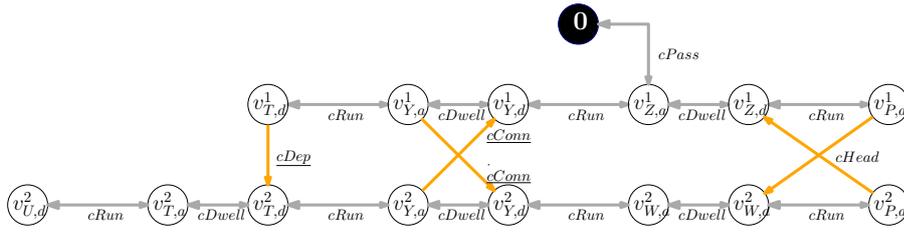


Fig. 4 Event Scheduling Problem for two train runs on the infrastructure of EBL (Figure 2). The white nodes represent the discrete events, and the black node represent a zero event. The grey right arrows coincide with constraints (1), the left-right arrows with (2), and the orange edges with (7)

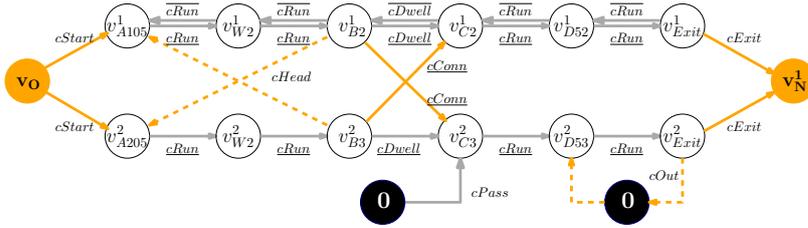


Fig. 5 Alternative Graph of two train runs in Ypslikon (see Figure 3). The white nodes correspond to the discrete events, the black nodes indicate zero events, and the orange ones the start node (v_0) and the end node (v_N). The solid grey lines represent constraints (1), the solid orange edges represent constraints (7), and the dashed ones represent different types of constraints (11)

While the scheduling versions of ESP, AG, and FP need no initialization, their rescheduling versions would need the initial positions of all the trains. Mazzarello and Ottaviani (2007) models the initial position of a train z in AG with a "position node" that is inserted into the train path. The position node is connected with v_0 through a constraint of type (1), where f_{S_1, S_2}^z denotes the current time. Due to their similarity with AG, one can imagine to apply the same initialization to the other models presented so far.

3.2 Discrete time formulations

Several approaches for scheduling and rescheduling that allow routing choices are based on discrete-time models. Using discrete time sets may be counter-intuitive but is a powerful method for exploiting degrees of freedom for (re)routing, while keeping the complexity limited. Packing problems are the the most prominent approaches in the literature. Their focus is finding conflict-free routes and passing times for all trains. This section presents the following models:

- Arc Packing Problem (APP) and its weak version (APP');
- Path Packing Problem (PPP);
- Arc Configuration Problem (ACP);
- Path Configuration Problem (PCP);
- Resource Tree Conflict Graph (RTCG) and Tree Conflict Graph (TCG);

- Resource Conflict Graph (RCG);
- REFormulated Simultaneous train Rerouting and Rescheduling (REF-SRR).

APP, ACP, PPP, and PCP have been used to route trains and produce aperiodic timetables (Caprara et al, 2002; Borndörfer and Schlechte, 2007; Fischer and Helmberg, 2010). RTCG has been applied to allocate blocks in microscopic timetabling (Caimi, 2009; Caimi et al, 2011). RCG has been used by Caimi (2009) for microscopic scheduling, and a version called Static Train Dispatching has been used by Fuchsberger (2012) for microscopic rescheduling. REF-SRR has been proposed by Meng and Zhou (2014) for microscopic rescheduling.

As for continuous time models, the discrete events are the arrivals at and departures from relevant infrastructure points. While the stations are the relevant infrastructure points for APP, ACP, PPP, and PCP, the endpoints of the single resources of the infrastructure are the relevant infrastructure points for RTCG, TCG, RCG, and REF-SRR. Each event can take place at several times. APP, APP', ACP, PPP, PCP, and REF-SRR contain a node for each time and place that can host such an event. RCTG, TCG, and RCG contain a tree indicating the possible routings on the infrastructure resources for each time when a train can enter the controlled area or start a run from a platform. The left hand side of Figure 6 shows the time-space grid for the macroscopic scheduling problem presented in section 1 in APP, APP', ACP, PPP, and PCP. The left hand side of Figure 7 shows the routing trees in station Ypslikon for the two trains of the sample microscopic scheduling problem mentioned in Section 1. In addition, both graphs contain an artificial source node s_z and an artificial sink node t_z for each train z . These nodes are depicted as black disks in both figures.

APP, ACP, RTCG, and TCG include a binary variable x_a^z for each action a separating two events that can be consecutive for train z . Each variable indicates whether the corresponding action is scheduled or not. These actions coincide with directed edges and can be:

- *aStart*: arcs connecting s_z to all nodes corresponding to the first station of train z ;
- *aEnd*: arcs connecting all nodes corresponding to the last station of z to t_z ;
- *aRun*: arcs connecting the specific time and place where a run starts to the end time and place (i.e. these arcs define the route choice and the running time);
- *aDwell*: dwells in stations from specific arrival times to specific departure times;
- *aInfeasibility*: arcs connecting a node that is not the last on the path of z to t_z .

Figures 6 and 7 show some of these actions for the sample problems mentioned above. For each node v , let $\delta_+(v)$ and $\delta_-(v)$ be the outgoing and ingoing arcs respectively. As each train can be scheduled at most once, and the path should be continuous, the following constraints must be satisfied.

$$\sum_{a \in \delta_+(s_z)} x_a^z \leq 1 \quad \forall z \quad (14)$$

$$\sum_{a \in \delta_+(v)} x_a^z - \sum_{a \in \delta_-(v)} x_a^z = 0 \quad \forall v \notin \{s_z, t_z\}, \forall z \quad (15)$$

For APP, (15) ensures route continuity, as well as time continuity. As the the stopping pattern influences the speed that is actually realisable, Fischer and Helmberg (2010)

extend ACP to consider both the rolling stock and the stopping pattern for defining the realisable speed on a section. PPP, PCP, and RCG include a unique binary variable x_p^z for an entire chain of such actions from the beginning of a train path to the end. The left hand side of Figure 6 depicts these variables as blue and red paths; the left hand side of Figure 7 shows these variables as blue and red nodes. In this case, continuity constraints (15) are implicitly assumed in the choice of the paths. The variables only have to satisfy the equivalent of (14):

$$\sum_p x_p^z \leq 1 \quad \forall z \quad (16)$$

While the paths in Fuchsberger (2012)'s RCG formulation are associated with maximal speed profiles, Caimi (2009) proposed a RCG formulation for scheduling in regions with low traffic density which considers different speed profiles.

Interactions between trains in APP, APP', ACP, PPP, PCP, RTCG, and TCG correspond to allocation conflicts. If a line between two stations that can host at most one train at a time or a single infrastructure resource is considered, then a conflict occurs if the occupation intervals of different trains overlap. If x_a^z and x_b^w correspond to conflicting allocations of an infrastructure resource r , APP' and TCG prevent the simultaneous allocation through

$$x_a^z + x_b^w \leq 1 \quad (17)$$

APP, PPP, RTCG, and RCG model conflicting allocations using conflict graphs. For a given infrastructure resource r , a conflict graph contains a node for each edge connected with that resource. Two nodes are linked together by an edge if either they correspond to the same train, or their allocation intervals of the resource overlap (see e.g. Herrmann (2006) for further information about conflict graphs). Let \mathcal{C}_r be the set of maximal cliques of the conflict graph on the infrastructure resource r . Then, the following constraints prevent conflicts:

$$(APP, APP', RTCG) \sum_{(z,a) \in C} x_a^z \leq 1 \quad \forall C \in \mathcal{C}_r, r \quad (18)$$

$$(PPP, RCG) \sum_{(z,p) \cap C \neq \emptyset} x_p^z \leq 1 \quad \forall C \in \mathcal{C}_r, r \quad (19)$$

The right hand side of Figure 6 shows the conflict graph for the sample macroscopic scheduling problem presented in Section 1, assuming that both trains use the same track between Testadt and Ypslikon. The nodes correspond to the edges between Testadt and Ypslikon, i.e. the edges in the orange box on the left hand side. The red and blue edges of the conflict graph connect all nodes corresponding to the one train, and the black edges coincide with conflicts. The right hand side of Figure 7 shows the conflict graph for resource $w8$ (i.e. in the orange box of the left hand side) in the sample microscopic scheduling problem presented in Section 1.

Fischer et al (2008) include capacity constraints in PPP, indicating how many trains can be hosted and how many edges can be assigned to each direction. Fuchsberger (2012)'s RCG model splits all train routes in one arriving to the station and one leaving it. Scheduled connections are used to identify incompatible stairways, i.e. pairs of an inbound and an outbound blocking time stairways of trains with a scheduled connection but without enough time to perform it. For every such pair, a variable indicating whether the connection has to be forced during operations is defined.

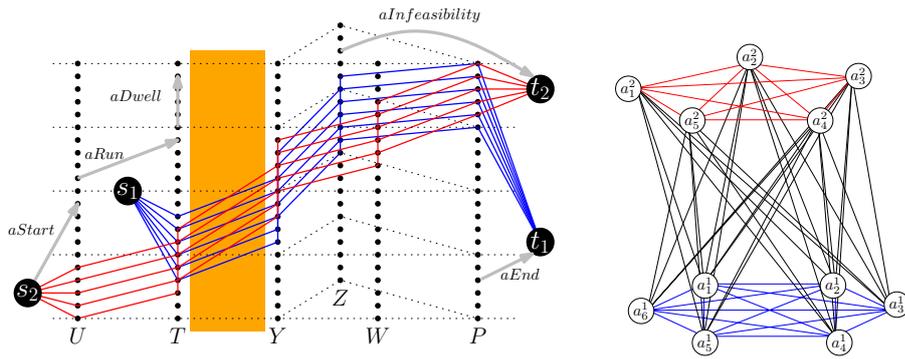


Fig. 6 APP/ACP arcs and PPP/PCP model for the sample macroscopic scheduling problem on the infrastructure in Figure 2. On the left hand side, the small nodes correspond to possible departure and arrival times at stations; the large nodes represent sources and sinks; the grey arcs represent the different types of constraints for APP/ACP; the blue paths correspond to possible runs for the first train in PPP/PCP; the red paths to the possible runs of the second train in PPP/PCP. The right hand side shows the conflict graph between T and Y (i.e. the orange box on the left hand side)

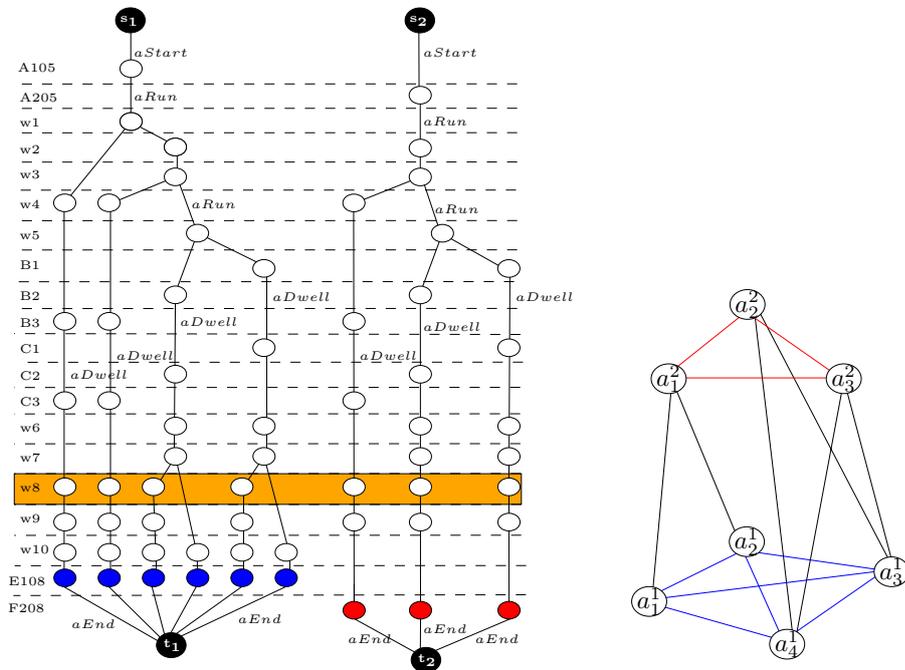


Fig. 7 RTCG/TCG tree and RCG model. On the left hand side, the black nodes indicate the sources and sinks; the white nodes correspond to the route choice points in the topology of Ypslikon (Figure 3); the arcs correspond to the binary variables of RTCG and TCG model; the blue nodes correspond to the possible paths of the first train in RCG; the red nodes correspond to the possible paths for the second train in RCG. The right hand side shows the conflict graph connected with resource w8. The nodes correspond to the nodes in the orange box on the left hand side. Red and blue edges connect all nodes associated with one train, and black edges indicate overlapping blocking time intervals

ACP and PCP are extensions of APP and PPP respectively. Instead of using conflict graphs directly, ACP and PCP use them to identify configurations. A configuration q is a set of trips on a track r that are not conflicting with each other. Thus, it satisfies

$$|q \cap C| \leq 1 \quad \forall C \in \mathcal{C}_r \quad (20)$$

A configuration q for a track r is a path from an artificial source node s_r to an artificial sink node t_r . The edges on the path either correspond to runs on the track or connect s_j to departure nodes, t_j to arrival nodes, or an arrival to possible departures of follow-up runs. For each such edge e , let y_e be a binary variable, indicating whether the edge is part of the configuration. $\delta_+(v)$ and $\delta_-(v)$ denote the outgoing and ingoing arcs for each node v of this graph. A conflict-free schedule is obtained by choosing a unique continuous path in this graph and limiting the choice of actions a in APP to the ones that are contained in the configuration path. This corresponds to substituting (18) with the following constraints, and the obtained model is ACP:

$$\sum_{e \in \delta_+(v)} y_e - \sum_{e \in \delta_-(v)} y_e = 0 \quad \forall e \notin \{s_r, t_r\}, \forall r \quad (21)$$

$$\sum_{e \in \delta_+(s_r)} y_e \leq 1 \quad \forall r \quad (22)$$

$$x_a - y_a \leq 1 \quad \forall a \quad (23)$$

For each configuration q , PCP includes a unique binary variable y_q indicating whether q is assigned in the schedule. Thus, continuity constraints (21) are no longer needed. Let Q_r be the set of configurations on track r . PCP is obtained by substituting (19) with the following constraints:

$$\sum_{q \in Q_r} y_q \leq 1 \quad \forall r \quad (24)$$

$$\sum_{a \in p} x_p - \sum_{a \in q} y_q \leq 0 \quad \forall a \quad (25)$$

REF-SRR contains binary variables $x_{i,j}^z$ indicating whether train z passes infrastructure nodes i and j sequentially; $a_{i,j,t}^z$ indicating whether train z arrives at j from i at time t ; $d_{i,j,t}^z$ indicating whether train z departs from i towards j at time t ; $y_{i,j,t}^z$ indicating whether train z is between i and j at time t . The relations between time variables $a_{i,j,t}^z, d_{i,j,t}^z$ correspond to running and dwell times and are modelled similarly as in continuous time formulations. Route continuity and uniqueness are ensured by constraints (14) and (15) in the variables $x_{i,j,t}^z$. As the decision variables correspond to the arrival and departure times and the route choices, this model includes constraints modelling the time-space interdependencies for all trains. Meng and Zhou (2014) also includes passing constraints (cfr. discrete time formulations) for preventing departures ahead of schedule. Conflict-free operations are modelled as clique constraints (18) in $y_{i,j,t}^z$ for each time t and track (i, j) .

As for continuous time formulations, while scheduling approaches need no initialization, Fuchsberger (2012)'s and Meng and Zhou (2014)'s rescheduling approaches need the initial position of the trains. Fuchsberger (2012) uses a rolling time horizon, which fixes the blocking time stairways of train movements that have already started

or are starting too early in the future and uses predictions for the entrance times of trains that are approaching the controlled area.

4 Results

Thanks to the presentation of the previous section, two main categories of mathematical models of railway operations can be distinguished. The formulations in the first category model time with **continuous** variables and safety with minimal time differences between **pairs** of consecutive trains on each infrastructure resource. This category includes ESP, AG and FP. The second category contains models based on **discrete** time variables, modelling conflicts as **cliques** of a conflict graph. APP, ACP, RTCG, PPP, PCP, RCG, and SRR are in this second category. While variables in APP, ACP, and RTCG represent single train actions, variables in PPP, PCP, and RCG model entire paths, and variables in SRR correspond to route choices and to events that can only take place at a discrete set of points in time. Some models cannot be classified in any of these categories because they rely on **discrete** representations of time and model conflicts **pairwise**. Examples of such models are APP' and TCG (i.e. the weak versions of APP and RCTG) . Table 1 shows the results of the analysis of the models from Section 3 with respect to the functional requirements listed in Section 2.

Table 1 Models and functional requirements: × means that the model satisfies the functional requirement; (×) means that the requirement is satisfied off-line (i.e. for timetabling); ◦ means that there is a model extension which satisfies the requirement

		continuous time			discrete time						
		ESP	AG	FP	APP'	TCG	APP/ ACP	RTCG	REF- SRR	PPP/ PCP	RCG
infrastructure	macroscopic	×			×		×			×	
	microscopic		×	×		×		×	×		×
rolling stock	max. speed	×	×	×	×	×	×	×	×	×	×
	real. speed		◦				◦			×	×
operations	timetable		×	×					×		×
	closed tracks		×	×					×		×
	re-timing	(×)	×	×	(×)	(×)	(×)	(×)	×	(×)	×
	re-ordering	(×)	×	×	(×)	(×)	(×)	(×)	×	(×)	×
	re-routing		◦	×	(×)	(×)	(×)	(×)	×	(×)	×
	connections		◦								×
	cancel train		◦	×	(×)	(×)	(×)	(×)	×	(×)	×
		pairwise conflicts					on tracks		on paths		
							conflict cliques				

4.1 Infrastructure

ESP, APP, ACP, APP', PPP, and PCP are based on a macroscopic topology, which is not suitable for modelling safety constraints of fixed block signalling systems precisely. Still, for each of these macroscopic models, there is a microscopic model that considers time and conflicts analogously. ESP, AG, and FP model time with continuous variables and require a minimum time separation between pairs of trains using the same infrastructure resource. APP' and TCG model discrete time choices and require that at most one allocation from each pair of conflicting allocations of an infrastructure resource is assigned. The other models limit time choices to discrete sets and model headway constraints using conflict graphs. APP, ACP, RTCG, and REF-SRR describe single activities (run, dwell), and PPP, PCP, and RCG describe entire paths.

4.2 Rolling Stock

All models are able to describe train dynamics considering maximum speed profiles. For timetabling the stopping patterns are usually known in advance. Thus, approaches modelling maximum speed profiles (that consider the planned stops) suffice to generate feasible schedules. If stopping patterns are not fixed in advance, it is necessary to consider the effect of stopping on the minimum feasible running time. As PPP, PCP, and RCG consider entire train paths, the actual running times are computed during the preprocessing phase, after having chosen which paths to include into the model. In the other models, the dependence of running times on stops has to be modelled within the optimization problem. Fischer and Helmberg (2010)'s ACP extension provides the actually realisable speed profiles as functions of rolling stock and stopping pattern. Note that this is the only model including variables that model explicitly whether a train stops at an infrastructure point or just passes it. In any case, if a train has to stop or it runs with limited speed on a track, it is not possible to enter the next contiguous track at maximum speed. D'Ariano et al (2007b) proposed an iterative approach in combining an AG with a second step updating the realisable speeds according to the solution found by the AG. RTCG computes conflicts using the maximal speed profiles allowed. If it cannot find a conflict free trajectory for each train, then the train is cancelled.

4.3 Operations

Models conceived for timetabling rather than dispatching do not usually include representations of operations related features. In addition, operations related features are represented differently by the models that have already been applied to real-time operations (i.e. AG, FP, RCG, and REF-SRR).

First, the constraints given by the planned timetable may be included as a set of passing constraints (i.e. inequalities imposing the departure times from stations to be greater to or equal to the scheduled times), as part of the objective function (i.e. terms that penalize early and late departures as well as too early or late arrivals and which

usually depend on the amount of delay), or by considering them in the preprocessing phase when choosing the paths.

Second, FP, RCG, and REF-SRR represent routing possibilities by binary variables that can be forced to zero to represent closed tracks. The classical AG contains no such variable and does not support re-routing. If a track is closed, trains are forced to pass it either before its closure or after its reopening time. D'Ariano et al (2014)'s AG extension contains binary variables for routing possibilities. Thus, closed tracks can be represented as by FP.

Third, all models presented include representations of either intervention features or their off-line counterpart. Formulations for timetabling support timing, ordering, routing, and/or scheduling/not scheduling trains, which are the off-line counterparts of re-timing, re-ordering, re-routing, and cancellation of trains. (Re-)timing is performed in models with continuous time variables by changing the value of time variables. In models considering discrete time choices, (re-)timing corresponds to a different choice for time-space combination. ESP, AG, and FP formulate (re-)ordering options as pairs of alternative arcs. The other formulations contain no such arc, and re-ordering is performed implicitly by choosing times and routes. All formulations based upon discrete times feature (re-)routing. Also D'Ariano et al (2014)' AG extension and FP feature re-routing. Breaking connections is featured only by RCG and Corman et al (2012)'s AG extension. ESP and the classical AG formulation aim at finding suitable times for previously selected discrete events. Thus, if a train cannot be scheduled because it conflicts with other services, the solution may either not exist or schedule the train after all other services. In contrast, train cancellation is represented by all other models as not scheduling the train. This coincides with all routing or path variables taking value zero. Note that Meng and Zhou (2014)'s REF-SRR formulation imposes that each train is scheduled exactly once, but this constraint can be relaxed to model that each train is scheduled at most once.

5 Conclusions and future work

The main differences between models are the representations of conflicts (pairwise or as cliques) and time (continuous or discrete). As expected, the formulations that have been developed for off-line scheduling do not satisfy all the functional requirements. In particular, they do not include any representation of delays and closed tracks, and they only consider off-line versions of dispatching actions. In contrast, the four formulations that have been proposed for real-time operations satisfy most requirements related to infrastructure and operations. In particular, known versions of AG and RCG satisfy all these requirements and also the requirements related to rolling stock. Although being different from each other, the presented models show some similarities. Thus, some models can be considered as microscopic versions of macroscopic models or extensions of models that have less features. Due to their similarities with microscopic rescheduling models, the macroscopic scheduling models presented in this work (ESP, APP, ACP, PPP, PCP) may be extended to satisfy the missing requirements. Doing so, it would be possible to consider traffic management at different hierarchical levels, depending on the size of perturbations. Conversely,

due to its similarity with PPP, it may be possible to extend RCG to a sort of microscopic PCP and apply column generation techniques to be able to deal with large instances as reported by Borndörfer and Schlechte (2007).

Basing on the analysis presented in this work, the next steps are to give real-time traffic management an appropriate mathematical representation exploiting the similarity of RCG to PPP, to find an appropriate solving strategy and, finally, to implement a decision support tool for dispatchers.

References

- Baccelli FL, Cohen G, Olsder GJ, Quadrat JP (1992) Synchronization and Linearity: An Algebra for Discrete Event Systems. Wiley series in probability and mathematical statistics, John Wiley and Sons, Chichester, England
- Borndörfer R, Schlechte T (2007) 05. Models for Railway Track Allocation. In: Liebchen C, Ahuja RK, Mesa JA (eds) 7th Workshop on Algorithmic Approaches for Transportation Modeling, Optimization, and Systems (ATMOS'07), Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Dagstuhl, Germany, OpenAccess Series in Informatics (OASICs), vol 7, DOI 10.4230/OASICs.ATMOS.2007.1170
- Brünger O, Dahlhaus E (2008) Running time estimation. In: Hansen IA, Pahl J (eds) Railway Timetabling and Operations, 2nd edn, Eurail press, Hamburg, Germany, chap 4, pp 65–89
- Cacchiani V, Huisman D, Kidd M, Kroon L, Toth P, Veelenturf L, Wagenaar J (2014) An overview of recovery models and algorithms for real-time railway rescheduling. *Transportation Research Part B: Methodological* 63(0):15 – 37, DOI 10.1016/j.trb.2014.01.009
- Caimi G, Chudak F, Fuchsberger M, Laumanns M, Zenklusen R (2011) A new resource-constrained multicommodity flow model for conflict-free train routing and scheduling. *Transportation Science* 45(2):212–227, DOI 10.1287/trsc.1100.0349
- Caimi GC (2009) Algorithmic decision support for train scheduling in a large and highly utilised railway network. Dissertation no. 18581, ETH Zurich
- Caprara A, Fischetti M, Toth P (2002) Modeling and solving the train timetabling problem. *Operations Research* 50(5):851–861, DOI 10.1287/opre.50.5.851.362
- Cordeau JF, Toth P, Vigo D (1998) A survey of optimization models for train routing and scheduling. *Transportation Science* 32(4):380–404
- Corman F, Meng L (2013) A review of online dynamic models and algorithms for railway traffic control. In: Intelligent Rail Transportation (ICIRT), 2013 IEEE International Conference on, pp 128–133, DOI 10.1109/ICIRT.2013.6696281
- Corman F, D'Ariano A, Pacciarelli D, Pranzo M (2010) A tabu search algorithm for rerouting trains during rail operations. *Transportation Research Part B: Methodological* 44(1):175 – 192, DOI 10.1016/j.trb.2009.05.004
- Corman F, D'Ariano A, Pacciarelli D, Pranzo M (2012) Bi-objective conflict detection and resolution in railway traffic management. *Transportation Research Part C: Emerging Technologies* 20(1):79 – 94, DOI 10.1016/j.trc.2010.09.009, special issue on Optimization in Public Transport+ISTT2011 Special issue on Optimiza-

-
- tion in Public Transport+International Symposium on Transportation and Traffic Theory (ISTTT), Berkeley, California, July 1820, 2011
- D'Ariano A, Pacciarelli D, Pranzo M (2007a) A branch and bound algorithm for scheduling trains in a railway network. *European Journal of Operational Research* 183(2):643 – 657, DOI 10.1016/j.ejor.2006.10.034
- D'Ariano A, Pranzo M, Hansen I (2007b) Conflict resolution and train speed coordination for solving real-time timetable perturbations. *Intelligent Transportation Systems, IEEE Transactions on* 8(2):208–222, DOI 10.1109/TITS.2006.888605
- D'Ariano A, Corman F, Pacciarelli D, Pranzo M (2008) Reordering and local rerouting strategies to manage train traffic in real time. *Transportation Science* 42(4):405–419, DOI 10.1287/trsc.1080.0247
- D'Ariano A, Samà M, D'Ariano P, Pacciarelli D (2014) Evaluating the applicability of advanced techniques for practical real-time train scheduling. *Transportation Research Procedia* 3(0):279 – 288, DOI 10.1016/j.trpro.2014.10.007, 17th Meeting of the {EURO} Working Group on Transportation, EWGT2014, 2-4 July 2014, Sevilla, Spain
- Fischer F, Helmberg C (2010) Dynamic Graph Generation and Dynamic Rolling Horizon Techniques in Large Scale Train Timetabling. In: Erlebach T, Lübbecke M (eds) 10th Workshop on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems (ATMOS'10), Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Dagstuhl, Germany, OpenAccess Series in Informatics (OASICs), vol 14, pp 45–60, DOI 10.4230/OASICs.ATMOS.2010.45
- Fischer F, Helmberg C, Janßen J, Krostitz B (2008) Towards Solving Very Large Scale Train Timetabling Problems by Lagrangian Relaxation. In: Fischetti M, Widmayer P (eds) 8th Workshop on Algorithmic Approaches for Transportation Modeling, Optimization, and Systems (ATMOS'08), Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Dagstuhl, Germany, OpenAccess Series in Informatics (OASICs), vol 9, DOI 10.4230/OASICs.ATMOS.2008.1585
- Fuchsberger M (2012) Algorithms for railway traffic management in complex central station areas. Dissertation no. 20398, ETH Zurich
- Herrigel S, Laumanns M, Szabo J, Weidmann U (2013) Periodic railway timetabling with hierarchical decomposition in the pesp model. In: *Proceedings of RailCompenhagen*
- Herrmann TM (2006) Stability of timetables and train routings through station regions. Dissertation no. 16398, ETH Zürich
- Kroon L, Huisman D, Abbink E, Fioole PJ, Fischetti M, Maròti G, Schrijver A, Steenbeek A, Ybema R (2009) The new dutch timetable: The or revolution. *Interfaces* 39(1):6–17, DOI 10.1287/inte.1080.0409
- Lu Q, Dessouky M, Leachman R (2004) Modeling train movements through complex rail networks. *ACM Transactions on Modeling and Computer Simulation* 14(1):48–75
- Lusby RM, Larsen J, Ehrgott M, Ryan D (2011) Railway track allocation: models and methods. *OR Spectrum* 33(4):843–883, DOI 10.1007/s00291-009-0189-0
- Mascis A, Pacciarelli D (2002) Job-shop scheduling with blocking and no-wait constraints. *European Journal of Operational Research* 143(3):498 – 517, DOI 10.1016/S0377-2217(01)00338-1

-
- Mazzarello M, Ottaviani E (2007) A traffic management system for real-time traffic optimisation in railways. *Transportation Research Part B: Methodological* 41(2):246 – 274, DOI <http://dx.doi.org/10.1016/j.trb.2006.02.005>
- Meng L, Zhou X (2014) 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(0):208 – 234, DOI 10.1016/j.trb.2014.05.005
- Mu S, Dessouky MM (2014) Control rules for dispatching trains on general networks with multiple train speeds. In: *Conference Proceedings Transport Research Arena 2014*, Paris, France
- Peeters L, Kroon L (2001) A cycle based optimization model for the cyclic railway timetabling problem. In: Voss S, Daduna JR (eds) *Computer-Aided Scheduling of Public Transport, Lecture Notes in Economics and Mathematical Systems*, vol 505, Springer Berlin Heidelberg, pp 275–296, DOI 10.1007/978-3-642-56423-9-16
- Radtke A (2008) Infrastructure modelling. In: Hansen IA, Pachel J (eds) *Railway Timetabling and Operations*, 2nd edn, Eurail press, Hamburg, Germany, chap 3, pp 47–63
- Serafini P, Ukovich W (1989) A mathematical model for periodic scheduling problems. *SIAM J Disc Math* 2(4):550 – 581
- Törnquist J (2006) Computer-based decision support for railway traffic scheduling and dispatching: A review of models and algorithms. In: Kroon LG, Möhring RH (eds) *5th Workshop on Algorithmic Methods and Models for Optimization of Railways (ATMOS'05)*, Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Dagstuhl, Germany, OpenAccess Series in Informatics (OASICS), vol 2, DOI 10.4230/OASICS.ATMOS.2005.659
- Yan C, Yang L (2012) Mixed-integer programming based approaches for the movement planner problem: Model, heuristics and decomposition. Tech. rep., URL <http://www.informs.org/Community/RAS/Problem-Solving-Competition/2012-RAS-Problem-Solving-Competition>, 2012 RAS Problem Solving Competition Report