Optimal train routing and scheduling for managing traffic perturbations in complex junctions

Paola Pellegrini, Grégory Marliere, Joaquin Rodriguez

Transportation Research Part B: Methodological, Volume 59, January 2014, p58–80, © 2013 Elsevier, http://dx.doi.org/10.1016/j.trb.2013.10.013

Optimal train routing and scheduling for managing traffic perturbations in complex junctions

Paola Pellegrini, Grégory Marlière
and Joaquin Rodriguez

Ifsttar – ESTAS
Univ. Lille Nord de France
rue Élisée Reclus 20, 59666 Villeneuve d'Ascq, Lille, France
paola.pellegrini@ifsttar.fr, gregory.marliere@ifsttar.fr,
joaquin.rodriguez@ifsttar.fr

Abstract

Real-time traffic management in railway aims to minimize delays after an unexpected event perturbs the operations. It can be formalized as the real-time railway traffic management problem, which seeks for the best train routing and scheduling in case of perturbation, in a given time horizon. We propose a mixed-integer linear programming formulation for tackling this problem, representing the infrastructure with fine granularity. This is seldom done in the literature, unless stringent artificial constraints are imposed for reducing the size of the search space. In a thorough experimental analysis, we assess the impact of the granularity of the representation of the infrastructure on the optimal solution. We tackle randomly generated instances representing traffic in the control area named triangle of Gagny, and instances obtained from the real timetable of the control area including the Lille-Flandres station (both in France) and we consider multiple perturbation scenarios. In these experiments, the negative impact of a rough granularity on the delay suffered by trains is remarkable and statistically significant.

Keywords: real-time railway traffic management problem, mixed-integer linear programming, track-circuit, complex junction, routing, scheduling

1 Introduction

Railway timetable is designed so that trains can be operated on the available infrastructure without the emergence of *conflicts*, where a conflict is represented by multiple trains concurrently claiming a portion of track. At peak hours, the capacity of the infrastructure is often fully exploited. When an unexpected event perturbs the operations, conflicts emerge and trains must be delayed for sequencing their use of the critical portions of track. This sequencing is particularly important at *junctions*, where different

Transportation Research Part B: Methodological, Volume 59, January 2014, p58–80, © 2013 Elsevier, http://dx.doi.org/10.1016/j.trb.2013.10.013

lines cross and share portions of track, and at stations, which are particular junctions in which trains may stop for loading and unloading purpose.

Traffic on the railway network is managed by *dispatchers*. They are in charge of smoothing operations in their *control areas*. If a control area includes a complex junction, the dispatcher task may become very challenging. At junctions, dispatchers may intervene on both train routes (*routing*) and train ordering along routes (*scheduling*) for minimizing delay propagation. In fact, for each pair of origin and destination (*o-d pair*), often multiple *routes* exist. Along these routes, trains can be stopped at selected location for scheduling purposes. Currently, few decision support tools are available for rerouting or rescheduling trains at junctions. The available tools, as for example the ARI system used in the Netherlands, may just reserve routes to trains on the basis of the timetable scheduling and on arrival time forecasts. Despite the undeniable aid of these tools, dispatchers must often take decisions autonomously (D'Ariano et al., 2008).

Several authors have proposed optimization algorithms for tackling the problem faced by dispatchers. We will refer to the formal problem tackled as the real-time railway traffic management problem (rtRTMP). In the literature, different variants of the rtRTMP have been tackled. These variants can be classified according to three criteria: the possibility of changing train routes, the granularity of the representation of the control area and the consideration of speed variation dynamics.

The limitation to the level of exploitation of capacity imposed by the impossibility of changing train routes is straightforward. D'Ariano et al. (2008) and Corman et al. (2010) empirically assess the large impact of this limitation, even when train routes are chosen through a heuristic.

The granularity of the representation of the control area is inversely proportional to the size of the track portions into which routes are split. As detailed in the following (Section 3), if the infrastructure is represented with fine granularity, then train routes are split into track-circuits and the route-lock sectional-release interlocking system can be modeled; if it is represented with rough granularity, then train routes are split into block sections (or even longer portions of track) and only a route-lock route-release interlocking can be modeled. Intuitively, as granularity becomes finer, the precision of the analysis, and hence the suitability of the decisions made, may increase as well. A first step in the assessment of the impact of granularity on the solution quality has been proposed by Corman et al. (2009b).

The consideration of speed variation dynamics is computationally costly and, then, hardly possible in real-time. However, the greater realism of models considering these dynamics (*variable-speed*) compared to the ones neglecting them (*fixed-speed*) is undeniable. Rodriguez (2007) reports some experimental evidence on the comparison between the solutions returned by an algorithm when considering either a variable or a fixed-speed model. The solutions of the two models are assessed in simulation. According to Rodriguez (2007), the increase of realism achieved through the variable-speed model does not pay, due to the huge additional computational effort for calculating speed variation dynamics and to the consequent limited effort available for the search space exploration. The general validity of these results has not been proven, but the number of algorithms considering fixed-speed models present in the literature, compared to the number of those considering variable-speed ones, shows that the community is quite unanimous on this validity: even if introducing a strong hypothesis, fixed-speed models are able to both

Transportation Research Part B: Methodological, Volume 59, January 2014, p58–80, © 2013 Elsevier, http://dx.doi.org/10.1016/j.trb.2013.10.013

capture many critical elements characterizing reality and supply solutions with a high practical relevance.

In this paper, we propose a mixed-integer linear programming (MILP) formulation for the rtRTMP. Besides considering constraints that model the relevant issues which emerge in the practice of railway traffic management, it finds the optimal solution to the rtRTMP allowing train routing along all possible routes existing in the control area. It solves instances which represent the infrastructure with fine granularity: through this formulation, we can apply either the route-lock route-release interlocking system, or the route-lock sectional-release one. Moreover, the formulation can incorporate constraints which allow its utilization in a rolling-horizon framework: we may consider time elapse in the optimization process, by solving instances representing subsequent time intervals and by considering, during each solution, the impact of the decisions previously made. In this research step, we do not consider speed variation dynamics.

We test the MILP formulation on two types of instances: random instances representing traffic in the triangle of Gagny, and perturbations of real instances representing traffic in the control area including the Lille-Flandres station, both in France. In the experimental analysis, we assess the improvement of the solution quality which we can achieve by fining the granularity of the representation of the control area, i.e., by moving from modeling a route-lock route-release interlocking system to modeling a route-lock sectional-release one. In particular, we show that granularity may have a remarkable impact on the quality of the optimal solution. We assess this impact in a thorough analysis, considering different scenarios and two objective functions measuring either total or maximum delay.

The rest of the paper is organized as follows. Section 2 reports the most relevant contributions on the solution algorithms for the rtRTMP, while Sections 3 and 4 depict the infrastructure on which this problem must be modeled and the main characteristics of the problem itself, respectively. Section 5 describes the formulation proposed in this paper. Section 6 presents the experimental setup and the instances tackled and Section 7 shows the results of the computational analysis. Section 8 reports an additional computational analysis aimed to assess the practical applicability of the formulation proposed. Finally, Section 9 concludes the paper.

2 Literature review

Several solution approaches have been proposed for tackling the rtRTMP. In this section we report the most relevant contributions, grouping them according to the granularity of the representation of the control area considered, as shortly presented in the Introduction and further detailed in Section 3.

Most of the contributions proposed in the literature represent the infrastructure with the rough granularity, i.e., splitting tracks into block sections. In some cases, the infrastructure is even represented in terms of *track segments* grouping sequences of block sections on which the train ordering cannot be changed. This is often done for modeling very large control areas. In this framework, Dessouky et al. (2006) describe a branch-and-bound algorithm for rescheduling trains using fixed routes. Törnquist and Persson (2007) propose a mixed-integer linear programming formulation allowing the change of

Transportation Research Part B: Methodological, Volume 59, January 2014, p58–80, © 2013 Elsevier, http://dx.doi.org/10.1016/j.trb.2013.10.013

both train scheduling and routing in a control area representing the whole South Traffic District of Sweden. However, this formulation is not strong enough for allowing the solution of the instances tackled in the paper, which include approximately 200 track segments. Hence, the authors propose strategies for reducing the feasible region of the instances and obtaining tractable sub-instances, strategies which are improved by Törnquist (2007). Törnquist Krasemann (2012) proposes a greedy heuristic for the rtRTMP, modeling track segments. All these solution approaches deal with fixed-speed models.

Several papers model the rtRTMP as a scheduling problem through an alternative graph: trains correspond to jobs and block sections correspond to machines. Mazzarello and Ottaviani (2007) propose a three step heuristic algorithm: the first step of the procedure solves a scheduling problem in which only train order can be modified; the second step deals with the route modification option; finally, the third step reassesses the train schedule with the new routes selected. D'Ariano et al. (2007a) propose a branch-and-bound algorithm for solving the scheduling problem with fixed routes. It is combined with a tabu search in charge of changing train routes by D'Ariano et al. (2008). This algorithm is improved by Corman et al. (2010) and it is used by D'Ariano and Pranzo (2009) for dealing with the rtRTMP considering a long time horizon split in sub-periods. Corman et al. (2009a) apply the algorithm by D'Ariano et al. (2007a) and two other existing heuristics for assessing the potential of the green wave policy. This policy consists in possibly increasing stop times at station for having trains always running at their scheduled speed between stations. Corman et al. (2012a) tackle a bi-objective variant of the rtRTMP using algorithms based on the branch-and-bound by D'Ariano et al. (2007a). Most of these papers consider fixed-speed models. The others consider an approximated computation of speed variation dynamics. The basis for this computation is proposed by D'Ariano et al. (2007b): the authors present a heuristic algorithm for the rtRTMP with fixed routes. It is based on an iterative two step procedure: first the algorithm by D'Ariano et al. (2007a) seeks for a solution with the fixed speed model; then a check is made for verifying the actual feasibility of the solution found in case of consideration of speed variation dynamics. If the solution is unfeasible, the time to be spent by a specific train on the first block section where a conflict emerges is enlarged for allowing braking, and the solution is re-optimized from there on through the fixed-speed algorithm. The experimental analyses proposed in these papers always consider the route-lock route-release interlocking system, and in the description of the model the infrastructure is represented with rough granularity. However, a sophistication of the model based on alternative graphs can be used to consider the route-lock sectional-release interlocking system. Corman et al. (2009b) describe how this sophistication might be implemented. The authors also propose a comparison between the route-lock route-release interlocking system and the sectional-lock sectional-release one. This comparison is made in terms of delay propagation and computation time for achieving the optimal solution. The conclusion is that the merits of the approximation of the sectional-lock sectional-release interlocking system in terms of reduction of delay propagation are not sufficient to justify the increase of computation time with respect to the route-lock route-release one. This conclusion is considered to be extensible to the route-lock sectional-release interlocking system. The analysis is carried out without considering the possibility of changing train routes. The introduction of this possibility may change the result of the comparison proposed: the difference between the two interlocking systems emerges only when two

Transportation Research Part B: Methodological, Volume 59, January 2014, p58–80, © 2013 Elsevier, http://dx.doi.org/10.1016/j.trb.2013.10.013

trains travel along only partially coincident routes. Hence, by changing train routes, the optimization may allow to exploit the route-lock sectional-release interlocking system to a higher extent.

Other papers consider the finest granularity for the representation of the control area, i.e., they represent the infrastructure in terms of track-circuits. Rodriguez (2007) proposes two constraint programming algorithms, considering a fixed-speed and a variable-speed model, respectively. The latter does not consider proper speed variation dynamics, but it constraints traveling times to be coherent with train braking and acceleration in case of conflict. Caimi et al. (2011) and Caimi et al. (2012) describe two exact algorithms. The authors limit the size of the feasible region by considering only some possible scheduling and routing options for each train. The number of routes considered is increased progressively as far as the feasible region is empty. The algorithms do not explicitly tackle either a fixed or variable-speed model: no delay along the route can be assigned to trains. The issue of speed variation dynamics is left to the preceding control area by, in case, imposing that a train enters later than it could in the considered one. Lusby et al. (2012) propose a heuristic algorithm considering a variable-speed model. This heuristic is based on a set packing formulation and it considers a discretized time horizon. For having the tractability of realistic size instances, the discretization step is of 15 seconds. The authors recognize that this actually represents a constraint which artificially limits the control area capacity exploitation, but they consider this limit acceptable. The algorithm takes as input the available *paths* for each train, in terms of vectors of binary pairs of track-circuit (possibly sequence of track-circuits) and time interval. Including the consideration of variable train speed, several paths must be taken into account for each route, even if the impact of the speed dynamics may not be very strong due to the time discretization. Initially, only one path is considered for each train. If no feasible solution exists, paths are added by either imposing delays or adding available train routes.

3 The infrastructure of a control area

A control area is a part of the railway network managed by one dispatcher. Its infrastructure contains portions of track on which the presence of a train is detected by an electrical device. These portions are called *track-circuits*. Sequences of track-circuits are typically grouped into *block sections*, which start and end with a light signal. Several block sections can share track-circuits, for example in presence of a switch: two block sections may have the same entry signal, share the first track-circuits, split through the switch and have different exit signals. A sequence of block sections connecting an o-d pair in the control area composes a *route*. The origin and the destination of a route can be either platforms, if the control area includes one or more stations, or track-circuits at the border with limitrophe control areas.

The aspect of the light signal at the beginning of a block section imposes the behavior to be held by the driver facing it: proceeding at the scheduled speed (green aspect), braking for being able to stop by the following signal (yellow aspect), or stop (red aspect). Different signaling systems exist, typically with different numbers of aspects: three being the most common configuration, further aspects (*restrictive aspects*) may separate the

Transportation Research Part B: Methodological, Volume 59, January 2014, p58–80, © 2013 Elsevier, http://dx.doi.org/10.1016/j.trb.2013.10.013

green and the red ones, with a consequently larger number of block sections available for braking. In general terms, if the signaling system has n possible aspects, then $n-2$ block sections are available for braking.

When a train enters the first track-circuit of a block section, all the following ones which belong to the same block section are *reserved* for the train itself. Until the train ends the *utilization* of a track-circuit, where the utilization includes both reservation and physical *occupation*, no other train can utilize it.

The minimum headway between trains is defined through *blocking times*. These blocking times include the approach time, which is the running time between a signal that provides a potentially restrictive aspect (i.e., the first signal with aspect different from the green one) and the following signal. (Pachl, 2002) When applying this concept (*blocking time theory*) in modeling, this translates into a reservation of a track-circuit starting as soon as a train occupies the first track-circuit of the $n-2^{nd}$ preceding block section. In a three-aspect system, for example, the reservation of a track-circuit starts before the train enters the preceding block section. In addition, before a train enters a block section, some time must be allowed for route formation and for taking into account the signal visibility distance (U.I.C., 2004). In the following we will refer to the sum of these times simply as *formation time*. After a train exits a block section, some time must elapse before the track-circuits of the block section itself can be utilized by another train. In this time, the route is released (*release time*) (U.I.C., 2004).

As mentioned in both the Introduction and Section 2, this infrastructure can be modeled with different granularities. In particular, in the literature two main possibilities have been considered for modeling control areas including complex junctions. The infrastructure can be represented in terms of either block sections or track-circuits:

- **block sections:** the only available information is on trains being or not in block sections; no detail exist on train exact positions within couples of signals. The only possible interlocking system is the route-lock route-release one (Theeg et al., 2009). Here, the utilization of track-circuit *tc* locks block section *bs* including *tc* along the route traveled, and all block sections sharing with *bs* a track-circuit, even different from *tc*;

- **track-circuits:** train locations are known in terms of track-circuits. Both the route-lock route-release and the route-lock sectional-release interlocking systems (Theeg et al., 2009) are possible. In the latter, the utilization of *tc* locks only block sections including *tc* itself.

The modeling in terms of block sections may be implemented in two ways: either the track-circuits are actually known by the model and the moment in which a track-circuit reservation ends is properly managed; or only block sections are actually known by the model (all track-circuits in a block section are merged into a unique entity) and the compatibility between the concurrent utilization of pairs of block sections is ensured.

Figure 1 depicts an example of the infrastructure of a simple control area. Track circuits are named tc and signals are named s, both indexed with a progressive number. Signals concern the availability of block sections in a precise direction: for example, signal s2 concerns block section s2-s4 including tc1, tc2 and tc3, in this order, and block section s2-s5 including tc1, tc2 and tc6, in this order.

Transportation Research Part B: Methodological, Volume 59, January 2014, p58–80, © 2013 Elsevier, http://dx.doi.org/10.1016/j.trb.2013.10.013

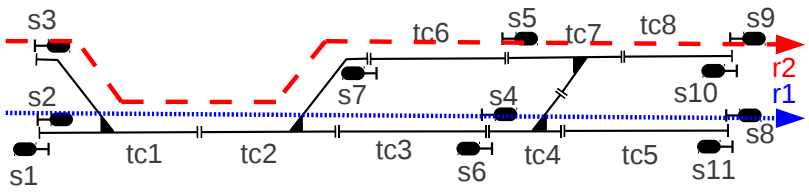

Figure 1: Example of the infrastructure present in a control area.

Suppose that two trains (t1 and t2) cross the control area: t1 going from s2 to s8 (using route r1 including block sections s2-s4 and s4-s8) and t2 going from s3 to s9 (using route r2 including block sections s3-s5 and s5-s9). Train t1 may arrive in the control area at second 70 at the earliest, and train t2 at second 75; t1 enters the preceding block section, in the limitrophe control area, earlier than t2, and t2 needs 30 seconds to cross it. Both trains need 30 seconds for completely traversing each track-circuit, and 10 seconds for clearing it. We set 15 and 5 seconds as formation and release time, respectively, we consider a three-aspect signaling system and we implement the blocking time theory.

By representing the infrastructure using any of the two different granularities which have been proposed in the literature, t1 is allowed to cross the control area first, since it enters the preceding block section before t2. At that time, block section s2-s4 is available and t1 can reserve it 15 seconds before entering the preceding block section itself (formation time). Moreover, 15 seconds before its earliest possible occupation of this block section, the following one, namely s4-s8, is available too. Hence, it can start the occupation as soon as possible, i.e., at second 70, and it reserves s4-s8 from second 55. It runs on each of the five track-circuits for 30 seconds, and it clears them 10 seconds after the head of the train entering the following one: it physically occupies tc1 between 70 and 110, tc2 between 100 and 140, tc3 between 130 and 170, tc4 between 160 and 200 and tc5 between 190 and 230. Figure 2 depicts the space-time diagram of this journey. The end of the utilization of track-circuits, and hence their availability to train t2, depends on the interlocking system applied:

- **route-lock route-release** (Figure 2, top): t1 utilizes all track-circuits of s2-s4 between second 55 minus the running time on the preceding block section and second 170, and all track-circuits of s4-s5 between second 55 and second 230. After it exits each block section, it still holds its utilization for the 5 seconds of the release time. Hence, t1 actually unlocks the first block section as a whole at second 175. Only at this point, t2 can start its reservation of tc1 and tc2, which belong to the first block section along its route, together with tc6. Due to the three-aspect signaling system, it is 15 seconds after this moment that t2 can enter the preceding block section, where it has to spend 30 seconds: it will start occupying tc1 at second 220, remaining there until second 260; it will occupy tc2 between 250 and 290, tc6 between 280 and 320, tc7 between 310 and 350 and tc8 between 340 and 380. The utilization of s2-s5 ends at second 295 and the one of s5-s9 at second 385.

- **route-lock sectional-release** (Figure 2, bottom): each track-circuit is unlocked

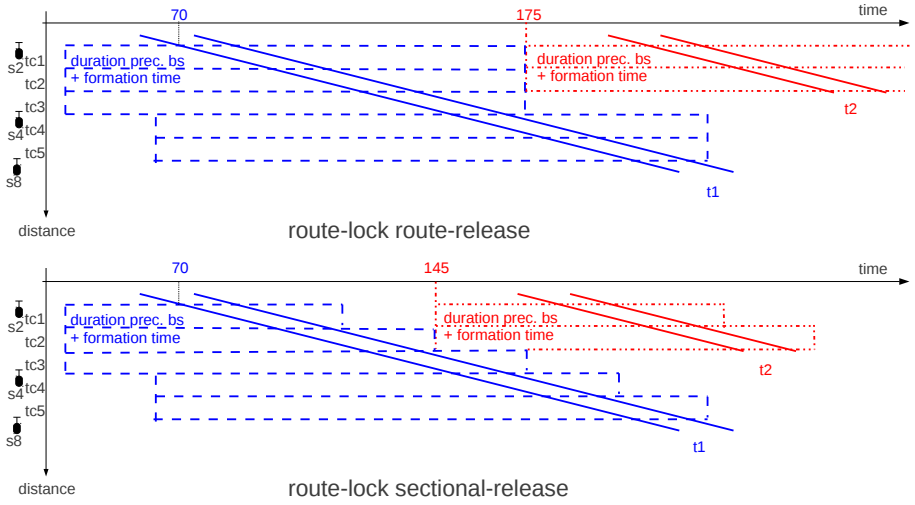

Figure 2: Space-time diagram of the journey of two trains in the control area shown in Figure 1: representation of track-circuits along route s2-s8. The top figure corresponds to the block section modeling. The bottom figure corresponds to the track-circuit one.

5 seconds after t1 has exit it, hence at second 115 for tc1, 145 for tc2, 175 for tc3, 205 for tc4 and 235 for tc5. The reservation of block section s3-s5 by t2 can then start at second 160: t2 spends second 160 to 190 in the preceding block section, and then occupies track-circuits along route r2 in the following intervals: between 190 and 230 for tc1, between 220 and 260 for tc2, between 250 and 290 for tc6, between 280 and 320 for tc7 and between 310 and 350 for tc8. The utilization of each track-circuit ends 5 seconds after the corresponding end of occupation.

Even if this example is trivial, it shows how the different interlocking systems, and hence different granularities of the representation, imply a different duration of the utilization of the infrastructure: with finer granularity, we can detect the actual utilization needs, and this may permit the better exploitation of the control area capacity.

4 The real-time railway traffic management problem

When an unexpected event occurs, trains may suffer a delay at their arrival in the control area. This delay is typically named *primary delay* and it may cause the emergence of conflicts within the control area itself. The additional delay due to these conflicts is named *secondary delay*. The goal of dispatchers is the minimization of this secondary delay.

According to the literature, various objective functions may be considered for the rtRTMP. In this paper, we consider the two most common ones: first, we minimize

Transportation Research Part B: Methodological, Volume 59, January 2014, p58–80, © 2013 Elsevier, http://dx.doi.org/10.1016/j.trb.2013.10.013

the maximum secondary delay suffered by any train (D'Ariano et al., 2007a); second, we minimize the total secondary delay (Rodriguez, 2007). Multiple sets of constraints characterize the problem.

First of all, **time concerning constraints**: a train cannot be operated earlier than its arrival in the control area (or than its scheduled departure from a platform), and it must occupy each track-circuit along one route for a certain amount of time. In variable-speed models, this time depends on traffic conditions. In fixed-speed models, it is computed *a priori* as the *running time* in absence of conflicts. The time in which the train occupies consecutive track-circuits concurrently is named *clearing time*: its rear is still on the current track-circuit and its front has already entered the following one. If the control area includes a station and trains with passenger transfers are scheduled (*trains in connection*), then their arrival and departure time must be coherent.

Second, some **constraints for managing delays** may be imposed. Three cases are possible: no constraints, delay allowed at any signal and delay allowed only out of the control area. In absence of constraints, delay may be assigned anywhere in the control area: the underlying hypothesis is that the dispatcher can delay the train in any track-circuit along the route. The case of delay allowed at any signal represents the fact that trains stop in front of signals. If delay can be assigned only out of the control area, no difference exists between fixed and variable-speed models, but complex coordination issues between control areas may emerge. In this paper, we allow the assignment of delay at any signal.

Third, **constraints due to the change of rolling stock configuration** may have to be imposed. In particular, the arrival and departure time of trains resulting from the turnaround, join or split of one another must be coherent.

Fourth, **capacity constraints** require that at most one train uses a block section at a time. When using a fixed-speed model, the blocking time theory described in Section 3 is typically implemented: all track-circuits belonging to a block section must be reserved for a train before it enters the first track-circuit of the $n - 2^{nd}$ preceding block section (where n is the number of aspects in the signaling system), minus the formation time. If representing the infrastructure with rough granularity, the reservation ends shortly after the train has exit the last track-circuit of the block section, due to the release time. If using fine granularity, it ends shortly after the train has exit the track-circuit itself.

5 Mixed-integer linear programming formulation

The mixed-integer linear programming formulation which we propose in this paper tackles the rtRTMP when the infrastructure is represented with fine granularity, i.e., in terms of track-circuits. As we show in this section, it can model either the route-lock sectional-release interlocking system, or the route-lock route-release one. In the following, we will refer to the formulation modeling the route-lock sectional-release interlocking system as the *SR formulation*, and to the one modeling the route-lock route-release one as the *RR formulation*. We proposed a preliminary version of this formulation in Pellegrini et al. (2012, 2013).

In addition to the existing track-circuits in the control area, we introduce two dummy ones: tc_0 and tc_∞. They represent the entry and the exit locations of the control area,

Transportation Research Part B: Methodological, Volume 59, January 2014, p58–80, © 2013 Elsevier, http://dx.doi.org/10.1016/j.trb.2013.10.013

respectively. The former precedes all track-circuits corresponding to route origins and the latter follows all track-circuits corresponding to route destinations. Their running time is null.

We consider routes which do not include any stop within their starting and ending point. If a train is actually scheduled to stop in an intermediate point of the infrastructure, we consider this train as two different ones using the same rolling stock.

For describing the SR formulation, we use the following notation:

T, R, TC	set of trains, routes and track-circuits, respectively,
ty_t	type corresponding to train t (indicating characteristics as weight, length, engine power, etc.),
$PL \subset TC$	set of track-circuits corresponding to platforms (if the control area includes a station),
$e(tc, r)$	indicator function: 1 if track-circuit tc belongs to an extreme (either the first or the last) block section on route r, 0 otherwise,
$R_t \subseteq R$	set of routes that can be used by train t,
TC^r	set of track-circuits composing route r,
$TC_t \subseteq TC$	set of track-circuits that can be used by train t ($TC_t = \bigcup_{r \in R_t} TC^r$),
$bs_{r,tc}$	block section including track-circuit tc along route r,
$p_{r,tc}, s_{r,tc}$	track-circuits preceding and following tc along route r, respectively,
$ref_{r,tc}$	reference track-circuit for the reservation of tc along route r: first track-circuit of the $n - 2^{nd}$ block section preceding $bs_{r,tc}$, with n number of aspects characterizing the signaling system,
$TC(tc, tc', r)$	set of track-circuits between tc and tc' along route r, tc and tc' included,
$rt_{ty,r,tc}, ct_{ty,r,tc}$	running and clearing time of track-circuit tc along route r for a train of type ty, respectively,
for_{bs}, rel_{bs}	formation and release time for block section bs, respectively,
$init_t$	earliest time at which train t can be operated: either expected arrival in the control area or expected departure from a platform within the control area,
$sched_t$	earliest time at which train t can reach its destination given $init_t$ and the route assigned to t in the timetable,
$i(t, t')$	indicator function: 1 if trains t and t' use the same rolling stock and t' results from the turnaround, join or split of train t, 0 otherwise,
$c(t, t')$	indicator function: 1 if trains t and t' are in connection, 0 otherwise,
$ms_{t,t'}, ms^c_{t,t'}$	minimum separation between the arrival of a train t and the departure of another train t' using the same rolling stock, or being in connection, respectively,
M	large constant.

5.1 Decision variables

We define both continuous and binary decision variables.

First of all, we define **continuous variables**, all non-negative:

$$D = \text{maximum secondary delay suffered by any train;}$$

Transportation Research Part B: Methodological, Volume 59, January 2014, p58–80, © 2013 Elsevier, http://dx.doi.org/10.1016/j.trb.2013.10.013

for all trains $t \in T$:

$$D_t = \text{secondary delay suffered by train } t \text{ when completing its route;}$$

for all triplets of train $t \in T$, route $r \in R_t$ and track-circuit $tc \in TC^r$:

$$o_{t,r,tc} = \text{time in which } t \text{ starts the occupation of } tc \text{ along route } r,$$

$$d_{t,r,tc} = \text{delay suffered by } t \text{ in } tc \text{ along route } r \text{ (defined if } bs_{r,tc} \neq bs_{r,s_{tc,r}}\text{);}$$

for all pairs of train $t \in T$ and track-circuit $tc \in TC_t$:

$$sU_{t,tc} = \text{time in which } tc \text{ starts being utilized by } t,$$
$$eU_{t,tc} = \text{time in which } tc \text{ ends being utilized by } t.$$

Remark that, for trains making intermediate stops within the infrastructure, since we consider their routes split, considering variables D_t as the secondary delay computed at the end of t's route corresponds to considering the secondary delay at each intermediate stop of the original route, if any.

Moreover, we define **binary variables**:
for all pairs of train $t \in T$ and route $r \in R_t$:

$$x_{t,r} = \begin{cases} 1 & \text{if } t \text{ uses } r, \\ 0 & \text{otherwise;} \end{cases}$$

for all triplets of train $t, t' \in T$ and track-circuit $tc \in TC_t \cap TC_{t'}$:

$$y_{t,t',tc} = \begin{cases} 1 & \text{if } t \text{ utilizes } tc \text{ before } t' \ (t \prec t'), \\ 0 & \text{otherwise } (t \succ t'). \end{cases}$$

Figure 3 shows the role of these variables in a single track-circuit and in a portion of the example depicted in Figure 1, corresponding to trains t1 and t2 along route r1. We depict track-circuit occupation as a rectangle with solid borders and reservation as a rectangle with dashed borders: the horizontal dimension represents time. The sum of reservation and occupation time corresponds to utilization. The reservation of a track-circuit starts for_{bs} time units before the occupation of the first track-circuit in the preceding block section bs (we represent the case of a three-aspect signaling system: $n - 2 = 1$), and it ends when the train starts the occupation of the track-circuit itself. A further reservation time follows the occupation and it lasts as long as the release time imposes it (rel_{bs}). Each track-circuit is occupied for a running time rt plus a clearing time ct: they depend on the track-circuit, on the route on which it is used and on the type of train using it. Moreover, a delay d may be assigned to the train in the last track-circuit of the block sections (tc3 and tc5). For track-circuits in block section s4-s8, the reference one ref along route r1 is tc1.

5.2 Route-lock sectional-release (SR) formulation

As anticipated in Section 4, we consider two alternative objective functions for the rtRTMP:

Transportation Research Part B: Methodological, Volume 59, January 2014, p58–80, © 2013 Elsevier, http://dx.doi.org/10.1016/j.trb.2013.10.013

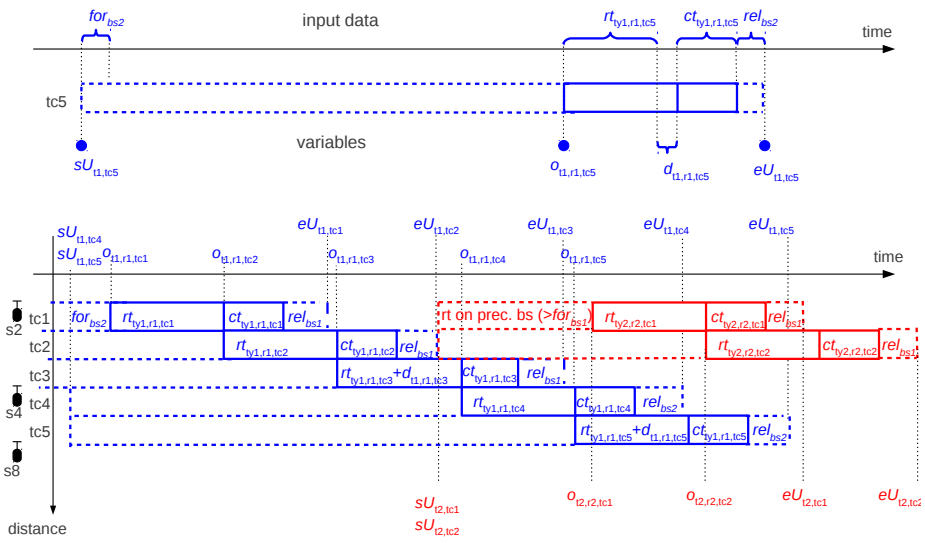

Figure 3: Graphical representation of data and variables through a Gantt diagram. Representation of the utilization (reservation + occupation + formation and release times) of track-circuit tc5 by train t1, and of both train journeys depicted in the space-time diagram in the bottom part of Figure 1.

1. the minimization of the maximum secondary delay suffered by any train

$$\min D, \tag{1}$$

2. the minimization of the total secondary delay suffered by trains

$$\min \sum_{t \in T} D_t. \tag{2}$$

Time concerning constraints

$$o_{t,r,tc} \geqslant init_t \, x_{t,r} \forall t \in T, r \in R_t, tc \in TC^r, \tag{3}$$

$$o_{t,r,tc} \leqslant M x_{t,r} \, \forall t \in T, r \in R_t, tc \in TC^r, \tag{4}$$

$$o_{t,r,tc} \geqslant o_{t,r,p_{r,tc}} + rt_{r,ty_t,p_{r,tc}} x_{t,r} \, \forall t \in T, r \in R_t, tc \in TC^r, \tag{5}$$

$$\sum_{r \in R_t} x_{t,r} = 1 \, \forall t \in T, \tag{6}$$

$$\sum_{\substack{r \in R_{t'}, \\ tc \in TC^r: \\ p_{r,tc} = tc_0}} o_{t',r,tc} \geqslant \sum_{\substack{r \in R_t, \\ tc \in TC^r: \\ s_{r,tc} = tc_\infty}} o_{t,r,tc} + (ms^c_{t,t'} + rt_{r,ty_t,tc}) x_{t,r} \, \forall t, t' \in T : c(t,t') = 1, \tag{7}$$

Transportation Research Part B: Methodological, Volume 59, January 2014, p58–80, © 2013 Elsevier, http://dx.doi.org/10.1016/j.trb.2013.10.013

$$D_t \geqslant \sum_{r \in R_t} o_{t,r,tc_\infty} - sched_t \; \forall t \in T \tag{8}$$

$$D \geqslant D_t \; \forall t \in T. \tag{9}$$

Constraints (3) state that trains cannot be operated earlier than $init_t$. Constraints (4) impose that the start of the occupation of all track-circuits along a route is set to zero if the route itself is not used. Constraints (5) state that a train cannot start occupying track-circuit tc along a route if it has not spent in the preceding track-circuit at least its running time, if the route is used. Constraints (6) ensure that exactly one route is used by each train. If the control area includes a station and trains in connection are scheduled, then we must impose Constraints (7). They state that a minimum separation of duration $ms^c_{t,t'}$ must be ensured between t's arrival and t''s departure, if they are in connection. The spatial coherence is ensured by the routes available for the trains: the destination ($s_{r,tc} = tc_\infty$) of any route available for the arriving train corresponds to a platform, as well as the origin ($p_{r,tc} = tc_0$) of any route available for the departing one. Constraints (8) and (9) impose the coherence of variables D and D_t, respectively. If the objective function (2) is considered, neither variable D nor Constraints (9) need to be set.

Constraints for managing delay

$$d_{t,r,tc} = o_{t,r,s_{r,tc}} - o_{t,r,tc} - rt_{r,ty_t,tc} x_{t,r} \; \forall t \in T, r \in R_t, \; tc \in TC^r : bs_{r,tc} \neq bs_{r,s_{r,tc}}. \tag{10}$$

For each track-circuit tc which can be used by train t and which closes its block section, the delay variable $d_{t,r,tc}$ assumes value equal to the moment in which train t starts occupying the track-circuit which follows tc, minus the moment in which it starts occupying tc itself, minus the running time $rt_{r,ty_t,tc}$: Constraints (10) ensure these relations. Remark that the last track-circuit of a block section is not necessarily long enough for hosting the whole train. If this is not the case, also the preceding track-circuit will be occupied longer if the train is delayed. In our experimental analysis this event seldom occurred (most of the last track-circuit of block sections are long enough to contain a train), but in principle it is an approximation which we introduced in the model.

Constraints due to the change of rolling stock configuration

$$\sum_{\substack{r \in R_t, \\ tc \in TC^r: \\ p_{r,tc} = tc_0}} o_{t,r,tc} \geqslant \sum_{\substack{r \in R_{t'}, \\ tc \in TC^r: \\ s_{r,tc} = tc_\infty}} o_{t',r,tc} + (ms_{t,t'} + rt_{r,ty_{t'},tc}) x_{t',r} \; \forall t, t' \in T : i(t',t) = 1, \tag{11}$$

$$\sum_{\substack{r \in R_t, \\ tc \in TC^r: \\ p_{r,tc} = tc_0}} sU_{t,tc} \leqslant \sum_{\substack{r \in R_{t'}, \\ tc \in TC^r: \\ s_{r,tc} = tc_\infty}} eU_{t',tc} \; \forall t, t' \in T : i(t',t) = 1, \tag{12}$$

$$\sum_{r \in R_t : tc \in TC^r} x_{t,r} = \sum_{r \in R_{t'} : tc \in TC^r} x_{t',r} \; \forall t, t' \in T : i(t',t) = 1, tc \in PL. \tag{13}$$

Transportation Research Part B: Methodological, Volume 59, January 2014, p58–80, © 2013 Elsevier, http://dx.doi.org/10.1016/j.trb.2013.10.013

Similarly to Constraints (7), Constraints (11) state that a minimum separation of duration $ms_{t',t}$ must be ensured between t''s arrival and t's departure, if t results from t''s turnaround, join or split. Constraints (12) ensure that the track-circuit where the turnaround, join or split takes place is reserved by t' until it arrives at the platform, plus the release time, and then it is immediately reserved by t. We must impose an inequality for allowing joins: the reservation of the resulting train starts with the ending of the reservation of the first train arriving. For example, let t result from the join of trains t' and t'', and let t' be the first train arriving at the platform ($eU_{t',tc} < eU_{t'',tc}$). By setting the inequality constraints, we ensure that train t reserves the platform itself immediately after the reservation by t': the platform cannot be utilized by any train other than t'' until t's departure. The consequent variable values will be such that $sU_{t,tc} < eU_{t'',tc}$: the constraints must necessarily be inequality ones. We manage the capacity issues arising with two trains reserving concurrently the same track-circuit (t and t'' in the example) as described in the following. Besides the train temporal coherence, we must ensure local coherence: trains using the same rolling stock must use routes including the same platform. Constraints (13) guarantee this local coherence. Of course, if the routes available for the two trains share only one platform, i.e., if we are not allowed to impose platform modifications, these constraints will be trivially met. Otherwise, routes with different o-d pairs will belong to R_t: a unique origin and multiple destinations will exist for trains arriving at a platform, and multiple origins and a unique destination will exist for trains departing from a platform.

Capacity constraints

$$sU_{t,tc} = \sum_{r \in R_t : tc \in TC^r} \left(o_{t,r,ref_{r,tc}} - for_{bs}\, x_{t,r} \right) \quad \forall t \in T,\ tc \in TC_t : (\nexists\, t' \in T : i(t',t) = 1) \vee (\forall\, r \in R_t :\ ref_{r,tc} \neq s_{r,tc_0}), \tag{14}$$

$$sU_{t,tc} \leqslant \sum_{r \in R_t : tc \in TC^r} \left(o_{t,r,ref_{r,tc}} - for_{bs}\, x_{t,r} \right) \quad \forall t \in T,\ tc \in TC_t : (\exists\, t' \in T : i(t',t) = 1), (\exists\, r \in R_t :\ ref_{r,tc} = s_{r,tc_0}), \tag{15}$$

$$eU_{t,tc} = \sum_{\substack{r \in R_t : \\ tc \in TC^r}} o_{t,r,ref_{r,tc}} + ul_{t,r,tc} \quad \forall t \in T,\ tc \in TC_t, \tag{16}$$

$$y_{t,t',tc} + y_{t',t,tc} = 1 \quad \forall t, t' \in T,\ tc \in TC_t \cap TC_{t'}, \tag{17}$$

$$eU_{t,tc} - M(1 - y_{t,t',tc}) \leqslant sU_{t',tc} \quad \forall t, t' \in T : tc \in TC_t \cap TC_{t'} : i(t,t') \sum_{r \in R_t} e(tc,r) = 0 \wedge i(t',t) \sum_{r \in R_{t'}} e(tc,r) = 0 \tag{18}$$

$$eU_{t',tc} - M y_{t,t',tc} \leqslant sU_{t,tc} \quad \forall t, t' \in T : tc \in TC_t \cap TC_{t'} : i(t,t') \sum_{r \in R_t} e(tc,r) = 0 \wedge \tag{19}$$

Transportation Research Part B: Methodological, Volume 59, January 2014, p58–80, © 2013 Elsevier, http://dx.doi.org/10.1016/j.trb.2013.10.013

$$i(t',t)\sum_{r\in R_{t'}} e(tc,r) = 0.$$

Constraints (14) state that a train's utilization of a track-circuit starts as soon as the train starts occupying the track-circuit $ref_{r,tc}$ along one of the routes including it, minus the formation time. According to Constraints (15), if we are considering a track-circuit of the first $n-2$ block sections ($ref_{r,tc} = s_{r,tc_0}$) and the concerned train t results from the turnaround, join or split of one or more other trains ($i(t',t) = 1$), the relation between the beginning of utilization and occupation may hold as an inequality. This inequality is due to the need of keeping platforms utilized, as explained when describing Constraints (12). For Constraints (16), the utilization of a track-circuit tc lasts as long as the train utilizes it along any route including tc ($r \in R_t : tc \in TC^r$), plus the formation time. For each route, for ease of visualization, we refer to this quantity as utilization length $ul_{t,r,tc}$, which includes the time necessary for the train to traverse all the track-circuits existing between the reference $ref_{r,tc}$ and tc itself, plus the delay possibly accumulated, plus the formation and release times:

$$ul_{t,r,tc} = \sum_{tc'\in TC(ref_{r,tc},tc,r)} rt_{r,ty_t,tc'}x_{t,r} + \sum_{\substack{tc'\in TC(ref_{r,tc},tc,r):\\ bs_{r,tc'}\neq bs_{r,s_{r,tc'}}}} d_{t,r,tc'} + (ct_{r,ty_t,tc} + for_{bs} + rel_{bs})\,x_{t,r}$$

Constraints (17) to (19) are disjunctive constraints imposing that track-circuit utilization by two trains do not overlap. Hence, at most one train utilizes a track-circuit at any time and capacity constraints are respected. Constraints (18) and (19) ensure that, if $t \prec t'$ on tc, then t's utilization ends before the utilization of train t' starts. Instead, if $t' \prec t$ on tc, then t''s utilization must end before t's utilization can start. If two trains use the same rolling stock, the constraints do not apply to track-circuits belonging to extreme block sections ($i(t,t')\sum_{r\in R_t} e(tc,r) = 1$ or $i(t',t)\sum_{r\in R_{t'}} e(tc,r) = 1$). If a train does not use a track-circuit, its utilization starts and ends at time zero. Hence, for ensuring feasibility, the value of the constant M must be at least as high as the latest end of a track-circuit utilization.

The fact that a train entering a track-circuit is still utilizing the preceding one (for both the clearing and the release time) ensures the feasibility of routes assigned to trains going in opposite directions.

5.3 Route-lock route-release (RR) formulation

As explained in Section 4, the difference between route-lock sectional-release and route-lock route-release (which is the only possibility when modeling the infrastructure with rough granularity) consists in the ending time of the utilization of track-circuits: in the latter the utilization of track-circuit tc along block section $bs_{r,tc}$ lasts until the train has exited the last track-circuit belonging to any block section sharing with $bs_{r,tc}$ one or more track-circuits.

Hence, through the formulation described in Section 5.2, we can consider a route-lock route-release interlocking system by adding a set of constraints. These constraints link the end of the utilization of track-circuit tc to the beginning of the occupation of the first track-circuit of a different block section. In particular, let $refEnd_{r,tc}$ be the last

Transportation Research Part B: Methodological, Volume 59, January 2014, p58–80, © 2013 Elsevier, http://dx.doi.org/10.1016/j.trb.2013.10.013

track-circuit of $bs_{r,tc}$; if train t uses route r, then its utilization of tc ends when t begins occupying $s_{r,refEnd_{r,tc}}$ plus the clearing time of $refEnd_{r,tc}$ and the release time:

$$eU_{t,tc} \geqslant \sum_{\substack{r \in R_t: \\ tc \in TC^r}} o_{t,r,s_{r,refEnd_{r,tc}}} + (ct_{r,ty_t,refEnd_{r,tc}} + rel_{bs})\, x_{t,r} \quad \forall t \in T,\ tc \in TC_t. \quad (20)$$

For preserving feasibility, we must modify Constraints (16) by setting them as inequality constraints: the end of the utilization is greater than or equal to the exit of the train from the track-circuit.

5.4 Optimization in a rolling-horizon framework

For both the SR and the RR formulation, the optimization may be included in a closed-loop optimization: several optimizations are performed subsequently for scheduling and routing trains during a long time horizon, e.g., one day, and the decisions resulting from each optimization are implemented as they are produced. In this case, a rolling-horizon framework must be considered: the time interval for a single optimization advances throughout the day. The time intervals for two subsequent optimizations are often overlapping. In this framework, we must ensure the compatibility of the decisions made in the time interval being tackled and the previously made ones. Previously made decisions can be either modifiable or not.

In both cases, we must account for the utilization of part of the infrastructure by trains which are not part of the optimization (they enter the control area before the beginning of the current time interval) but currently utilize part of the infrastructure itself; let these trains be grouped in set $\mathcal{T}$. If previously made decisions are non-modifiable, this utilization concerns the whole route traversed by these trains. Otherwise, it concerns only the partial route traversed up to the last track-circuit whose utilization started before the beginning of the current time interval. Let this last track-circuit be named ltc, and let $ltc = tc_\infty$ if previously made decisions are non-modifiable. We then define a further set of binary variables $y_{t,t',tc}$ for all triplets $t \in T$, $t' \in \mathcal{T}$ and $tc \in TC_t \bigcup TC(ltc, tc_\infty, r)$, with r being the route traversed by t'. The set of track-circuits for which these additional variables shall be defined may be further reduced after the analysis of the infrastructure, due for example to the absence of switches allowing the reordering of t and t'.

If previously made decisions are modifiable, we must include in the optimization (and hence in set T) all trains belonging to $\mathcal{T}$ for which $ltc \neq tc_\infty$, i.e., all trains in $\mathcal{T}$ which have not started the utilization of their final track-circuit before the beginning of the current time interval. The routes available for these trains include all partial routes connecting ltc to the train original destination, or to the available platforms if platform modification is allowed. Finally, we must include in the formulation the constraints ensuring the time coherence of track-circuit utilization with the already traversed partial route.

Finally, in a rolling-horizon framework, we must include in set T all trains which enter the control area after the end of the current time interval and use the same rolling stock as a train in T. By doing so we ensure that the platform used for the turnaround, join or split is utilized for the whole time between the first arriving train reaching the platform and the last departing one leaving it. Moreover, we guarantee that, when rerouting and

Transportation Research Part B: Methodological, Volume 59, January 2014, p58–80, © 2013 Elsevier, http://dx.doi.org/10.1016/j.trb.2013.10.013

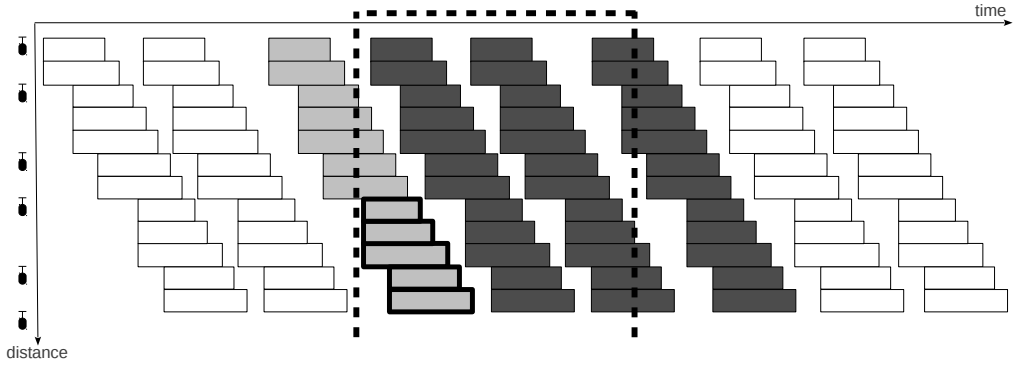

Figure 4: Space-time diagram representing a sequence of trains traveling along a mono-directional line, considered in a rolling-horizon framework. The white trains are not part of the current optimization (dashed lines). The dark-grey trains are those included in T. The light-grey one composes set $\mathcal{T}$. If previously made decisions are modifiable it also belongs to set T: the partial route highlighted through the thick contour is subject to optimization.

rescheduling trains, we preserve enough capacity for not incurring in a deadlock with respect to the trains which have to leave the platform. For the same reason, if some trains leave a platform at any time in the day using a rolling stock already in place (which had been used for a train arrived the previous evening, for example), they are included in set T until the beginning of the time interval follows their arrival in the control area. Once a train has been considered as part of set T, it will be part of this set for all instances corresponding to a time interval preceding or including its arrival in the control area. In this way, we can ensure that the rerouting and rescheduling set for these trains are optimal with respect to the whole time horizon preceding the moment in which the final decisions have to be made.

Figure 4 represents an example of train sequence to be considered in a rolling-horizon framework on a space-time diagram. For sake of simplicity, we consider a mono-directional line and we focus on five block sections. The current time interval is included between the dashed lines. The trains whose track-circuit utilization is represented in white are excluded from the current optimization: either they were part of a previous optimization, or they will be part of a subsequent one. The trains whose track-circuit utilization is represented in dark-grey are those entering the control area within the current time interval: they are included in set T. The train for which we use the light-grey color is the only one belonging to set $\mathcal{T}$: it has entered the control area before the beginning of the current time interval but it has not exited it yet. If previously made decisions are modifiable, it also belongs to set T: ltc is the last track-circuit of the third block section traversed: it is the last one that the train has started utilizing before the beginning of the current time interval. In this case, the partial route highlighted through the thick contour is part of the optimization.

In Pellegrini et al. (2013), we presented a preliminary analysis in which we applied our

Transportation Research Part B: Methodological, Volume 59, January 2014, p58–80, © 2013 Elsevier, http://dx.doi.org/10.1016/j.trb.2013.10.013

formulation in a rolling-horizon framework, where previous decisions are not modifiable.

6 Experimental setup

In the experimental analysis, we assess the impact of the granularity of the representation of the infrastructure on the quality of the solution, i.e., on the delay imposed to trains according to the optimal routing and scheduling returned by either the SR or the RR formulation.

Indeed, the RR formulation assesses routing and scheduling solutions based on a longer headway time between trains than the SR one. To fairly compare SR and RR solutions, we compare the impact of the routing and scheduling decisions implied by the solutions themselves, eliminating the difference in headway times. To this aim, we assess each RR optimal solution S^* through the SR formulation, where we impose train routing and scheduling to be those of S^* itself. Here, we interpret scheduling decisions as train ordering, hence considering only the value of y variables rather than the exact timing in which trains are scheduled to utilize track-circuits in S^*. In the following we will refer to the so re-computed objective function value as the result of the RR_rec formulation.

As reported in Section 5, we consider two alternative objective functions: the maximum delay suffered by any train (max delay), and the total delay suffered (tot delay). We perform parallel experimental analyses for these two functions, using the same experimental setup.

We run the experiments on Intel Xeon twelve core 2.67GHz processor with 24 GB RAM, under Linux Ubuntu distribution version 12.04., using the IBM ILOG CPLEX Concert Technology for C++ (IBM ILOG CPLEX version 12, default parameter settings) (IBM Corporation, 2012) for implementing our formulations. For each run, we impose a limit of 125 hours of CPU time. This time limit is clearly not in line with real-time purposes, when a solution is typically needed in three minutes (Rodriguez, 2007). We set it so high because we aim at measuring the impact of the interlocking system on the optimal solution. This analysis is finalized at identifying the most appropriate interlocking system, rather than the real-time application of the formulation.

For speeding up the solution process, we incorporate both the SR and the RR formulations in a two-step cycle. In the first step, we perform an optimization without changing train routes, i.e., imposing the use of the route fixed in the timetable. In the second step, we use the solution so obtained as starting point for the optimization with all possible train routes. If the first step returns a solution with cost equal to zero, then the perturbation does not directly concern the instance tackled, and there is no need for the second step. For the first optimization step, we allow CPLEX running for sixty seconds, or until it finds any integer solution if it does not manage to do so within this time. We impose this restriction because we are not really interested in the optimal solution with no route changes, unless its cost is equal to zero. If its identification is too time consuming, we interrupt the first step and directly investigate the situation in which we can change routes. By performing this two step optimization we make the most of the available time under two perspectives: on the one hand, we obtain very quickly a feasible and quite good solution; on the other hand, we get a starting solution for CPLEX exploration of the search space, which becomes very large when considering all rerouting

Transportation Research Part B: Methodological, Volume 59, January 2014, p58–80, © 2013 Elsevier, http://dx.doi.org/10.1016/j.trb.2013.10.013

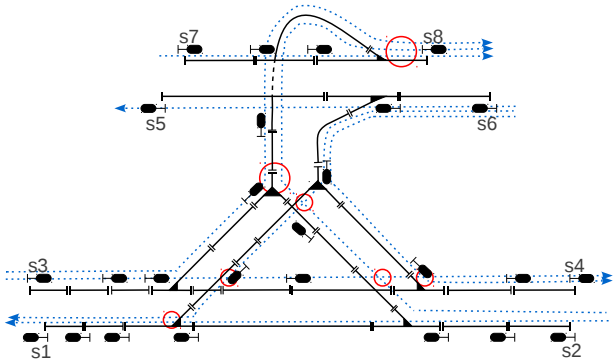

Figure 5: Infrastructure of the control area referenced as triangle of Gagny. Dotted arrows represent the possible routes. Circles show the potential conflicts.

possibilities.

We divide the experimental analysis in two parts. First, we propose a brief analysis on instances representing traffic on a rather simple, at least apparently, junction known as the triangle of Gagny, in France. In this analysis we show in detail the difference between the block section and the track-circuit models in a possible practical application. Then, we propose a thorough analysis on instances representing traffic in a very complex junction, namely the control area including the Lille-Flandres station, in France. In both cases, we consider formation and release times of 15 and 5 seconds, respectively, for all block sections and a three aspect signaling system. Track-circuit running and clearing times depend on the characteristics of the track-circuit itself (its length, its curve, its gradient), on the route along which it is used (which implies the maximum speed attainable, e.g., if the previous or the next switch are nearby, the maximum speed will be quite low) and on the train type (which implies its length, its weight, its acceleration and braking capability), as mentioned in Section 5, and are expressed in seconds.

6.1 Instances representing traffic in the triangle of Gagny

The instances tackled in the first part of our analysis represent traffic in the junction named triangle of Gagny. Figure 5 depicts the infrastructure characterizing this junction. It includes eight possible routes. Only one route exists for connecting each o-d pair: no rerouting is possible. Conflicts may emerge in many locations, as shown in the figure. Routes include 3 to 10 track-circuits, and 2 to 6 block sections. The total running time varies between 60 and 342 seconds for a conventional train. In the figure, we name only signals which correspond to either the entry or the exit point of a route, for ease of visualization.

Based on this infrastructure, we solve 301 instances considering the total delay objective function. In the first instance, we consider three trains: t1 goes from s3 to s8 and enters the control area at second 100; t2 goes from s3 to s4 and enters the control area at second 212; t3 goes from s6 to s4 and enters the control area at second 162.

Transportation Research Part B: Methodological, Volume 59, January 2014, p58–80, © 2013 Elsevier, http://dx.doi.org/10.1016/j.trb.2013.10.013

Table 1: Mean number of variables and constraints in the three sets of instances representing traffic in the triangle of Gagny.

horizon	# continuous var.	# binary var.	# constraints in SR formulation	# constraints in RR formulation
150 sec	270	166	926	1000
300 sec	268	169	932	1005
450 sec	273	176	961	1036

We obtain the remaining 300 instances by generating them randomly, and we group them in three sets of 100 instances each. In the instances of the first set, 10 trains enter the control area within 150 seconds. In the instances of the second set, 10 trains enter the control area within 300 seconds. In the instances of the third set, 10 trains enter the control area within 450 seconds. By varying the duration of the time horizon and keeping fix the number of trains, we simulate three situations with decreasing train density in the control area. For each train, both the route and the specific time at which it enters the control area are randomly drawn considering a uniform probability distribution: the route is randomly drawn for the list of the eight possible ones; the entrance time is randomly drawn from the set of integer number between zero and the duration of the time horizon characterizing the instance. Table 1 reports the mean size of the formulations corresponding to these sets of instances. In the table, each set is identified by the duration of the time horizon in which trains enter the control area, and for each set we report the corresponding mean number of variables, distinguished in continuous and binary, and the mean number of constraints in the SR and the RR formulations. On the one hand, the number of continuous variables depends on the number of trains in an instance and on the number of track-circuits it may use. On the other hand, both the number of binary variables and the number of constraints depend on the number of potential conflicts in terms of pair of trains and common track-circuits: the time at which these trains may use the common track-circuits does not play a role in the model definition. As a consequence, the number of neither variables nor constrains varies noticeably as a function of the duration of the time horizon, and hence as a function of the train density.

6.2 Instances representing traffic in the Lille-Flandres station

In the main part of the experimental analysis, we tackle instances representing perturbations of the timetable of a weekday in 2002 in the control area including the main station of Lille in the North of France, i.e., the Lille-Flandres station. In particular, we consider a Wednesday timetable including 589 trains. We do not have any information on connections, and hence we do not consider Constraints (7) presented in Section 5. Being the Lille-Flandres a terminal station, all rolling stocks are used for both an arriving and a departing train, but for what concerns the first trains departing in the morning (which arrived the day before to the platform) and the last ones arriving at night (which will leave the platform the day after): for almost any train t (97.11% of the total) a t' exists

Transportation Research Part B: Methodological, Volume 59, January 2014, p58–80, © 2013 Elsevier, http://dx.doi.org/10.1016/j.trb.2013.10.013

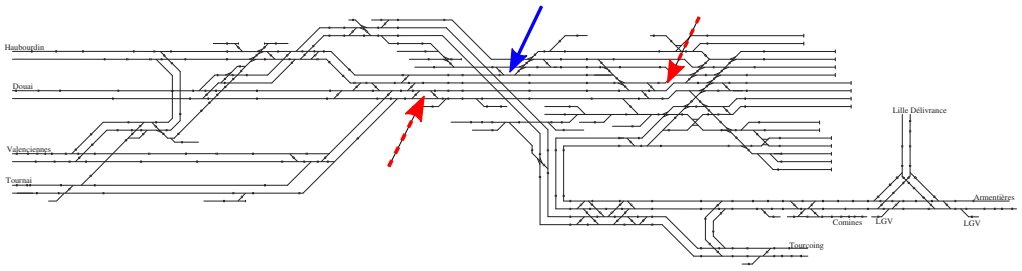

Figure 6: Infrastructure of the control area including the Lille-Flandres station. The solid arrow indicates the unavailable track-circuit in the partially disrupted scenario. The dashed arrows indicate the further track-circuits unavailable in the severely disrupted scenario.

such that $i(t,t') = 1$ or $i(t',t) = 1$. Besides 259 turnarounds, the timetable contains 8 joins and 10 splits.

Figure 6 depicts the infrastructure of the control area: the station is linked to seven regional, national and international lines and it has 17 platforms. All routes either depart or arrive at the station: either their initial or their final track-circuit is a platform. A total of 2409 routes exist and they are composed by 299 track-circuits. The routes include 9 to 35 track-circuits (mean = 24), 2 to 13 block sections (mean = 5), and they have a total running time of 141 to 707 seconds (mean = 343) and a total length of 950 to 11500 meters (mean = 4331). The track-circuits measure between 10 and 1710 meters (mean = 177) The trains considered in this analysis measure 100 meters each.

We consider three different *infrastructure scenarios*: as often done in the literature (see, e.g., Corman et al. (2010)) we limit the number of available train routes by forbidding the use of selected track-circuits, simulating for example maintenance works. In the so called ***fully functioning*** scenario, all the existing routes are available; in the so called ***partially disrupted*** scenario, about 70% of routes are available; in the so called ***severely disrupted*** one, this percentage is slightly higher than 40%. More in detail, in the fully functioning scenario, trains have 1 to 72 available routes (mean = 10); in the partially disrupted scenario, trains have 1 to 36 available routes (mean = 6) and in the severely disrupted one, trains have 1 to 10 available routes (mean = 3). In Figure 6, we indicate with arrows the track-circuits which are unavailable in the partially disrupted (solid arrow) and in the severely disrupted (both solid and dashed arrows) scenarios. In these experiments, we consider the platform assigned to trains as non modifiable. If the route assigned to a train according to the timetable is not available in either the partially disrupted or the severely disrupted scenario, in the first optimization step explained at the beginning of this section we consider the first available route in the list of those which the train itself can use.

Starting from the original timetable, we impose a delay to 20% of trains which do not represent shunting movements: we randomly select the trains to be delayed and we randomly draw their delay in the interval between 5 and 15 minutes (Lusby et al., 2012). Both these random selections are based on uniform probability distributions. By repli-

Transportation Research Part B: Methodological, Volume 59, January 2014, p58–80, © 2013 Elsevier, http://dx.doi.org/10.1016/j.trb.2013.10.013

cating the random assignment of train primary delay 30 times, we obtain 30 different perturbed one-day timetables. For each of these 30 perturbed one-day timetable, we consider the peak-time of the day. By varying the duration of the time horizon ending at 7:30 am, we assess the performance of both the SR and the RR formulations on different size instances. In particular, we consider the 30 instances, one for each perturbed one-day timetable, including all trains entering the control area between 7:10 and 7:30 am (set I20), those including all trains entering the control area between 7:00 and 7:30 am (set I30), and so on up to the interval 6:30 to 7:30 am (I60). The exact characteristics of each instance depend on the specific perturbed one-day timetable originating it. Table 2 reports a summary of these characteristics and of the corresponding size of the formulations. In particular, the first part of the table indicates three quantities for each set of instances: the number of trains, the number of trains suffering a primary delay and the total primary delay suffered by trains, measured in seconds. For each of these quantities, the table reports the mean calculated on the 30 instances of the set, and the minimum and maximum values between parenthesis. Remark that a few instances do not include any of the 20% of trains suffering a primary delay; in this case the total primary delay associated to the instance is null. The second, third and fourth parts of the table report, for each infrastructure scenario, the number of variables, split into continuous and binary ones, and the number of constrains in the SR and the RR formulations: the values reported correspond to the mean computed on the 30 instances of each set. The number of variables is equal for the two formulations, while the number of constraints is slightly higher for the latter: it includes constraints for guaranteeing that we unlock concurrently all partially overlapping block sections, as explained in Section 5.3. The number of continuous variables grows almost proportionally to the number of trains and the number of routes available in the different sets of instances and infrastructure scenarios. The number of both binary variables and constraints, instead, grows more than proportionally, since they strongly depend on the number of possible conflicts, which intuitively grows more than proportionally with respect to, for example, the number of trains.

As explained in Section 5, the value of the constant M needs to be at least equal to the latest end of a track-circuit utilization for ensuring the coherence of Constraints (18) and (19). We set the value of $M = 86400$, i.e., midnight computed in seconds, for ensuring this coherence.

7 Computational results

The main indication of the results of our computational analysis is that there is an actual difference between optimizing traffic when considering different granularities for the representation of the infrastructure. This is due to the different interlocking systems that can be modeled, based on different granularities of the representation: route-lock sectional-release with the fine granularity, and route-lock route-release with the rough one.

This indication supports the observation made on a similar comparison by Corman et al. (2009b). In particular, Corman et al. (2009b) compared the solutions obtained when applying the route-lock route-release to an approximation the route-lock sectional-

Transportation Research Part B: Methodological, Volume 59, January 2014, p58–80, © 2013 Elsevier, http://dx.doi.org/10.1016/j.trb.2013.10.013

Table 2: Summary of the main characteristics of the instances tackled representing the Lille-Flandres station. The first quantity reported is the mean value, while the ones in parenthesis are the minimum and the maximum values, respectively, with respect to the 30 instances of each set.

	# trains	# trains suffering primary delay	total primary delay (sec)
I20	15 (11, 18)	2 (0, 5)	616 (0, 704)
I30	22 (19, 24)	3 (1, 7)	609 (305, 760)
I40	29 (26, 32)	5 (3, 8)	611 (305, 796)
I50	36 (33, 40)	6 (3, 12)	617 (305, 805)
I60	45 (42, 47)	8 (3, 15)	619 (305, 833)

fully functioning scenario

	# continuous var.	# binary var.	# constraints in SR formulation	# constraints in RR formulation
I20	7153	2885	26524	27224
I30	10785	6466	46319	47357
I40	13113	11349	66564	67933
I50	15970	16893	90346	92005
I60	20156	26400	129315	131432

partially disrupted scenario

	# continuous var.	# binary var.	# constraints in SR formulation	# constraints in RR formulation
I20	4772	2587	19399	20051
I30	7083	5698	34369	35327
I40	8774	9853	50788	52040
I50	10653	14666	69805	71320
I60	13358	23034	101523	103403

severely disrupted scenario

	# continuous var.	# binary var.	# constraints in SR formulation	# constraints in RR formulation
I20	2520	1388	10067	10569
I30	3715	3223	18370	19114
I40	4836	6000	29233	30227
I50	5903	8972	40611	41823
I60	40611	14054	59097	60598

release interlocking systems considering fix train routes, as mentioned in Section 2. Even if the latter system is not practically applicable for safety reasons, the results obtained by considering it are judged indicative of the results obtainable with the route-lock sectional-release system. The difference between the systems in terms of secondary delay appears evident in the results presented. The authors considered this difference small enough to justify the use of the route-lock route-release system in further studies. As in the reference paper, also in our results the difference does not exceed a value of a few minutes when considering sets of few dozens of trains. However, we consider these few minutes to be

Transportation Research Part B: Methodological, Volume 59, January 2014, p58–80, © 2013 Elsevier, http://dx.doi.org/10.1016/j.trb.2013.10.013

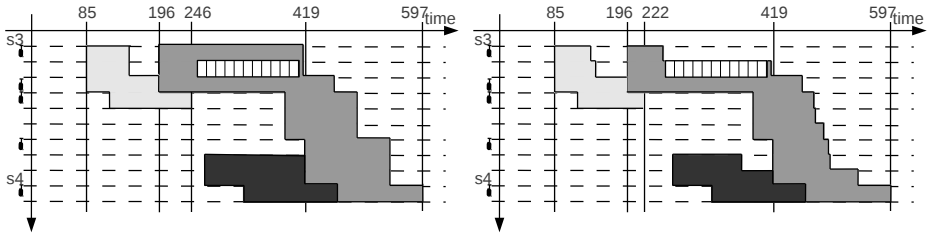

Figure 7: Gantt diagram associated to trains along t2's route according the RR (left) and the RR_rec (right) formulations. t1 → light-grey; t2 → grey; t3 → dark-grey; delay → box containing vertical dashes.

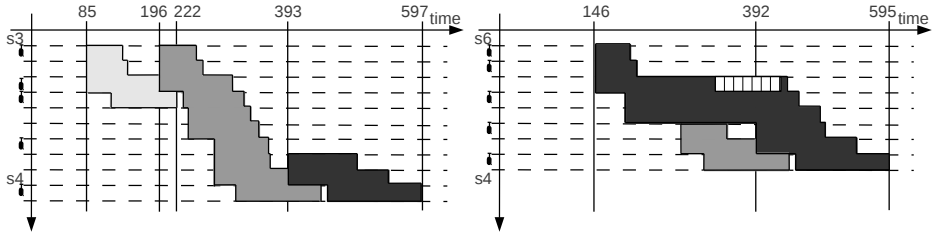

Figure 8: Gantt diagram associated to trains along t2's (left) and t3's routes (right) according the SR formulation. t1 → light-grey; t2 → grey; t3 → dark-grey; delay → box containing vertical dashes.

important in practical terms, especially when considering traffic at a main station as the Lille-Flandres one.

7.1 Triangle of Gagny

In the first part of this analysis, we focus on the detailed observation of the solution returned by the SR and the RR formulations when minimizing total delay for the three train instance on the triangle of Gagny, described in Section 6.1.

Figure 7 shows the Gantt diagrams associated to the three trains along t2's route according the RR formulation (Figure 7 left) and after its reassessment into the RR_rec one (Figure 7 right). Figure 8 shows the Gantt diagram associated to t2's (Figure 8 left) and t3's routes (Figure 8 right) according the SR formulation. Track-circuit utilizations of t1 are shown in light-grey, the ones of t2 and t3 in grey and dark-grey, respectively. A box containing vertical dashes indicates when a train is delayed. We do not show either track-circuit or signal names for ease of representation. Given the small size of the infrastructure, shown in Figure 5, it is anyway simple to identify them.

Train t1 travels along its route never crossing or following another train. Hence, it gets no secondary delay, no matter the formulation used. Train t2 follows t1 at the beginning of its route, until the third block section, where the routes of the two trains diverge. Due

Transportation Research Part B: Methodological, Volume 59, January 2014, p58–80, © 2013 Elsevier, http://dx.doi.org/10.1016/j.trb.2013.10.013

Table 3: Mean difference (RR_rec-SR) and mean percentage difference ((RR_rec-SR)/SR %) between the secondary delay obtained by implementing the optimal solutions of the RR_rec and the SR formulations on the instances of the three sets representing traffic on the triangle of Gagny. Delay difference is expressed in seconds.

Minimization of total secondary delay		
150-sec instances	300-sec instances	450-sec instances
68 sec	57 sec	54 sec
4.97%	5.22%	6.38%

to the application of the blocking time theory and the three aspect signaling system, train t2 can start utilizing the second block section (and hence occupying the first one after waiting for the 15 seconds of formation time) after t1 has finished utilizing it (at second 196). The running time for t2 along the first block section is 36 seconds; hence it may potentially start occupying the second block section at second 247. However, it can start occupying the second block section after t1 has finished locking the third one, plus the formation time. For the RR formulation, in the third block section, t1 needs to travel across two track-circuits before unlocking the block section needed by t2: the unlock occurs at second 246. For the SR formulation, instead, the block section is unlocked at second 222, after t1 exiting the track-circuit common to t1's and t2's block sections. The two formulations see, hence, a difference in the assessment of the secondary delay due to t2 following t1.

For the RR formulation, t2 necessarily suffers a delay of 14 seconds ($247 + 15 - 248$). After this delay, for the same type of reasoning made for t1 and t2, if t2 passes before t3 on the common track-circuits, then t3 must be stopped: it will suffer a secondary delay of 142 seconds. The total delay of this solution is then the sum of the 14 seconds suffered by t2 and the 142 seconds suffered by t3, that is, 156 seconds. Instead, having t3 passing before t2 and not being stopped, t3 suffers no delay, and the delay suffered by t2 amounts to 155 seconds. This second solution, hence, appears preferable when minimizing total delay. Even in the re-computation of the solution value according to the RR_rec formulation, this delay cannot be reduced: the solution of the RR formulation translates into having t3 passing before t2 on the common track-circuits, with t2 hence being able to enter the common track-circuits with 155 seconds of delay.

For the SR formulation, instead, it is immediately recognized that t2 does not need to be delayed due to t1, and that its passing straight, before t3, implies only t3 to be delayed of 128 seconds. If t2 has to wait, instead, it will be, as for the RR_rec formulation, delayed of 155 seconds.

Both solutions are feasible for the two formulations; however, the longer (and actually not practically necessary) utilization of a track-circuit by t1 implies that delay is overestimated by the RR formulation.

For understanding how this longer utilization of track-circuits may impact the results even on such a small control area, we solve random instances containing trains arriving

Transportation Research Part B: Methodological, Volume 59, January 2014, p58–80, © 2013 Elsevier, http://dx.doi.org/10.1016/j.trb.2013.10.013

in the control area within 150, 300 and 450 seconds. The results are reported in Table 3. The table details the mean, across the 100 instances of each set, of the difference of total secondary delay suffered by trains after the optimization: the difference appears in absolute terms in the first line of the table, measured in seconds, and in percentage terms in the second line. When the traffic is more intense, i.e., when the trains arrive in the control area within the shortest time horizon, both the total delay suffered by trains and the impact of granularity are higher. However, the difference between the results of the RR_rec and the SR formulations decreases slower than the total delay as a function of the increase of the time horizon: the percentage difference between the results of the two formulations is the highest in the case of the 450-second instances. For example, the mean total delay (in seconds) for the SR formulation passes from 1133 in the 450-second instances to 1627 in the 150-second ones. The difference between RR_rec and SR formulations, instead, varies of only 14 seconds, as shown in the table.

7.2 Lille-Flandres station

The thorough analysis which we perform on instances representing traffic in the Lille-Flandres station clearly shows that the granularity of the representation of the infrastructure has an impact on the quality of the optimal solution.

Table 4 reports the mean difference between the results obtained through the RR_rec and the SR formulations, which correspond to rough and fine granularity of the representation of the infrastructure, respectively. We consider here only the instances for which both formulations found the optimal solution (and proved its optimality) within the time limit of 125 hours of CPU time. We group the results as a function of both the set of instances and the infrastructure scenario. When the instances include few trains (I20 and I30), the difference between the results is in average quite small, despite always being statistically significant in favor of the the SR formulation, according to the Wilcoxon rank-sum test with a confidence level of 0.95. The total or maximum secondary delay amounts in average to few seconds in the smaller instances and with the fully functioning infrastructure. Nonetheless, these small absolute values correspond to non-negligible percentage of the secondary delay suffered by trains. The quantities in parenthesis report the number of instances which could be solved to optimality by both the RR and the SR formulation within the time limit. As also discussed in the following (in the comments to Tables 5 and 6), these quantities show that the difficulty of the instances increases with the size of the instances themselves. Moreover, finding the solution which minimizes total delay and proving its optimality appear harder than finding the one minimizing maximum delay. In particular, rather few instances of sets I50 and I60 could be solved within the time limits by both the SR and the RR formulations minimizing total delay in the fully functioning and partially disrupted scenarios. The worst case is represented by set I60 in the fully functioning scenario: only for 7 of the 30 instances, both formulations could complete the search process within the time limit. In particular, the RR formulation could not complete the search process for 22 instances; the SR formulation could not complete the search process for 12 of these instances and for a further one. Even when the number of instances solved to optimality becomes rather small, anyway, the first feasible solutions are found by the two formulations in few seconds.

The consequent small number of results might bias the mean and percentage differ-

Transportation Research Part B: Methodological, Volume 59, January 2014, p58–80, © 2013 Elsevier, http://dx.doi.org/10.1016/j.trb.2013.10.013

Table 4: Mean difference (RR_rec-SR) and mean percentage difference ((RR_rec-SR)/SR %) between the secondary delay obtained by implementing the optimal solutions of the RR_rec and the SR formulations on the instances of each set for which the optimal solution was found by both formulation within the time limit, and for the three infrastructure scenarios. Delay difference is expressed in seconds. Between parenthesis we indicate the number of instances solved to optimality by both formulations within the time limit.

Minimization of maximum secondary delay

infrastructure scenario	I20	I30	I40	I50	I60
fully functioning	7 (30)	8 (30)	12 (30)	17 (29)	13 (30)
	3.85%	3.49%	4.48%	4.56%	3.05%
partially disrupted	12 (30)	16 (30)	15 (29)	8 (30)	9 (29)
	6.52%	6.67%	5.00%	2.01%	1.94%
severely disrupted	24 (30)	30 (30)	68 (30)	57 (30)	49 (30)
	9.96%	10.49%	19.88%	13.48%	10.38%

Minimization of total secondary delay

infrastructure scenario	I20	I30	I40	I50	I60
fully functioning	16 (30)	32 (30)	66 (26)	96 (13)	18 (7)
	7.02%	10.16%	14.57%	12.44%	2.89%
partially disrupted	20 (30)	35 (30)	71 (29)	110 (17)	62 (8)
	8.33%	8.84%	11.54%	11.18%	5.24%
severely disrupted	60 (30)	146 (30)	391 (30)	456 (28)	525 (21)
	14.15%	19.52%	34.76%	27.44%	43.93%

ences reported: these values can be computed only for the easiest instances. However, the table shows that, even when only the easiest instances are considered, the results are still in favor of the SR formulation. The main reason for comparing the formulations' results only on the instances in which both RR and SR achieve the optimal solution is the fact that we aim to assess the impact of the granularity of the representation of the infrastructure without introducing a bias due to the solution time, which is indeed a consequence of the formulation, of the solver and of the hardware used. By considering only optimal solutions, we eliminate this bias since the optimal remains of course constant through formulations, solvers and hardware. All solved instances are considered, even if corresponding to results which show up as outliers in the distribution of the results themselves. Outliers are observations that are numerically distant from the rest of the data in a distribution: they are all the observations which are either smaller than the first quartile minus 1.5 times the interquartile range or larger than the third quartile plus 1.5 times the interquartile range. We decided to take into account these results as well, since all instances are equally relevant in the scope of this paper.

Figure 9 depicts the whole distribution of the results for the instances of set I30 through boxplots: each box represents the distribution of the observations corresponding

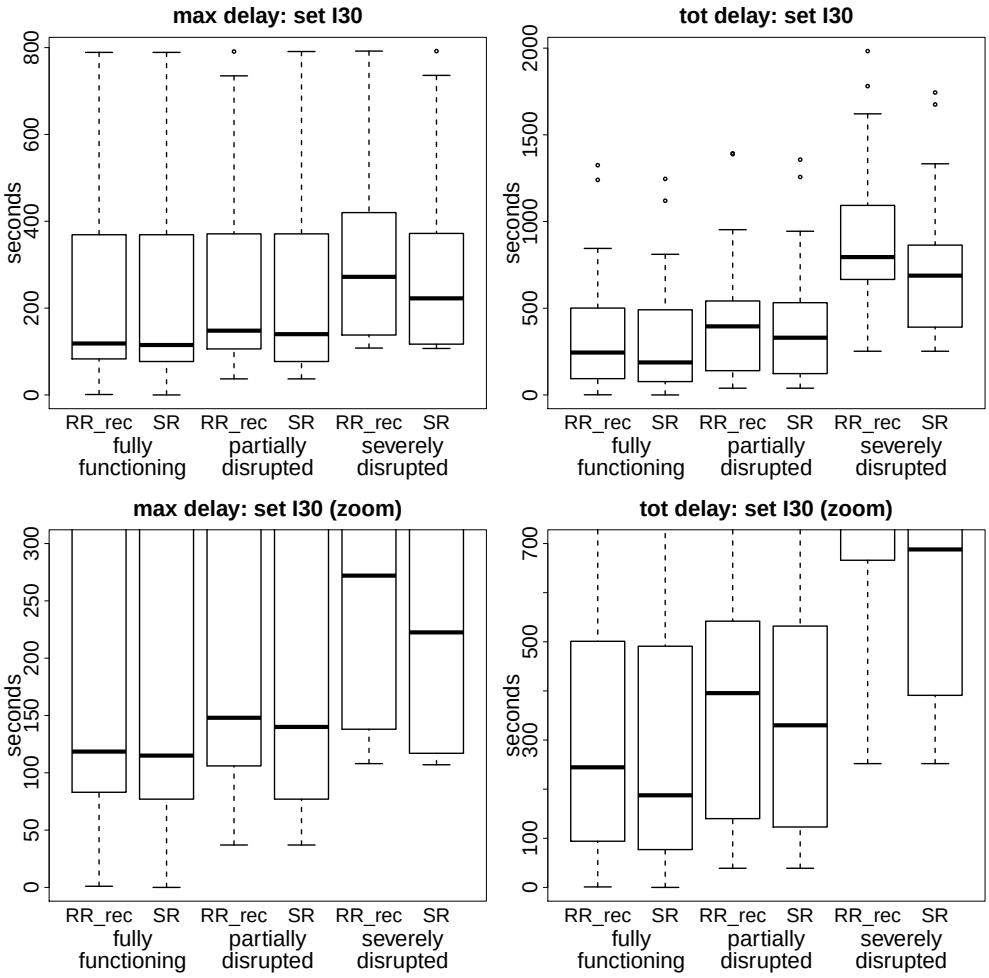

Figure 9: Boxplots of the distributions of the optimal solution value for the RR and the SR formulations.

to the 30 perturbed instances tackled. The horizontal line within the box represents the median of the distribution, while the extremes of the box represent the first and third quartiles, respectively; the whiskers show the smallest and the largest non-outliers in the data-set and dots correspond to the outliers. We are showing here only the results for set I30 since we could solve to optimality, within the time limit, all its instances with both formulations, and since the conclusions which we can draw based on this set are qualitatively equivalent to those based on the other sets.

This representation shows that the difference between the optimal objective function value of RR_rec and SR increases when passing from the fully functioning, to the partially

Transportation Research Part B: Methodological, Volume 59, January 2014, p58–80, © 2013 Elsevier, http://dx.doi.org/10.1016/j.trb.2013.10.013

disrupted, to the severely disrupted infrastructure scenario. The reason of this increase is the number of available routes: intuitively, the larger the number of available routes, the fewer conflicts exist and the fewer trains have to queue for traversing the control area. The fewer conflicts, the fewer scheduling choices to be made, and the smaller the impact of the granularity.

Due to the rather large range of optimal objective function values, between 0 and 800 seconds as for maximum delay (Figure 9, top - left) and between 0 and 2000 seconds as for total delay (Figure 9, top - right), the difference implied by the rough and fine granularity appears almost null in some cases, even if it is always remarkable as emerging from Table 4. This is particularly true in case of the minimization of the maximum delay. The bottom plots of the figure show the same distributions as the top ones, but with a larger scale on the y-axis representing seconds of secondary delay. Also here it is observable that, even if the differences are sometimes small, the SR formulation performs better than the RR_rec one in all scenarios. The difference increases as a function of the infrastructure disruption, and it is larger for total than for maximum delay. It has to be remarked that, anyway, the delay considered in the objective function concerns only one train in the latter case, and hence that small absolute differences are anyway relevant.

For getting a hint on the quality of the solutions under a different perspective, we compare the two formulations after evaluating the solutions themselves according to the alternative objective function: we assess in terms of total delay the routing and scheduling choices made by the two formulations when minimizing the maximum delay, and *viceversa*. Focusing on the severely disrupted scenario, when considering the solutions obtained by the RR formulation minimizing the maximum delay, and assessing them through the RR_rec formulation minimizing total delay, we find a median total delay of 1307 seconds. The corresponding value for the solution of the SR formulation found minimizing maximum delay is of 1201 seconds. Always focusing on the severely disrupted scenario, the RR_rec formulation minimizing the total delay imposes a median maximum delay of 337 seconds, while the SR formulation imposes a median maximum delay of 253 seconds. These values confirm the conclusions drawn when discussing Figure 9: the optimal solution found representing the infrastructure with fine granularity is better than the one found with rough granularity, even when we consider a perspective on secondary delay which is different from the one used in the optimization.

In all the instances tackled and for both objective functions, rerouting plays a major role in the optimization. Figure 10 depicts the mean percentage number of trains which are rerouted in the optimal solution of the SR and RR formulations minimizing total delay. The general trend of the values represented indicates that the RR formulation tends to reroute fewer trains than the SR one: the reason for these fewer trains rerouted may be found in the smaller advantage that rerouting offers when the train is not recognized to quickly unlock block sections, as is the case for the RR formulation. As an example of this smaller advantage, let us consider the case depicted in Figure 11. It represents three train journeys in a portion of the infrastructure. This case actually emerged in one of the instances tackled, and rerouting was actually seen as a disadvantage by the RR formulation. The left plot shows the situation emerging with the scheduled routes: train t1 (dotted line) travels along a conflict-free route, train t2 (dashed line) travels along a route which is in conflict with the one of train t3 (solid line), and t2 passes first. Train t3 suffers a delay when exiting the infrastructure. As shown in the center and right plots,

Transportation Research Part B: Methodological, Volume 59, January 2014, p58–80, © 2013 Elsevier, http://dx.doi.org/10.1016/j.trb.2013.10.013

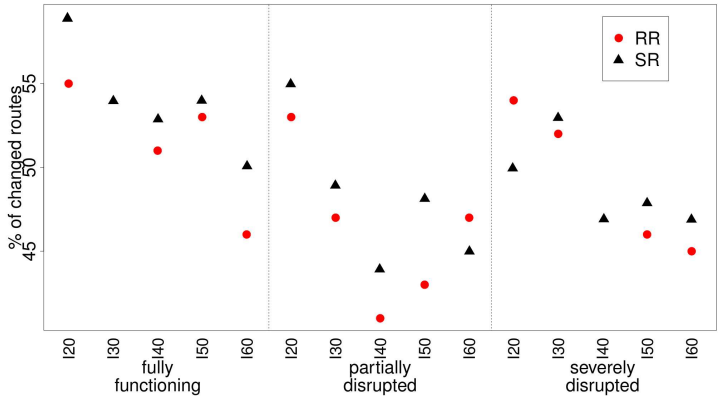

Figure 10: Mean percentage number of rerouted trains.

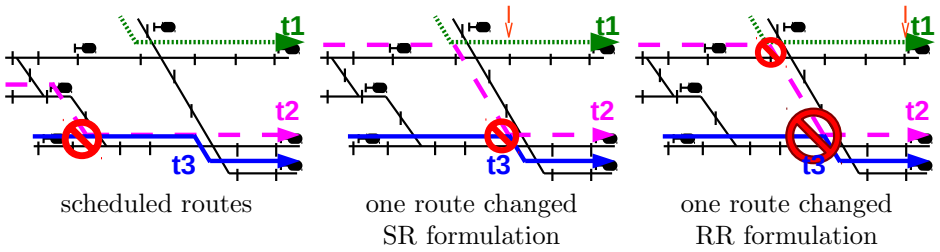

Figure 11: Comparison of the evaluation of rerouting by the RR and the SR formulation. Example with three trains. t1 $\prec$ t2 $\prec$ t3. When train t2 is rerouted, the RR formulation detects a conflict that does not exist for the SR one. The vertical down arrow shows the point which decides the moment when the block section containing a track-circuit in common to t1 is available for t2.

train t2 can travel along an alternative route, which uses one track-circuit in common with the route of t1. The center plot shows the implication of this rerouting for the SR formulation: the critical block section between t1 and t2 becomes available to t2 after t1 has left the common track-circuit (vertical down arrow in the figure). However, the magnitude of the conflict between t2 and t3 is smaller than when t2 uses its scheduled route. The right plot shows the corresponding assessment made by the RR formulation: t1 passes first, and the block section containing the common track-circuit becomes available for t2 only after t1 has passed the last signal shown (vertical down arrow in the figure). Hence, a conflict emerges and t2 must be delayed. In turn, this implies the increase of the magnitude of the conflict between t2 and t3: t3 will be delayed even more than when t2 uses the scheduled route. As a consequence, the RR formulation discards this rerouting option and maintains the scheduled route in the final solution.

What clearly emerges from the analysis, besides the trends implied by the different scenarios, is the relevance of rerouting: in all scenarios, more than 40% of trains are

Transportation Research Part B: Methodological, Volume 59, January 2014, p58–80, © 2013 Elsevier, http://dx.doi.org/10.1016/j.trb.2013.10.013

Table 5: CPU time when minimizing **maximum delay**.

CPU time (sec) for the first optimal solution

		fully functioning		partially disrupted		severely disrupted	
		SR	RR	SR	RR	SR	RR
I20	mean	17	83	7	28	2	4
	median	4	20	3	6	2	3
I30	mean	42	291	17	49	7	11
	median	17	76	7	32	6	10
I40	mean	15076	544	20	83	12	27
	median	37	198	5	37	7	19
I50	mean	193	15828	71	181	32	56
	median	85	262	19	51	22	42
I60	mean	264	2583	180	15272	89	139
	median	124	424	96	193	66	87

CPU time (sec) for the whole search

		fully functioning		partially disrupted		severely disrupted	
		SR	RR	SR	RR	SR	RR
I20	mean	20	133	11	43	4	5
	median	6	33	6	14	3	3
I30	mean	154	6886	50	131	8	15
	median	49	161	18	53	7	15
I40	mean	15149	2157	33	245	16	35
	median	70	385	18	85	9	28
I50	mean	2151	20802	108	753	45	95
	median	223	1092	49	339	32	81
I60	mean	506	11618	263	15925	115	5148
	median	320	3734	172	780	125	203

rerouted through the optimization.

Although in this paper we do not focus on the boost of the solution time, Tables 5 and 6 show the mean and median computation time necessary for finding the first optimal solution and for performing the whole search. As previously mentioned, the first feasible solution is found in few seconds. Once the first feasible solution is found, CPLEX keeps exploring the search space for improving it, or for proving its optimality. As soon as CPLEX finds a better solution than the incumbent one, it returns its objective function value and continues the search for improvement. When it manages to prove the optimality of a solution, it stops the search. When it does not manage to prove this optimality and the search is interrupted due to the reach of the time limit (125 hours of CPU time, as mentioned in Section 6), we consider the incumbent solution at the end of the search as

Transportation Research Part B: Methodological, Volume 59, January 2014, p58–80, © 2013 Elsevier, http://dx.doi.org/10.1016/j.trb.2013.10.013

Table 6: CPU time when minimizing **total delay**.

CPU time (sec) for the first optimal solution

		fully functioning		partially disrupted		severely disrupted	
		SR	RR	SR	RR	SR	RR
I20	mean	43	922	19	180	6	10
	median	10	60	14	42	4	9
I30	mean	329	4279	116	1303	22	50
	median	89	323	78	162	22	29
I40	mean	16246	50516	526	32368	101	659
	median	559	2521	204	801	59	131
I50	mean	75187	261315	54698	200752	552	52492
	median	1643	450000	762	45172	274	1540
I60	mean	207151	323839	182897	302048	21530	155089
	median	65588	450000	15733	450000	1052	12649

CPU time (sec) for the whole search

		fully functioning		partially disrupted		severely disrupted	
		SR	RR	SR	RR	SR	RR
I20	mean	77	1385	26	197	6	11
	median	20	84	19	44	5	9
I30	mean	644	25461	176	3839	23	56
	median	159	507	88	197	22	29
I40	mean	19944	75176	786	38096	117	2323
	median	1046	7139	241	1751	64	131
I50	mean	95226	282950	72803	220681	1197	65348
	median	7618	450000	1504	162303	285	2481
I60	mean	235172	347886	188936	341411	24352	174269
	median	210836	450000	35237	450000	1330	47187

the optimal one, and the solution time equal to 450000 seconds. Let us remark that this time is the total CPU time, which corresponds to a much lower wall-clock time when exploiting CPLEX capabilities of parallel computation (about 12 hours on the hardware used for this analysis).

The RR formulation appears here much slower than the SR one. As remarked in Table 4, for both formulations, the identification of the minimum total-delay solution requires a much longer time than the identification of the minimum maximum-delay one. In almost all cases, i.e., for both objective functions, both formulations and all infrastructure scenarios, the instances representing 20 minute time horizons are solved in a time which is in line with the necessity of real-time traffic management, even in the worst case scenario in which only one CPU is available for the computation. The

Transportation Research Part B: Methodological, Volume 59, January 2014, p58–80, © 2013 Elsevier, http://dx.doi.org/10.1016/j.trb.2013.10.013

threshold often considered is three minutes (Rodriguez, 2007) (in some cases an even higher time of four minutes and thirty seconds may be accepted (Lusby et al., 2012)). As expected, the CPU time is typically smaller in case of the severely disrupted scenario, when fewer routes translate into fewer variables and constraints. The values reported may appear sometimes surprising, as the case of the set I40 for the SR formulation minimizing maximum delay on the fully functioning infrastructure (Tables 5): the mean computation time is much higher than the corresponding value for set I50. This high mean value is actually due to a single instance, in which the best known feasible solution with objective function value 155 and optimality gap of 45 seconds is found in about two minutes; however, the optimality gap cannot be reduced within the time limit. Excluding this instance, the mean computation time is be 154 seconds. When we add six trains to this same instance, i.e., we consider the corresponding one in set I50, the optimal solution has objective function value of 155, and only 3450 seconds are needed for completing the search. The same reasoning holds for the other apparently surprising cases. These results show that the number of trains present in an instance is only a proxy of its difficulty. Further research must be devoted to the analysis of the factors that actually explain an instance difficulty.

We consider these results quite encouraging, especially if we focus on the SR formulation. As it is suggested by the difference between the mean and the median CPU time, the formulation is quite slow for few instances, while it is rather fast in the large majority of the cases.

8 Practical applicability

Both the SR and the RR formulation proved to be able to solve to optimality realistic instances. In this section, we report the results of some experiments aiming to determine the practical applicability of the SR formulation for real-time purposes. We focus here only on the SR formulation since the results presented in Section 7 show that at the optimum it makes more advantageous choices than the RR formulation. Even if the optimum is not always reachable within the time limit imposed by a real-time implementation, aiming to the very best possible solution remains definitely the overall objective.

To understand the practical applicability of the SR formulation, we solve the same instances tackled in Section 7.2, which represent traffic in the Lille-Flandres station. We consider a time limit for CPLEX execution of three minutes (Rodriguez, 2007) of wall-clock time, on the hardware described in Section 6.

Table 7 reports the results of these further experiments. For assessing the performance of the SR formulation, we compare the results obtained within three minute computation (3M) to the optimal results achieved with no possible train rerouting (NOR). Such an optimal solution is always achieved in less than two CPU minutes. For each set of instances, infrastructure scenario and objective function, we report the mean objective function value achieved across the 30 corresponding instances through 3M and NOR, in seconds. The difference represents the advantage offered by the SR formulation within a time limit suitable for real-time purposes, compared to the best possible solution implementable with the scheduled routes. The best result for each set of instances,

Transportation Research Part B: Methodological, Volume 59, January 2014, p58–80, © 2013 Elsevier, http://dx.doi.org/10.1016/j.trb.2013.10.013

Table 7: Mean number of seconds of secondary delay avoided by running the SR formulation for three minutes (3M) compared to optimally solving each instance with no train rerouting (NO-R). The bold font corresponds to the best mean result.

Minimization of maximum secondary delay

infrastructure scenario	I20		I30		I40		I50		I60	
	NOR	3M	NOR	3M	NOR	3M	NOR	3M	NOR	3M
fully functioning	249	**182**	296	**229**	336	**263**	419	**372**	490	**436**
partially disrupted	315	**184**	367	**240**	396	**300**	454	**399**	520	**464**
severely disrupted	336	**241**	402	**286**	484	**342**	524	**423**	556	**472**

Minimization of total secondary delay

infrastructure scenario	I20		I30		I40		I50		I60	
	NOR	3M	NOR	3M	NOR	3M	NOR	3M	NOR	3M
fully functioning	450	**228**	656	**317**	971	**510**	1444	**1052**	2154	**1970**
partially disrupted	642	**240**	921	**396**	1261	**636**	1786	**1112**	2655	**2146**
severely disrupted	809	**424**	1569	**748**	2375	**1125**	3266	**1711**	4322	**3331**

infrastructure scenario and objective function is reported in bold, and it always corresponds to 3M: allowing rerouting is always advantageous, and the SR formulation is able to deal with multiple route choices even when a short computation time is allowed. The difference in favor of 3M is always statistically significant, according to the Wilcoxon rank-sum test with a confidence level of 0.95.

As a further indication on the practical usefulness of the SR formulation, we assess the advantages brought by the possibility of modifying the platform assignment with respect to the scheduled one, always within a three minute computation. In particular, for each train, we consider three alternative platforms: the scheduled platform and the two adjacent ones. These two adjacent platforms are selected considering the feasible train-platform coupling in the practice: high speed trains can be assigned to only four platforms.

Indeed, the consideration of alternative platforms strongly increases the number of routes available to each train. As discussed when describing Table 2, the number of both binary variables and constraints grows more than proportionally with respect to the number of train routes. Hence, including alternative platforms strongly increases the instance size.

To assess the advantages offered by platform modification while trying to eliminate the impact of this increase of the instance size, we perform the experiments with two different setups: one the one hand, we allow each train to use all the routes connecting its origin or destination line to one of the three alternative platforms; on the other hand, for each train, we randomly select a few routes, as many as are available to the train itself when no platform modification is possible, guaranteeing a fair distribution of this number of routes between the three alternative platforms. We will refer to the first setup as MOD-ALL-R and to the second as MOD-FEW-R. We compare the results achieved through

Transportation Research Part B: Methodological, Volume 59, January 2014, p58–80, © 2013 Elsevier, http://dx.doi.org/10.1016/j.trb.2013.10.013

Table 8: Mean secondary delay across the 30 instances of each set achieved by the SR formulation within three minute computation. Comparison between three setups: no platform modification (FIX), platform modification with all routes (MOD-ALL-R) and platform modification with a few routes (MOD-FEW-R). The best mean result for each set and infrastructure scenario is reported in bold. A superscript letter indicates which setup, if any, is significantly better according to Wilcoxon rank-sum test with a confidence level of 0.95: the letter a concerns the comparison between FIX and MOD-ALL-R, and the letter f concerns the comparison between FIX and MOD-FEW-R.

Minimization of maximum secondary delay

	fully functioning			partially disrupted			severely disrupted		
	MOD-ALL-R	MOD-FEW-R	FIX	MOD-ALL-R	MOD-FEW-R	FIX	MOD-ALL-R	MOD-FEW-R	FIX
I20	**169** **a**	170 f	182	**170** **a**	172 f	184	**186** **a**	213 f	241
I30	232	**202** **f**	229	**211** **a**	217 f	240	**233** **a**	254 f	286
I40	301	**257**	263 a	**276** **a**	289 f	300	**305** **a**	316 f	342
I50	418	**355**	372 a	401	**345** **f**	399	369 a	**356** **f**	423
I60	487	467	**436** **af**	518	484	**464** **af**	501	**465**	472

Minimization of total secondary delay

	fully functioning			partially disrupted			severely disrupted		
	MOD-ALL-R	MOD-FEW-R	FIX	MOD-ALL-R	MOD-FEW-R	FIX	MOD-ALL-R	MOD-FEW-R	FIX
I20	208 a	**197** **f**	228	**205** **a**	207 f	240	**250** **a**	321 f	424
I30	432	**295**	317 a	407	**325** **f**	396	**402** **a**	551 f	748
I40	853	653	**510** **af**	883	641	**636** **a**	**712** **a**	870 f	1125
I50	1414	**1028**	1052 a	1545	1326	**1112** **af**	2165	1805	**1711**
I60	2133	2108	**1970** **af**	2607	2534	**2146** **af**	3976	4082	**3331** **af**

these setups to the ones achieved by the SR formulation with no platform modification possibility, which we will refer to as FIX.

Table 8 reports the results of this analysis. The table is divided in two parts, for the results achieved when minimizing maximum and total secondary delay, respectively. For each of these parts, the three infrastructure scenarios are considered separately, as well as the set of instances. Each table element is the mean secondary delay (either maximum or total) achieved on the 30 instances of a set for an infrastructure scenario by the SR formulation within three minutes in one of the three considered setups: FIX, MOD-ALL-R and MOD-FEW-R. The best mean secondary delay for each set of instances and infrastructure scenario is reported in bold. Moreover, the table shows the results of a statistical significance test on pairs of setups, according to Wilcoxon rank-sum test with

Transportation Research Part B: Methodological, Volume 59, January 2014, p58–80, © 2013 Elsevier, http://dx.doi.org/10.1016/j.trb.2013.10.013

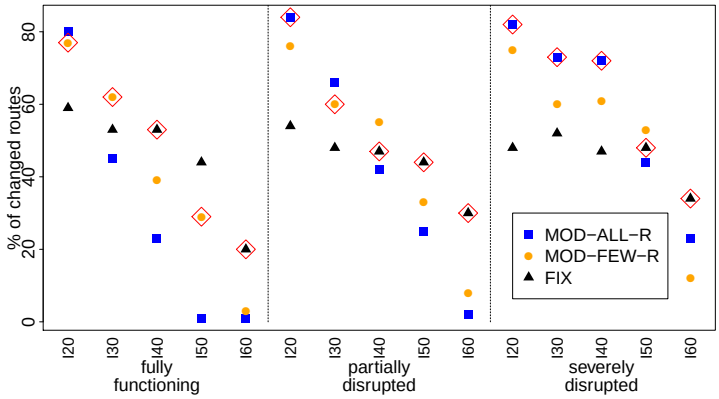

Figure 12: Mean percentage number of rerouted trains by MOD-ALL-R, MOD-FEW-R and FIX within three minute computation. The diamond shows which algorithm got the best mean performance.

a confidence level of 0.95. In particular, we test for significance the difference between FIX and MOD-ALL-R and between FIX and MOD-FIX-R. We indicate with the letter "a" the result of the first test, and with the letter "f" the result of the second. Such a letter follows as a superscript the mean secondary delay achieved by the significantly best setup. For example, in the fully functioning infrastructure scenario, when minimizing the maximum secondary delay for instances of set I20, FIX finds solutions that imply a mean maximum secondary delay of 182 seconds. The corresponding mean value for MOD-ALL-R is 169 seconds, while the one for MOD-FEW-R is 170 seconds. In both cases, the result achieved when considering platform modification is better than the one obtained with fix platforms, and the difference is statistically significant in both cases. Hence, a letter "a" follows the element corresponding to MOD-ALL-R, and a letter "f" follows the element corresponding to MOD-FEW-R. On the instances of set I30, instead, the difference is statistically significant in favor of MOD-FEW-R, while no significance is remarkable for the case of MOD-ALL-R. The meaning of this significance is that, if we consider further instances with similar characteristics to the ones belonging to set I30, we can expect to achieve a better solution through MOD-FEW-R than through FIX. Instead, there is no reason to expect MOD-ALL-R to be better than FIX, or viceversa.

In general, the best results are achieved by MOD-FEW-R: for the large majority of the observations it achieves better mean results than FIX, or they are statistically indistinguishable. Set I40 in the fully functioning scenario and total delay minimization, set I50 in the partially disrupted scenario and total delay minimization, and set I60 disregard the infrastructure scenario and the objective function are exceptions: FIX is significantly better than MOD-FEW-R in statistical terms. The table shows how the comparison is more favorable to MOD-FEW-R when smaller instances are tackled, or when fewer routes are available in the infrastructure due to disruption. A similar trend characterizes the comparison between FIX and MOD-ALL-R, even if FIX outperforms MOD-ALL-R in many more cases than MOD-FEW-R.

Transportation Research Part B: Methodological, Volume 59, January 2014, p58–80, © 2013 Elsevier, http://dx.doi.org/10.1016/j.trb.2013.10.013

Figure 12 shows the mean percentage number of trains which are rerouted in the best solution returned by MOD-ALL-R, MOD-FEW-R and FIX minimizing total delay. This figure is conceptually the same as Figure 10, and the triangles correspond to the same setup, apart from the fact that here a three minute computation time limit is imposed. Remark that the scale of the two figures is different for coping with the different data series to be represented. The biggest difference between the two series of triangles concerns the case of set I60 in the three infrastructure scenarios: within the short time limit, the SR formulation manages to reroute a much lower percentage of trains.

In Figure 12, a diamond highlights the result corresponding to the best mean performance for each set of instances and infrastructure scenario, shown in bold in Table 7. Such a best performance corresponds almost regularly to the setup which reroutes the largest percentage of trains. When this is not the case, the result corresponding to the best performing setup is the second highest, and we can distinguish two cases: either the difference with respect to the highest value is very small, as for set I20 in the fully functioning scenario, or this difference is not negligible and the results summarized in Table 7 actually could not indicate a statistically significant difference between the two concerned setups, as for the set I50 in the fully functioning scenario. In brief, then, the figure shows that the best performance is achieved when rerouting is maximally exploited. However, the figure also shows that the capability of profitably rerouting trains strongly diminishes with the increase of the number of trains involved in the search, i.e., with the size of the horizon considered in each instance, and with the diversity of the routes available to each train, i.e., using different platforms or not. An extreme behavior in this sense is achieved by MOD-ALL-R, for example, on sets I50 and I60 in the fully functioning infrastructure scenario. Here, the search space becomes so large that the optimization process does not manage to find, within the time limit, any feasible solution after the one excluding rerouting returned by the first optimization step (Section 6). The number of instances in which this inability of improving the no-rerouting solution increases with the number of trains involved: few instances start to be unmanageable for MOD-ALL-R in set I30, I50 and I60 in the fully functioning, partially disrupted and severely disrupted scenario, respectively. The impact of the number of available routes which the setup has to deal with and of the route diversity is evident since, when we decrease this number, the number of trains that becomes manageable within the time limit increases: for MOD-FEW-R the difficult instances in the three infrastructure scenarios start with I40, I50 and I60; for FIX the difficulty only concern some instances of set I60.

In summary, these results suggest that the increase of the instance size, and hence the possibility of modifying the platform assignment, is positively manageable up to a certain level, but it may become an issue when the combination of trains and available routes is too large, i.e., when too many potential conflicts are to be managed.

In practical implementations, then, the choice of whether to allow platform modification should be made as a function of the specific situation to be tackled. Such a choice actually characterizes also the current practice: dispatchers prefer not to modify platform assignments, unless in specific situations which they typically are able to identify *a priori*. This *a priori* identification may be translated into automatic triggers capable of detecting when platform modification may actually be advantageous in terms of optimization potentials.

Furthermore, an important pre-processing that could be implemented would be the

Transportation Research Part B: Methodological, Volume 59, January 2014, p58–80, © 2013 Elsevier, http://dx.doi.org/10.1016/j.trb.2013.10.013

reduction of the set of routes available to each train. The performance of MOD-FEW-R shows that decreasing the number of these available routes may allow a more efficient exploitation of the optimization capability of the SR formulation. However, in absence of an accepted rule for selecting which routes are more suitable to be eliminated, we randomly selected the available ones. On the one hand, this is a choice that avoids the bias due to an arbitrary decision. On the other hand, this possibly eliminates options which may be potentially advantageous based a study of the topology of the infrastructure. However, the basic principles of such a study would need to be agreed and justified.

9 Conclusions

In this paper, we proposed a mixed-integer linear programming formulation for the rtRTMP. It solves instances in which the infrastructure is represented with fine granularity. Through this granularity, the formulation can model either the route-lock sectional-release or the route-lock route-release interlocking system. Moreover, it selects among all possible train routes which can be practically exploited and it considers all possible train scheduling.

We applied this formulation to random instances representing traffic in the infrastructure named triangle of Gagny, and to instances obtained by perturbing the real timetable of a week day in the control area including the Lille-Flandres station, in France. For the latter, we considered multiple scenarios in terms infrastructure availability. Moreover, we considered two alternative objective functions to be minimized, namely the maximum secondary delay suffered by any train and the total secondary delay.

In this thorough experimental analysis we assessed the impact of the granularity of the representation of the control area, where the rough granularity corresponds to the route-lock route-release interlocking and the fine granularity to the route-lock sectional-release one. Our results show that modeling a rough granularity worsens the optimal solution quality. In other word, the solution chosen with rough granularity implies longer delay suffered by trains than the solution chosen with the fine granularity; this difference is statistically significant.

The fixed-speed model which we use for the comparison of the results is indeed an approximation. As often mentioned in the literature when this type of comparison is made (e.g., D'Ariano et al. (2008); Corman et al. (2010, 2012a,b)), the validity of the quantification of the differences would increase if speed variation dynamics were considered. We reckon the inclusion of speed variation dynamics as a future research topic.

The computation time for solving instances representing 20 minutes of traffic at peak-time is in line with its real-time purposes. When the duration of the time interval considered for each instance increases, the computational time for finding the optimal solution and proving its actual optimality becomes excessively long. In future research, we will propose possible methods for reducing the computation time required for solving large complex instances, through the insertion of valid inequalities and the tuning of the parameters of the CPLEX solver.

However, on an additional set of experiments, we showed that the proposed formulation can positively tackle all the instances considered even when a short computation time is allowed. In this case, the optimal solution may not be found or proved, but a

Transportation Research Part B: Methodological, Volume 59, January 2014, p58–80, © 2013 Elsevier, http://dx.doi.org/10.1016/j.trb.2013.10.013

feasible good quality solution is always returned. In this additional set of experiments, we also proposed an analysis on the advantages offered by the possibility of modifying the platform assignment. The results suggest that the increase of complexity that derives from this possibility does not necessarily allow a full exploitation of the optimization capability of the proposed formulation.

In the future, we will exploit the formulation proposed here for further analyzing and quantifying the impact of decisions which are to be taken *a priori* when performing the optimization, as for example the application of the optimization itself in a rolling-horizon framework.

References

Caimi, G., Chudak, F., Fuchsberger, M., Laumanns, M., and Zenklusen, R. (2011). A new resource-constrained multicommodity flow model for conflict-free train routing and scheduling. *Transportation Science*, 45(2):212–227.

Caimi, G., Fuchsberger, M., Laumanns, M., and Lüthi, M. (2012). A model predictive control approach for descrete-time rescheduling in complex central railway station approach. *Computers & Operations Research*, 39:2578–2593.

Corman, F., D'Ariano, A., Pacciarelli, D., and Pranzo, M. (2009a). Evaluation of green wave policy in real-time railway traffic management. *Transportation Research Part C*, 17:607–616.

Corman, F., D'Ariano, A., Pacciarelli, D., and Pranzo, M. (2010). A tabu search algorithm for rerouting trains during rail operations. *Transportation Research Part B*, 44:175–192.

Corman, F., D'Ariano, A., Pacciarelli, D., and Pranzo, M. (2012a). Bi-objective conflict detection and resolution in railway traffic management. *Transportation Research Part C*, 20:79–94.

Corman, F., D'Ariano, A., Pacciarelli, D., and Pranzo, M. (2012b). Optimal inter-area coordination of train rescheduling decisions. *Transportation Research Part E*, 48:71–88.

Corman, F., Goverde, R., and D'Ariano, A. (2009b). Rescheduling dense traffic over complex station interlocking areas. In Ahuja, R., Möhring, R., and Zaroliagis, C., editors, *Robust and Online Large-Scale Optimization: Models and Techniques for Transportation Systems*, volume 5868 of *Lecture Notes in Computer Science*, pages 369–386, Berlin, Germany. Springer Berlin / Heidelberg.

D'Ariano, A., Corman, F., Pacciarelli, D., and Pranzo, M. (2008). Reordering and local rerouting strategies to manage train traffic in real-time. *Transportation Science*, 42(4):405–419.

D'Ariano, A., Pacciarelli, D., and Pranzo, M. (2007a). A branch and bound algorithm for scheduling trains in a railway network. *European Journal of Operational Research*, 183:643–657.

Transportation Research Part B: Methodological, Volume 59, January 2014, p58–80, © 2013 Elsevier, http://dx.doi.org/10.1016/j.trb.2013.10.013

D'Ariano, A. and Pranzo, M. (2009). An advanced real-time train dispatching system for minimizing the propagation of delays in a dispatching area under severe disturbances. *Networks and Spatial Economics*, 9:63–84.

D'Ariano, A., Pranzo, M., and Hansen, I. (2007b). Conflict resolution and train speed coordination for solving real-time timetable perturbations. *IEEE Transactions on Intelligent Transportation Systems*, 8(2):208–222.

Dessouky, M., Lu, Q., Zhao, J., and Leachman, R. (2006). An exact solution procedure to determine the optimal dispatching times for complex rail networks. *IIE Transactions*, 38(2):141–152.

IBM Corporation (2012). User's manual for cplex. http://publib.boulder.ibm.com/infocenter/cosinfoc/v12r2.

Lusby, R., Larsen, J., Ehrgott, M., and Ryan, D. (2012). A set packing inspired method for real-time junction train routing. *Computers & Operations Research*.

Mazzarello, M. and Ottaviani, E. (2007). A traffic management system for real-time traffic optimisation in railways. *Transportation Research Part B*, 41:246–274.

Pachl, J. (2002). Spacing trains. In *Railway Operation & Control*, chapter 3, pages 38–90. VTD Rail Publishing, Mountlake Terrace, WA, USA.

Pellegrini, P., Marlière, G., and Rodriguez, J. (2012). Real Time Railway Traffic Management Modeling Track-Circuits. In D., D. and L., L., editors, *12th Workshop on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems*, volume 25 of *OpenAccess Series in Informatics (OASIcs)*, pages 23–34, Dagstuhl, Germany. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

Pellegrini, P., Marlière, G., and Rodriguez, J. (2013). A mixed-integer linear program for the real-time railway traffic management problem modeling track-circuits. In *5th International Seminar on Railway Operations Modelling and Analysis, RailCopenhagen 2013, Copenhagen, Denmark*.

Rodriguez, J. (2007). A constraint programming model for real-time train scheduling at junctions. *Transportation Research Part B*, 41:231–245.

Theeg, G., Maschek, U., and Nasedkin, O. (2009). Interlocking principles. In Theeg, G. and Vlasenko, S., editors, *Railway Signalling & Interlocking. International Compendium*, pages 61–112. Eurail press, Hambourg, Germany.

Törnquist, J. (2007). Railway traffic disturbance management - An experimental analysis of disturbance complexity, management objectives and limitations in planning horizon. *Transportation Research Part A*, 41:249–266.

Törnquist, J. and Persson, J. (2007). N-tracked railway traffic re-scheduling during disturbances. *Transportation Research Part B*, 41:342–362.

Transportation Research Part B: Methodological, Volume 59, January 2014, p58–80, © 2013 Elsevier, http://dx.doi.org/10.1016/j.trb.2013.10.013

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

U.I.C. (2004). Leaflet 406 "capacity".