

STEVEN HARROD THOMAS SCHLECHTE

A Direct Comparison of Physical Block Occupancy Versus Timed Block Occupancy in Train Timetabling Formulations

Herausgegeben vom
Konrad-Zuse-Zentrum für Informationstechnik Berlin
Takustraße 7
D-14195 Berlin-Dahlem

Telefon: 030-84185-0
Telefax: 030-84185-125

e-mail: bibliothek@zib.de
URL: <http://www.zib.de>

ZIB-Report (Print) ISSN 1438-0064
ZIB-Report (Internet) ISSN 2192-7782

A Direct Comparison of Physical Block Occupancy Versus Timed Block Occupancy in Train Timetabling Formulations

Steven Harrod** Thomas Schlechte*

April 5, 2013

Abstract

Two fundamental mathematical formulations for railway timetabling are compared on a common set of sample problems, representing both multiple track high density services in Europe and single track bidirectional operations in North America. One formulation, ACP, enforces against conflicts by constraining time intervals between trains, while the other formulation, RCHF, monitors physical occupation of controlled track segments. The results demonstrate that both ACP and RCHF return comparable solutions in the aggregate, with some significant differences in select instances, and a pattern of significant differences in performance and constraint enforcement overall.

1 Introduction

The railway timetabling literature contains many distinct mathematical formulations for the problem of scheduling multiple trains within a limited capacity railway line or network of lines. Some of this literature is twenty, thirty, and even forty years old. The majority of these models remain academic novelties, and only a few have seen any commercial application. A frequently cited obstacle to the application of mathematical programming to railway timetabling is that computing power is insufficient to resolve commercially relevant problems in an acceptable time frame.

A number of publications and case studies in the last decade demonstrate that this obstacle is falling away, particularly due to the stunning increase in addressable memory in common office computers, and the general improvement in algorithms and heuristics (Bixby 2002). As the application of

**University of Dayton, School of Business Administration, 300 College Park Dayton, OH 45469, Email steven.harrod@udayton.edu

*Zuse-Institute Berlin (ZIB), Takustrasse 7, 14195 Berlin-Dahlem, Germany, Email schlechte@zib.de

mathematical programming to train scheduling is imminent, it is appropriate to evaluate in greater detail how various formulations differ from each other in objective, operating rule enforcement, and performance. Most of the prior literature on train timetabling is problem or scenario specific. That is, the typical paper focuses on solving a particular scheduling problem.

This paper contributes to the literature a direct comparison of two distinct mathematical formulations for train timetabling that have both previously appeared in the literature independently. These two formulations are representative of two fundamental alternatives in operating rule constraint formulation. One general class of train timetabling formulation enforces a timed separation between scheduled trains, and the other class enforces a physical separation between trains. Under many scenarios and at an aggregate level, these two formulations return comparable results in twenty-six out of forty-one problem instances. However, as scenarios are evaluated in closer detail, differences in calculation speed and operating rule enforcement are revealed in the majority of problem instances. In addition, this paper demonstrates feasible methods by which two popular timetable data structures may be compared.

The evaluation consists of a two by two experimental control consisting of the above described formulations and two distinct groups of scheduling problems. The first problem group consists of a classical high density “European” passenger network, and the second problem group consists of a low density “North American” single track line. The characteristic European scheduling problem consists of hundreds of individual train schedules distributed over a high capacity network of multiple track lines (each track designated for a single direction of flow). In addition to operating and safety rules, these problems also frequently contain requirements of specific stopping times for individual trains at stations. In contrast, the characteristic North American scheduling problem considers a smaller set of trains distributed over a significantly more limited set of tracks. In the North American problem, tracks are bidirectional, and operating rule enforcement must consider the arbitration between conflicting, opposing direction trains.

The paper is organized as follows: § 2 gives a literature overview on railway capacity studies, § 3 describes a problem formulation for the *train pathing problem*, and § 4 presents two distinct integer programming formulations for consideration. Section 5 describes two alternative data formats, how to convert between these data formats, and provides one example of each. Finally, the strengths and weaknesses of each formulation as exhibited by trials on these data sets are reviewed in § 5.6.

2 Literature

Very few reports known to the authors compare formulations under identical data sets or problem scenarios. Kraft (1988) compares three analytical capacity formulas and two simulation models upon common data sets, but these are not mathematical programming formulations for scheduling. Mills et al. (1991) compares a discrete time interval scheduling formulation to a nonlinear, real valued time formulation within the same data set, and finds the nonlinear formulation offers very small improvement over the discrete time formulation, and requires a much greater solution time.

A number of surveys are available which describe the formulations of published models and organize them by selected characteristics. Harrod (2012) reviews formulations specifically relevant to timetable generation and optimization, and groups formulations according to their decision variable and constraint structure. Lusby et al. (2011) categorizes formulations according to the track structure they represent (single, double, or station tracks). Cacchiani and Toth (2012) provide a detailed review of a number of formulations for both deterministic scheduling and for robust scheduling (stochastic programming). Törnquist (2006) reviews 48 published formulations and includes detailed comparisons by network constraints, objective, solution method, and problem size. Cordeau et al. (1998) conducts a broader detailed survey of railroad optimization in all aspects: blocking, strategic planning, and dispatching. Bussieck et al. (1997) evaluates formulations particularly relevant to passenger railways in Europe, and periodic scheduling in particular. Assad (1980) provides an excellent reference for timetabling literature prior to 1980.

3 Problem Description

The scheduling problems evaluated in this paper are variously labeled “master scheduling”, “track allocation” or “timetabling”, but to be very specific this paper considers the *train pathing* problem. This is the scheduling process that determines the route or path of a train upon the railway network such that departure and arrival constraints, limits on available track, and conflicts with other trains are observed. It is not to be confused with various aggregate scheduling problems that defer the pathing and arbitration between conflicting trains to a later time period or authority. In particular, in North America a timetable rarely asserts a previously assigned train path. Instead, in North America the routing of a train on the track structure is left to the decision of dispatchers on the day of operation, as a tactical operating decision, even though schedule commitments may exist. Thus to avoid confusion, the equivalent term *Train Timetabling Problem* (TTP), which originates with Caprara et al. (2002), is not used in this paper.

Region	Pass. km./ Ton-km	Max. Train Length m.	Coal Traffic	Total Freight Share
Europe	0.15	750	23%	39%
North America	0.004	2,290	1%	17%

Table 1: Comparison of European and North American Networks (Vassallo and Fagan 2007, International Union of Railways 2008, Bureau of Transportation Statistics 2009, CER 2008)

Train pathing occurs at multiple planning events on the railway. It occurs at an early strategic time when service timetables are composed and public service commitments made. It occurs again closer to the day of operation when additional “extra” trains are included, such as maintenance trains and extra trains for random traffic increases, see e.g., Flier et al. (2009) and Cacchiani et al. (2010b).

Finally, train pathing occurs in response to delays and disruptions such as weather, equipment failure, and network emergencies. This is a tactical operating decision. The majority of this train pathing is still performed manually.

This paper considers two train pathing problem classes, one based upon a European scheduling problem and the other based upon a North American scheduling problems. Together these two scenarios represent the extremes of the full range of railway operating patterns and densities worldwide. The European railway network is characterized by multiple track lines carrying dense, high frequency traffic dominated by passenger trains. In contrast, 70% of the North American route network is single track (Richards 2006). Table 1 provides key network statistics for comparison.

The European and North American rail networks are both congested and busy in their own way. The European network is recognized for its dense traffic on multiple track lines at very close headways. The North American network is also congested on a different scale, because although the gross train frequency is not comparable, the underlying track infrastructure is severely restricted.

On an abstract level all problem formulations for the Train Pathing Problem (TPP) can be defined as follows:

Definition 1. (*Train Pathing Problem (TPP)*) *Let G be a (railway infrastructure) graph and I a set of trains, then define TPP as the problem of choosing conflict-free paths through G for each train $i \in I$ with maximal profit.*

The definition of “conflicts” is always related to the occupation of infrastructure resources in space and time. Due to the various safety systems there are many different ways of enforcing and specifying operational con-

licts and defining the discrete resources, e.g. set of *blocks* B or *tracks* J . Safety separation between trains is enforced by constraints that limit the simultaneous occupation of adjoining track segments. In North American practice these track segments are called *blocks* and are defined by the discrete control limits of the signal system. The same principle is used in Europe with the difference that in the planning stage an aggregation of several physical succeeding blocks called *tracks* is considered and hence the physical safety distance between trains is aggregated to time separation at the departure of the track. Basically, all problem formulations of the train pathing problem have one thing in common: a solution to TPP must specify the complete route of each train i through a network G and on the departure and arrival times at each control point of that route. Table 2 lists the used symbols for the TPP.

Table 2: Sets or Parameters for problem TPP

Symbol	Description
I	The set of requested trains
J or B	The set of network resources, e.g. tracks or blocks
G	The railway infrastructure graph

4 IP Formulations

Two formulations are compared in this paper, one originating with Borndörfer and Schlechte (2007) and another originating with Harrod (2011). The first one, the Arc Configuration Problem (ACP), is an extended formulation of classical packing formulations introduced by Brännlund et al. (1998) and Caprara et al. (2006). See also Cacchiani et al. (2010a) for a detailed analysis of alternative packing formulations for timetabling under timed intervals between trains. This formulation enforces against conflicts between arcs that represent discrete train moves by using artificial graphs for each block or track. The second formulation aggregates points of conflict by modeling each arc as a hyperarc on a hypergraph. Conflicts are then indexed not by pairs of arcs, but by unique conflict nodes shared by two or more arcs. Each hyperarc represents a train move from an origin node to a destination node and the arc also contains all other potential conflicting nodes. A major difference between both models is that the hyperarc model is very flexible with respect to the definition of conflicts, i.e., an almost arbitrary (but in some resource sense connected) set of hyperarcs can be restricted. In contrast, model ACP assumes an a priori decomposition of conflicts with respect to tracks

and utilizes the structure of given minimal headway times for all potential combinations of trains.

4.1 ACP formulation

The classical way of formulating the train pathing problem is by using single flow formulations for the trains which are coupled by additional packing constraints to exclude conflicts between those trains, see Brännlund et al. (1998) or Caprara et al. (2006). Borndörfer and Schlechte (2007) proposes an alternative formulation for the train pathing problem that handles conflicts by additional variables. This arc coupling formulation (ACP) is based on the concept of feasible arc *configurations*, i.e., sets of time-expanded arcs on a track without occupation conflicts. Given is a standard time expanded scheduling graph D , in which time expanded paths represent train routes and departure and arrival times for the individual train requests $i \in I$ in the subgraph $D_i \subseteq D$. Operational feasibility with respect to headway times between two consecutive trains is handled by artificial digraphs D_j for each $j \in J$. Each two train moves on track j which respect the minimal headway time are connected by artificial arcs in D_j . Section 5.3 shows how physical block separation rules are transformed to headway times. Table 3 lists all relevant objects for the ACP formulation.

Figure 1 shows the construction of a track digraph respecting headway times for successions of trains. On the left hand side the requested train moves on track $j = (u, v)$ are shown in a time expanded setting, i.e. six trains belonging to set A_Ψ^j with three different running times. Note that all of these arcs are binary variables x_a . On the right hand side introduce artificial dotted arcs which determine the first or last departure of a train on track j . In addition, each feasible consecutive succession of two trains are connected by an artificial dashed arc, i.e., if the difference between both departure times is larger or equal to the given headway time of that train combination. Hence, graph D_j models all potential orderings of the departures of trains on track j . Note that all arcs of D_j use binary variables y_a . In case of a transitive headway matrix, paths in D_j represent valid consecutive successions of discrete train moves on track j . Note that each arc representing a train move is contained in exactly one D_j and D_i and induces two variables x_a and y_a . More details on the construction of D_i and D_j can be found in Schlechte (2012).

Variables x_a , $a \in A_i$, $i \in I$ control the use of arc a in D_i and y_b , $b \in A_j$, $j \in J$ in D_j , respectively. Note that only arcs which represent track usage, i.e., $A_S = A_I \cap A_J$, have to be coupled in that formulation by two variables x_a and y_b . Finally, the ACP formulation reads as:

Table 3: Sets or Parameters of formulation ACP for problem TPP

Symbol	Description
$D_i = (V_i, A_i)$	The time expanded graph of train i
s_i	The artificial source node of train digraph D_i
t_i	The artificial sink node of train digraph D_i
$D_j = (V_j, A_j)$	The artificial track digraph of track j
s_j	The artificial source node of track digraph D_j
t_j	The artificial sink node of track digraph D_j
w_a	The profit weight of arc a
C	The set of all station conflict sets
A_c	The set of arcs belonging to station conflict $c \in C$
κ_c	The upper limit or capacity of station conflict $c \in C$
A_S	The set of arcs (representing train moves) to couple

(ACP)

$$\max \sum_{a \in A} w_a x_a \quad (i)$$

$$s.t. \quad \sum_{a \in \delta_{\text{out}}^i(v)} x_a - \sum_{a \in \delta_{\text{in}}^i(v)} x_a = 0, \quad \forall i \in I, v \in V_i \setminus \{s_i, t_i\} \quad (ii)$$

$$\sum_{a \in \delta_{\text{out}}^i(s_i)} x_a \leq 1, \quad \forall i \in I \quad (iii)$$

$$\sum_{a \in \delta_{\text{out}}^j(v)} y_a - \sum_{a \in \delta_{\text{in}}^j(v)} y_a = 0, \quad \forall j \in J, v \in V_j \setminus \{s_j, t_j\} \quad (iv)$$

$$\sum_{a \in \delta_{\text{out}}^j(s_j)} y_a \leq 1, \quad \forall j \in J \quad (v)$$

$$x_a - y_a = 0, \quad \forall a \in A_S \quad (vi)$$

$$\sum_{a \in A_c} x_a \leq \kappa_c, \quad \forall c \in C \quad (vii)$$

$$x_a, y_a \in \{0, 1\}, \quad \forall a \in A_I, a \in A_J. \quad (viii)$$

The objective, denoted in ACP (i), is to maximize the weight of the track allocation, which can be expressed as the sum of the arc weights w_a .

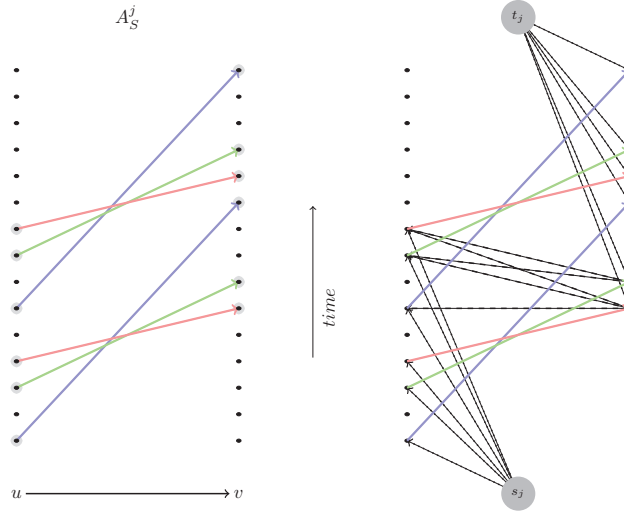


Figure 1: Example for the Construction of a Track Digraph.

Equalities (ii) and (iv) are well-known *flow conservation constraints* at intermediate nodes for all trains flows $i \in I$ and for all flows on tracks $j \in J$, (iii) and (v) state that at most one flow, i.e., train and track, unit is realized. Equalities (vi) link arcs used by train routes and track configurations to ensure a conflict-free allocation on each track individually. Packing constraints (vii) ensures that no station capacity is violated in its most general form. Finally, (viii) states that all variables are binary.

4.2 Hypergraph Formulation

The hypergraph model represents train movements as decision variables of sequential occupancy, or *transition*, between two controlled track segments, over an interval of one or more discrete time units. The controlled track segments are individually indexed by discrete time unit over the planning horizon, and form nodes within a time expanded graph. Additional nodes define zones of transition conflict between adjacent track segments. Let R denote the set of resource nodes either from occupation or transition of an hypergraph.

A hypergraph is a graph in which the definition of an edge is expanded to include any non-empty subset of nodes (see Gallo et al. (1993) for a formal definition). In contrast to most of the hypergraph literature, it is convenient in this setting to conceive a hyperarc $a = (u, v)$ as a set of tail nodes $T(a)$ with distinct entry tail $u \in T(a)$ and exactly one head node v .

Definition 2. (TPP hypergraph) Let R be a set of resource nodes and

$A \subseteq 2^R \times R$ a set called hyperarcs, then the TPP hypergraph is denoted by $H = (R, A)$.

In the case of a railway line, a track segment has frequently been modeled as an arc with nodes at the endpoints of the track segment. However, occupation of a track segment in practice may conflict other adjacent track segments. A discrete time graph model of the network likely will also require a train movement decision variable to register the occupation of multiple graph nodes each representing a labeled network resource for a discrete time unit. Thus a hypergraph provides the flexibility to encapsulate all of the variety of implied conflicts of a train movement variable. The binary decision variables are hyperarcs on this graph that potentially enclose multiple nodes.

The TPP can be seen as a *resource constrained hyperflow* (RCHF) in H . The resources basically limit the block occupation, the transitions between track segments, and the physical headways between trains. All of these features are presented in detail in Harrod (2011). For comparison here, all of these conflict categories are consolidated to a corresponding set of conflicting hyperarcs A_c for each conflict $c \in C$. Note that, parallel segments such as sidings or multiple through track are defined as a single block with a total capacity and hence an upper bound on the conflict set $\kappa_c \geq 1$. A detailed variant of the model and the sets A_c can be found in the Appendix A.

A hypergraph with its representative network of blocks is shown in Figure 2. Hollow nodes are track segment or block occupancies, while solid nodes represent potential transition conflicts between blocks, and are only defined where paths represent separate trains. Note how the transition conflict between the two labeled hyperarcs is indicated. A sample conflict is depicted by the highlighted hyperarc representing a starting occupancy of block 2 from time $t + 4$ and reaching block 3 at time $t + 6$, and the hyperarc from block 3 at time $t + 5$ to block 2 at time $t + 6$, working exactly in the opposite direction. The conflicting part is the resource node representing the (undirected) transition between block 2 and 3 at time $t + 5(+\epsilon)$. In addition, denote by $s_i \in R$ and $t_i \in R$ an artificial source and sink node for train i .

The (RCHF) formulation is presented below. It differs from Harrod (2011) in that the objective and coefficients have been revised to correspond with those of the Borndörfer et al. formulation.

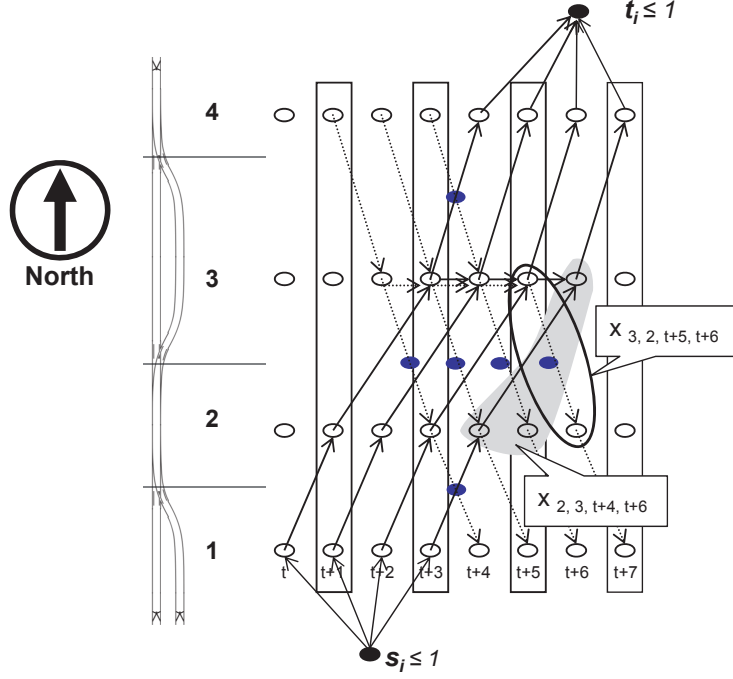


Figure 2: An example hypergraph of train paths with highlighted hyperarc variables $x_{2,3,t+4,t+6}$ and $x_{3,2,t+5,t+6}$ which do not conflict in block occupancy, but do conflict in transition node $(2, t+5)$.

(RCHF)

$$\max \sum_{a \in A} w_a x_a \quad (i)$$

$$s.t. \sum_{a \in \delta_{out}^i(s_i)} x_a \leq 1, \quad \forall i \in I \quad (ii)$$

$$\sum_{a \in \delta_{out}^i(v)} x_a - \sum_{a \in \delta_{in}^i(v)} x_a = 0, \quad \forall i \in I, \forall v \in R \setminus \{s_i, t_i\} \quad (iii)$$

$$\sum_{a \in \delta_{in}^i(t_i)} x_a \leq 1, \quad \forall i \in I \quad (iv)$$

$$\sum_{a \in A_c} x_a \leq \kappa_c, \quad \forall c \in C \quad (v)$$

$$x_a \in \{0, 1\}, \quad \forall a \in A. \quad (vi)$$

Objective (RCHF) (i) and constraints (ii)-(iv) define a set of independent single commodity flows. Resource constraints (RCHF) (iv) bind these flows together and regulate them as a multicommodity flow respecting the

limited resource. Note that constraints (RCHF) (ii) enforces a single departure for each train, and because it is an inequality also offers the option of removing the train from the schedule if it is not productive, while constraints (RCHF) (iii) enforces conservation of mass at the nodes (enforcing a single train path) and sinking constraints (RCHF) (iv) ensures a single terminal arrival. Resource constraints (RCHF) (v) enforces a conflict-free allocation of the resources, i.e., the block occupancy limit, the limits on transitions between track segments, and the limitation due to the headways or follow on spacing of trains. Finally, (RCHF) (vi) states that all hyperarc variables x_a have to be binary.

5 Sample Problems

Two sample data sets are considered, representing two typical, yet very different, operating problems. In each data set individual problems are composed by specifying requests for, possibly conflicting, train paths. Individual problems vary in size by the number of train paths requested and the portion of the network evaluated. The hypergraph formulation is solved in IBM ILOG CPLEX Optimization Studio 12.4.

The first data set originates with Borndörfer et al. (2006) and covers the network of lines connecting Fulda and Kassel to Hanover, Germany, and its suburbs. This Deutsche Bahn network consists of 1,460 km of route segments including a 250 km/h high speed line, various alternate lower speed routes, and suburban and terminal trackage. The data set presumes that all routes are double track and designated for one direction of flow. The train path requests in this data set are dominated by passenger traffic, but a wide variety of speeds and stopping patterns are represented.

The second data set originates with Harrod (2011) and describes an experimental bi-directional, single track line with equally spaced sidings. Documentation of a comparable service instance may be found in Holowaty et al. (2004) where a corridor from Richmond, Virginia to Charlotte, North Carolina is specified with sidings 7.5 miles (12 km) long spaced 15 miles (24 km) from center to center, or a block length of 7.5 miles where every alternate block contains a siding for its full length. The train path requests in this data set consist of two train types, a moderate speed freight train and a higher speed passenger train.

The primary challenge faced in comparing these data sets across formulations is how to translate the data format. Each data set originates with its respective formulation, and supporting data for the alternative formulation is not available. For example, data is not available for the Deutsche Bahn network to specify the physical track blocks, which are required for the hypergraph formulation. In each case, a reasonable plan of data translation has been specified and is described in the following sections.

5.1 European Sample Data

A map of the Fulda-Hannover network is displayed in Figure 3. As can be seen, alternate routes exist between many points. Close parallel lines represent that in some cases tracks in one direction between two endpoints have different timing characteristics than the matching track in the opposing direction. Track segments in the data set range from terminal tracks of 300 m. to a dedicated high speed rail segment between Fulda and Kassel of 93 km.

Two questions arise in converting the data set for solution under the hypergraph formulation. First, the hypergraph model requires as input data

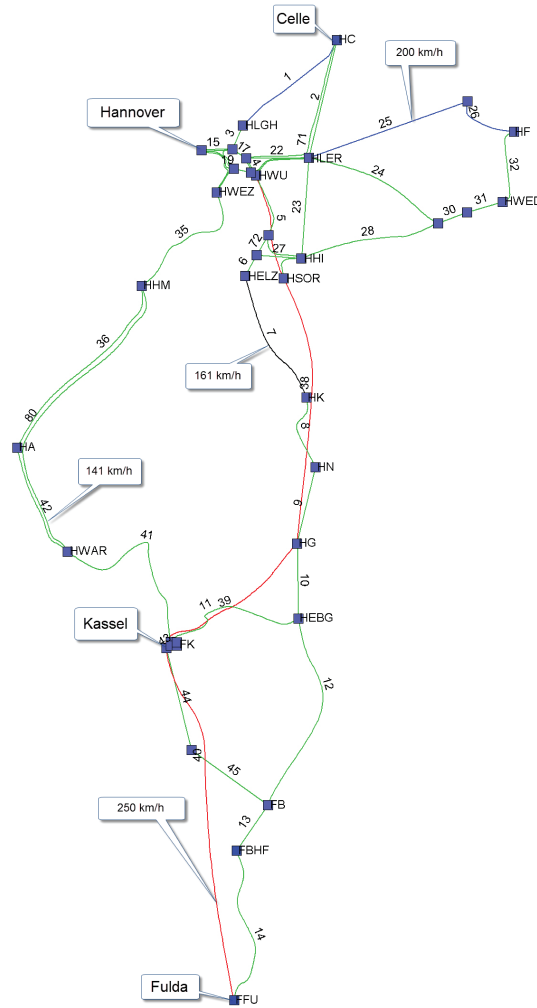


Figure 3: Map of Fulda-Hannover Region Rail Network

specification of all possible routes for a given train. In some cases path requests on the Fulda-Hannover network may take alternative routes. To limit problem complexity, each path request is assigned its minimum time route during pre-processing.

Second, the track segments must be divided into controlled blocks to support the train separation constraints of the hypergraph model. No detailed information on the true track block dimensions is available from Deutsche Bahn, so track segments are uniformly divided into blocks according to the following rule: tracks less than 3 km long are a single block, longer tracks less than 6 km are two equal blocks, tracks greater than 6 km are divided into equal blocks approximately 3.5 km long, and high speed (250 km/h) tracks are divided into equal blocks approximately 7 km long. All train paths in the hypergraph model enforce a two block physical separation between trains. In case of model formulation ACP the network consists of 37 stations and 120 track arcs. This translates to 870 track blocks in model RCHF.

5.2 North American Sample Data

A track diagram for representative single track problem 1400 is displayed in Figure 4. The experimental network is a single track line with long passing sidings at even numbered blocks, on which there is a dense flow of moderate speed traffic and a pair of mandatory higher speed trains which have priority dispatch. Nine network lengths, from $|B| = 3$ to $|B| = 19$ blocks long, provide nine trials.

The characteristics of the dense flow are given in Table 4. Any coefficient not displayed is zero. In addition, one roundtrip of a higher speed (3 time units per block) train, with no utility value ($c_p^r = 0$), is mandated by setting the corresponding source outflow constraint to an equality and has absolute priority over all other traffic. In each trial the total input train set is greater than the network capacity, and the solution will contain a subset of the available trains.

The last dispatched base flow train in each data set is priced such that any dispatch beyond the standard planning horizon pays a penalty that represents consumption of additional network time. This data set asserts traffic volume within the given planning horizon is the dominant objective.

5.3 On Block Separation and Headways

This section presents a transformation of railway systems operated with fixed block separation to a time separation formulation using headway times. The result is a formulation of the TPP in the format of the free available problem library TTPLib, see Erol et al. (2008) for the definition of the data structures.

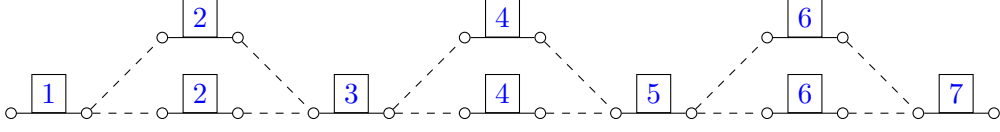


Figure 4: Seven Block Single Track Line With Sidings.

Visualize a single track line such as in Figure 4, where each track segment allows bidirectional movement and parallel tracks or sidings may exist. The numbered track segments represent specific limits of control for each track segment, called blocks. Generally, under traditional train control technology, locations of trains are defined by the block(s) they occupy, but no additional information clarifies the specific location of a train within the block. In order to translate a physical block separation constraint between pairs of trains into timed event separation between trains (in this case at “station” nodes), the network is segmented as follows:

Introduce for each block b two nodes to represent the entering and leaving of that block, denoted by l_b and r_b . Note for any adjoining, sequentially numbered blocks that $r_b \equiv l_{b+1}$. Due to the fact that each block may be traveled in either direction, define for each block a bidirectional track $e = (u, v)$ which consists of two directed arcs (u, v) and (v, u) with $u = l_b$ and $u = r_b$. In Figure 4 blocks two, four, and six contain two parallel tracks which in this case are a main line and a siding. The term *multiplicity* denotes the capacity of such a block, equivalent to the number of parallel, interchangeable tracks and the number of allowed simultaneous train occupancies. Of course, this value must be a natural number. To avoid overuse of the term “track”, individual tracks in a block are referred to as *lanes*.

The network equivalent of Figure 4 is shown in Figure 5. The marked or indicated seven blocks induce eight station nodes and twenty track arcs, i.e., $A = \emptyset$ and $|E| = 10$. From the left, node l_1 marks the left end of block one, and node r_1 the right end. The labels $a_{i,j}^{b,m}$ denote arcs representing movement in the direction from block i to block j and occupying lane m of block b . Block two begins at node r_1 and ends at node l_2 . It supports two lanes (multiplicity is two), and so there are two arcs in the direction of block three, $\{a_{2,3}^{2,1}, a_{2,3}^{2,2}\}$, and two arcs in the direction of block one, $\{a_{3,2}^{2,1}, a_{3,2}^{2,2}\}$. In addition the order of trains can only change “in” blocks with multiplicity larger than one, i.e., block 2, 4, and 6.

The mapping of the running times for leaving a block and entering a successive block is straight-forward, i.e., for $a_{u,v}^{b,m}$ define the running time for leaving the block associated with node i and entering the block associated with node j . Modeling the block separation by reasonable headways is more complex. Assuming a block headway of 2 intervening blocks, it is necessary to make sure that the next two blocks the train will enter are not allocated.

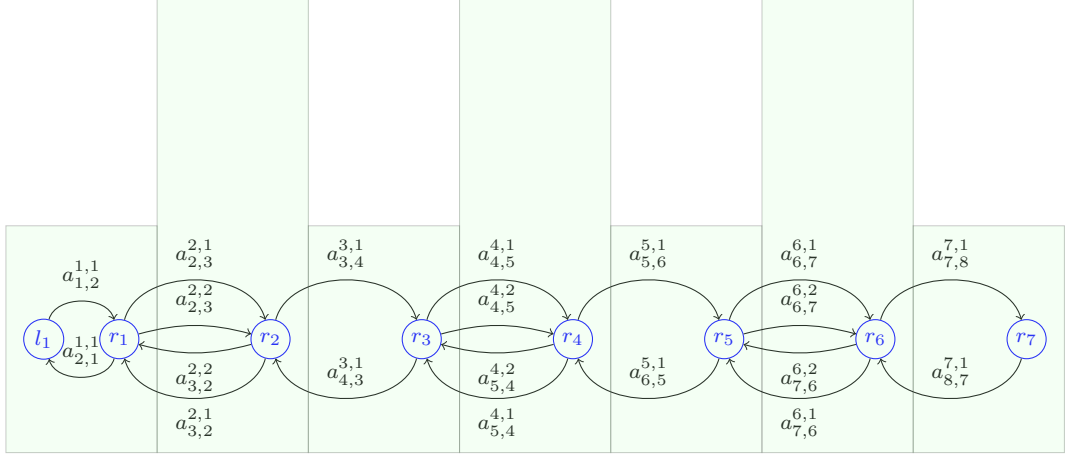


Figure 5: Network after transformation.

For example in Figure 5 the situation is as follows: Assume train p wants to run from l_1 to node r_3 (from block 1 to block 3) as well as train q . Then the minimal headway time $h_{p,q}$ for arc $a_{1,2}^{1,1}$ is the running time for train p on track $a_{1,2}^{1,1}$ and on the next track $a_{2,3}^{2,m}$ and vice versa. The same holds for $a_{2,3}^{1,m}$ and his successor $a_{2,4}^{1,m}$ with $m = 1, 2$. In case of various successor blocks, chose the maximum time to enforce conservative headway times, to enforce two block separation in any case. Due to the fact that the running times of all lanes are the same and the order in which a train traverses the blocks is also fixed, it is possible to compute $h_{p,q}$ for each track arc a priori.

This formulation is intended to enforce locally conflict free scheduling of running trains on blocks without specific knowledge of the physical geography of track blocks. In enforcing overtaking moves for trains, it is necessary to make sure that the number of trains on a block does not exceed the capacity. For example, in 4 the capacity of the block is one or two and therefore the capacity of the nodes is one or two with respect to their adjacent blocks, respectively. This will also avoid potential dead-lock situations in which two trains are facing each other. However, due to the splitting of each block into two stations with the same capacity situations can occur where the origin block capacity is violated. That is, note in Figure 5 that each node displayed represents two overlapping nodes, l and r , which are not simultaneously constrained in the formulation. The reason is that capacity constraints concerning both stations are missing in model ACP. The interested reader may find the transformed instances of Harrod (2011) accessible for download via TTPLib (2008).

5.4 Effects Due to Translation of Data

Note that the transformation of data sets from one format to another may lead to different solution spaces. The translation from headway times to physical blocks, from ACP to RCHF, leads to changes in the upper bound of flow capacity of the routes. The original European network source data, the native format data for ACP, contains elaborate matrices of timed headway limits between all types of trains on each track segment. When converted to RCHF, these limits are considerably simplified and consolidated to the rule of a physical two block separation.

For example, track segment “71-1” in the Fulda-Hannover network is 44 km long, and has a matrix of 36 specified timed separation limits between pairs of train types. These timed separations range from three to thirty-two minutes. After conversion to RCHF, this track segment has thirteen defined blocks (dispatch track segments) of 3.38 km length each. In the same native data set, the running time on this track segment varies by train type from 22 to 33 minutes.

This leads to fundamental upper bound limits on the flow capacity of this track segment. For example, for two trains of type Inter-City, the timed separation is four minutes, and the running time is 22 minutes. Consequently, a maximum of five Inter-City trains that may occupy track segment 71-1 under ACP. However under RCHF and a two block separation between occupied blocks, the theoretical occupancy is four trains ($13/3$), regardless of train type or speed. Whether RCHF has more or less capacity than ACP is a function of the combination of train type paths requested and whether the two block separation rule is more or less favorable to that train type than the timed separation rule.

In practice, the physical block separation model correctly emulates the reality of train dispatching with current technology. An alternative dispatching scheme based on timed separation is in force on a very limited number of advanced technology railway networks. Timed separation is also a feature of a proposed new dispatching technology called “positive train control” (see Hansen (2001), Kraft and Smith (2001), Transportation Research Forum (2004), Emery (2008), or Wendler (2009)).

5.5 Results

The computational results for both models begin with the comparison of the results for the European data sets of § 5.1. A selection of sixteen sets of train path requests are evaluated under each formulation. These range in dimension from the smallest, at 20 train path requests, to the largest, at 1140 train path requests. Results for the first general comparison are presented in Table 5. Capacity limits are not enforced within stations (platform tracks) in these solutions. Specifically in formulation ACP the station set C is

empty, and in formulation RCHF the capacity of the station blocks are set at nine (effectively larger than any expected demand). This is a frequent assumption in the literature (Caprara et al. (2002) asserts this condition), reflecting the large station facilities typical in Europe.

The table contains for both models ACP and RCHF the model sizes (columns $|I|$ (= number of train requests), ROWS and COLUMNS), the optimal solution value denoted by w , and the computation time in seconds (columns *sec.*). Furthermore, the last column Δ shows the objective difference between the optimal solutions found for model ACP and model RCHF in %.

On average in this first data set, the hypergraph model returns solutions 3% lower than the ACP model, but this average difference is the result of large differences in a few select problem instances. Nearly half of the problem instances demonstrate no or insignificant difference (0.1%) in solution value. One of the questions to consider is whether the modeling of blocks with a single direction of travel is too restrictive and unrealistic for dispatching. In Table 6 the data set is revised so that the hypergraph model allows bi-directional dispatching on all routes. Under this relaxation, the hypergraph model returns solutions on average 5% greater than the ACP model, but again this is because of a few select large differences, with a large number of insignificant differences.

The results for the North American data sets are presented in Table 7. Figure 6 presents graphical views of the ACP input data and the ACP solution of scenario 1399. These views were created with the 3D-visualization tool **TRAVIS**. On the left hand side the train request bounds on that line of five blocks (six nodes) can be seen. Note that crossing is allowed on even numbered track segments, which represents a block with capacity two. On the right hand side the solution schedule from ACP is shown with time increasing on the vertical axes.

5.6 Comparison of Formulation Capabilities

A few interesting characteristics of the performance of these two formulations should be noted. Generally, the RCHF formulation requires less computation time and computing resources (size of rows and columns array) than the ACP formulation. However, in very resource constrained problems such as the single track scenarios of Table 7, the RCHF formulation exhibits exponential increases in computation time, as displayed in Figure 7. Indeed, larger single track scenarios were not presented in this paper because they quickly exceed available memory and/or reasonable computing time under formulation RCHF.

The ACP formulation has difficulty enforcing safety rules at the microscopic level (at the level of individual signal protected blocks of track). Figure 8 displays the stringline diagram of train movements from the RCHF

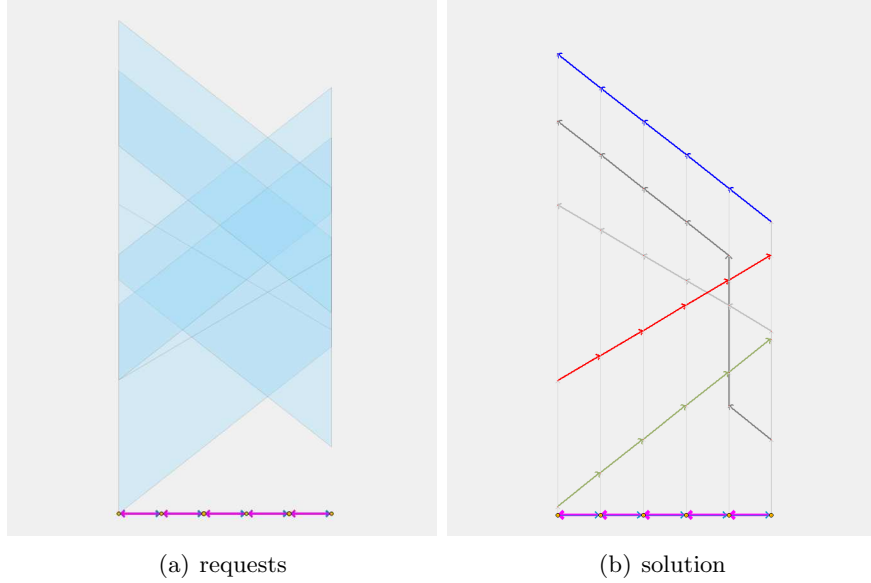


Figure 6: Screenshots for scenario 1399.

solution of scenario 1405. The diagram is drawn with proprietary software from the authors. Time is on the horizontal axis and the vertical axis is labeled with block identification numbers. Dark line train paths represent the faster priority traffic, and fine line paths represent the baseline traffic. Note how meets and passes between trains only occur on even numbered blocks, where sidings or double track are provided.

In spite of multiple attempts, variations, and adjustments, it was not possible to obtain the same precision of meet/pass enforcement under the ACP formulation. As described in § 5.3, the transformation of the microscopic network to a macroscopic network tends to overestimate the capacity of the blocks. Figure 9 displays the related solution stringline for the ACP formulation. Even numbered blocks are highlighted, and meets/passes that occur in odd numbered blocks (single track blocks) or that exceed block capacity are circled. These diagrams display a reordering of the trains asserted by ACP which does not violate the individual station node capacities. A full resolution of these solutions would require capacity pooling constraints that would modify the ACP model such that it might resemble the RCHF model.

The structure of RCHF is significantly different than ACP in the role that the side constraints play in the formulation. Side constraints are those constraints other than the pure network flow constraints. In formulation ACP these are referred to as coupling constraints, and in RCHF these are constraints (v). For the North American data set (Table 7), for example, 23% of the total rows in ACP are coupling constraints, and 55% of the

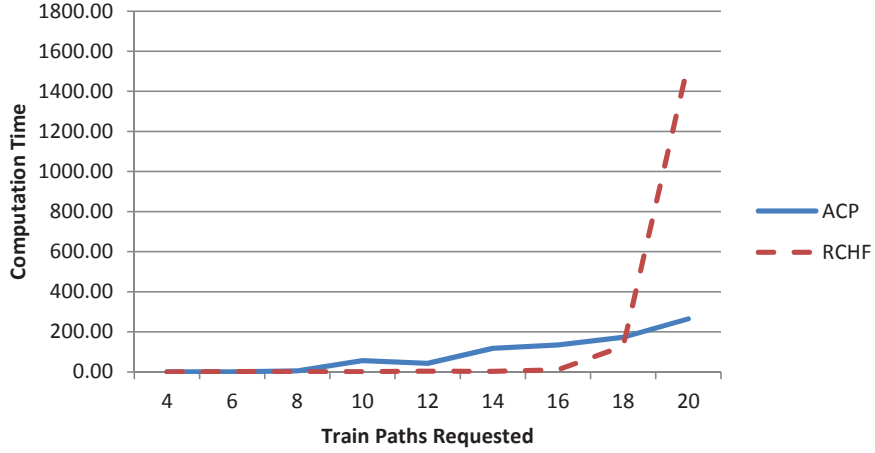


Figure 7: Comparison of Computation Times for Single Track Scenarios

rows in RCHF are side constraints (v). Surprisingly, these ratios are nearly constant for all nine problem instances, and thus detailed row counts are not presented.

5.7 Comment on Timed Separation Versus Physical Separation

The decision to model train separation by timed intervals or physical separation carries with it a fundamental set of operating rule enforcement protocols and an accompanying mathematical relationship amongst sets of trains. In §5.4 it was described how the choice of timed separation or physical separation may affect the theoretical flow capacity, and here is discussed how fundamental safety rules may be affected by this choice of rule enforcement. The value of explicitly modeling track infrastructure in a timetabling or dispatching formulation is best understood by examining typical train path interactions. The difficulty is that the true timed headway between two trains is frequently a function of the dispatch pattern of additional proximate train paths and the routing of the train path, both of which are frequently decision variables.

Figure 10 displays an example where three trains are traveling in the same direction along a signaled line with a typical red/yellow two block separation between any two trains. In the figure, two passenger trains, RE 103 and IC 208, are caught behind a slower freight train, GV 09. The timed headway between trains RE 103 and IC 208 would normally be a function of their track speed, but in this case their headway must be lengthened because train RE 103 is forced to travel at the slower speed of train GV 09. Thus, the time separation between any two trains is not only a function of

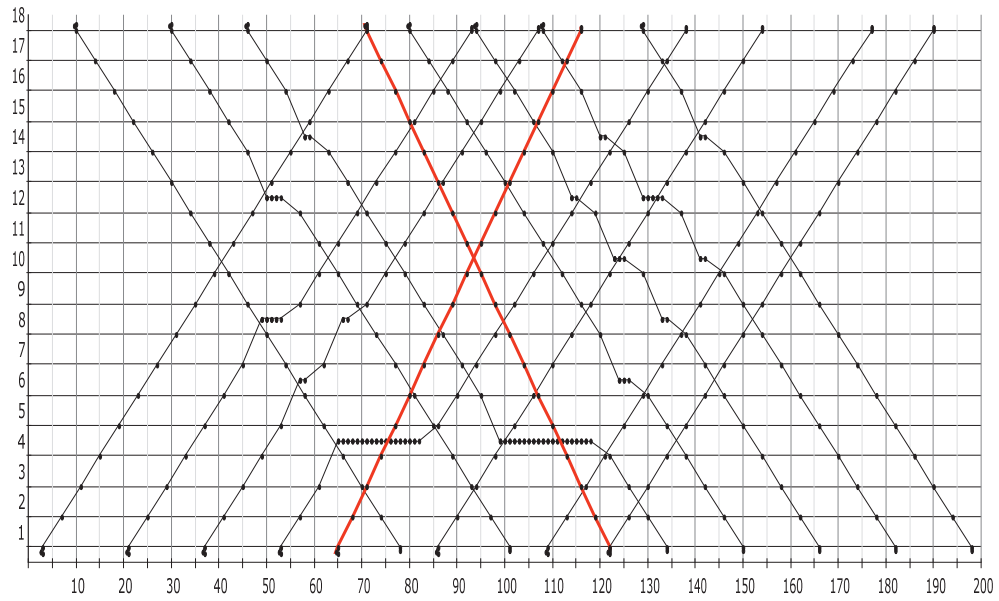


Figure 8: Stringline Diagram of Single Track Solution, Scenario 1405, from RCHF

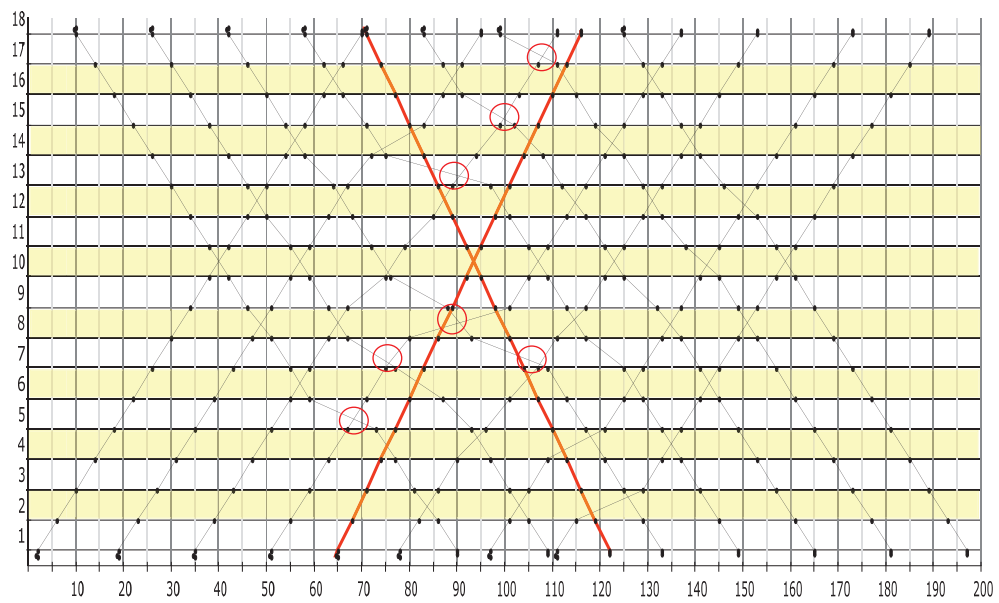


Figure 9: Stringline Diagram of Single Track Solution, Scenario 1405, from ACP

their joint performance, but the performance of any preceding train. This limiting function in fleets of trains is called “signal wake”, and is documented in Welch and Gussow (1986).

Signal wake is a significant concern in North American railways, but perhaps less of a concern in traditional European conditions. The reason for this is the structural design of many European networks where “stations” define all siding, signal, and meeting points, stations have relatively large (more than three) counts of tracks, and are located closely together. In such an environment, delaying a fleet of trains by a leading slower train is less likely, because there are more frequent opportunities to hold the delaying train in the clear. Many of the track segments in the native ACP data set (the European data set) contain long (more than 30 km) track segments without defined intermediate stations or sidings (although in reality satellite photos show that many stations do in fact exist). When converted to physical separation data sets for RCHF, these track segments would be subject to signal wake.

A second example is displayed in Figure 11, where train IC 208 is following GV 09. The question is, will IC 208 take the siding, or continue to follow GV 09? Or perhaps GV 09 should instead take the siding (the curved track), and allow IC 208 to overtake. That path choice is one or more decision variables, but the setting of those variables in turn affects the running time of these trains. Whichever train takes the siding must slow down to negotiate the points and curve at a slower speed than is necessary if the straight path through the switch is selected. For the timetabling model, this means that the headway time is not fixed, but is instead a function of the ultimate train path solution. In this paper, no attempt was made in the North American data set to model the diverging sidings as separate tracks with slower running times, so although this specific example is a general concern, it does not apply to these results.

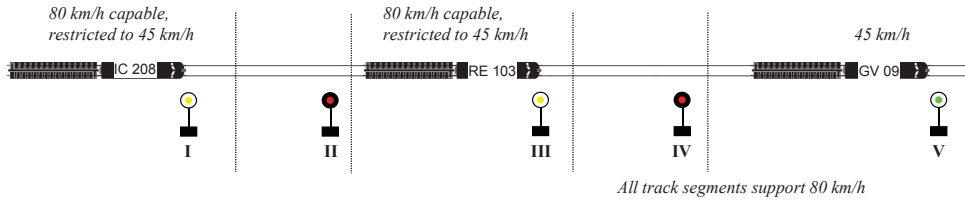


Figure 10: Train Headway Times Defined by Leading Train

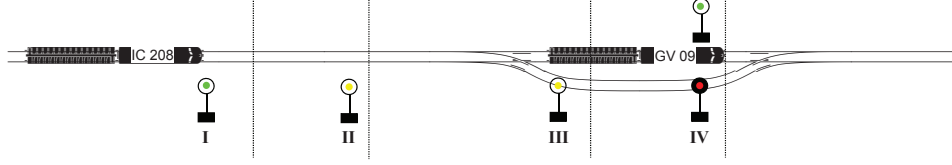


Figure 11: Train Headway Times Defined by Siding Selection

Table 4: Baseline Flow Train Parameters

Characteristic	Value
number of roundtrips	$(B - 1)/2$
time per block (speed)	4
optimal dispatch headway	16
minimum train separation in physical number of blocks	2
schedule slack time	25
utility value for scheduling a train (c_p^r)	\$1
stop cost per time unit (c_w^r)	\$0.01
penalty for network consumption beyond the planning horizon (c_{ed}^r)	\$0.0625

Table 5: Fulda-Hannover Solution Comparison, Uni-directional Dispatching in Hypergraph

SCENARIO	$ I $	ACP				RCHF				Δ
		ROWS	COLUMNS	w	sec.	ROWS	COLUMNS	w	sec.	
R07	184	45115	76534	1568737	153.14	72666	43069	1456059	20	-7.2
R09	114	2880	3759	798230	7.67	1869	1569	541347	2	-32.2
R11	98	23129	36658	777049	10.26	35504	21433	749448	11	-3.6
R13	28	704	765	284510	4.14	812	728	276103	1	-3.0
R14	33	4762	6568	299333	3.28	6462	4084	299111	2	-0.1
R15	31	8002	11922	220517	1.23	10865	6742	220225	4	-0.1
R16	30	16676	27300	195527	8.65	21402	12664	195324	8	-0.1
R25	117	2297	2968	840050	15.28	2532	2130	823799	3	-1.9
R26	118	15265	23677	917484	11.31	18358	10488	816047	7	-11.1
R27	118	25940	43903	887419	27.02	35147	19495	858140	11	-3.3
R28	102	61953	121007	915468	211.87	66822	36057	825526	18	-9.8
R29	20	295	323	175260	3.35	382	361	175179	0	0
R30	31	3150	4206	159589	2.62	4724	2777	159429	2	-0.1
R32	1140	82849	132400	919907	59.03	109739	63126	1113752	36	21.1
R47	25	5106	7250	208788	6.88	7543	4254	208614	4	-0.1
R48	41	23314	39927	312937	19.18	24793	13629	312667	9	-0.1

Table 6: Fulda-Hannover Solution Comparison, Bi-directional Dispatching with model RCHF

SCENARIO	$ I $	ACP				RCHF				Δ
		ROWS	COLUMNS	w	$sec.$	ROWS	COLUMNS	w	$sec.$	
R07	184	45115	76534	1568737	153.14	46108	43069	1629481	18	3.9
R09	114	2880	3759	798230	7.67	1727	1569	605919	2	-24.1
R11	98	23129	36658	777049	10.26	21211	21449	796695	10	2.5
R13	28	704	765	284510	4.14	622	587	253044	1	-11.1
R14	33	4762	6568	299333	3.28	3822	4084	299111	3	-0.1
R15	31	8002	11922	220517	1.23	6208	6758	220370	4	-0.1
R16	30	16676	27300	195527	8.65	12335	12682	195467	7	-
R25	117	2297	2968	840050	15.28	2266	2130	1010978	3	20.3
R26	118	15265	23677	917484	11.31	11054	10488	924276	7	0.7
R27	118	25940	43903	887419	27.02	20711	19495	947560	10	6.8
R28	102	61953	121007	915468	211.87	39812	36057	915264	17	-
R29	20	295	323	175260	3.35	382	361	175179	0	-
R30	31	3150	4206	159589	2.62	2672	2777	159443	2	-0.1
R32	1140	82849	132400	919907	59.03	82620	65870	1526555	32	65.9
R47	25	5106	7250	208788	6.88	4070	4254	208618	2	-0.1
R48	41	23314	39927	312937	19.18	13386	13629	312719	7	-0.1

Table 7: Single Track Solution Comparison

SCENARIO	$ I $	ACP				RCHF				Δ
		ROWS	COLUMNS	w	$sec.$	ROWS	COLUMNS	w	$sec.$	
1398	4	624	940	0.88	0.14	180	133	0.8750	0	-0.6
1399	6	2686	5272	2.21	0.50	792	591	2.6875	2	21.6
1400	8	6100	14186	3.70	5.51	1875	1357	4.8750	1	31.8
1401	10	10866	29490	5.82	56.26	3404	2431	6.0000	1	3.1
1402	12	16984	51938	7.40	42.85	5403	3813	8.0000	4	8.1
1403	14	24454	83845	9.10	118.25	7864	5503	10.0000	3	9.9
1404	16	33276	125442	10.92	134.67	10784	7501	11.9800	11	9.7
1405	18	43450	179575	12.45	172.97	14166	9807	12.8475	128	3.2
1406	20	54976	245934	14.16	264.59	18008	12421	13.8900	1541	-1.9

6 Conclusion

The solution of railway timetables is a crucial task at many management levels of railway systems. Timetable solutions must be determined repeatedly at strategic and tactical stages of network management. Most of these tasks continue to be performed manually, but advances in computer power and algorithms make automation of these tasks feasible. In addition, automated (mathematical programming) methods for timetable generation make possible wider exploratory analysis of network alternatives.

This paper directly compares two widely published mathematical structures for timetable optimization over a common set of sample problems. The first formulation, ACP, is characterized by enforcing train separation through a minimum time separation constraint at the entrance to a capacity limited track segment. The second formulation, RCHF, instead enforces train separation by enumeration of all individually controlled track segments and enumerating constraints that monitor the count of trains occupying each time expanded instance of a track segment. Sample problem sets represent both multiple track high density services in Europe and single track bidirectional operations in North America.

The contributions of this paper to the literature are that it establishes a relevant baseline comparison of these two formulations over practical, application relevant, sized problems, and that specific, tangible strengths and weaknesses may be identified for each formulation. The results demonstrate that both ACP and RCHF return comparable solutions in the aggregate. In nearly half the scenarios the solution train flow value is identical, and in total the average difference is less than five percent, but in some select instances the differences are severe. Where the solutions differ, there is no discernable pattern, and the differences are attributed to the unique specifications of the problem instances. In the majority of scenarios the RCHF formulation requires less computing resources, however, in some instances RCHF requires more resources than ACP and some classes of problem are not even tractable in RCHF. The ACP formulation experiences difficulty enforcing train separation on bidirectional single track lines. Neither formulation is absolutely superior over all scenarios and applications, but the results suggest that formulation RCHF should be attempted first if possible, that is if microscopic data is available and the infrastructure is limited such that operational feasibility must be established early in the timetable design process. Alternately, formulation ACP should be pursued if the RCHF formulation becomes intractable, or if only macroscopic network data is available.

This research was funded in part by a University of Dayton Research Council Seed Grant.

References

- Assad, A. A. (1980). Models for rail transportation. *Transportation Research Part A* 14, 205–220.
- Bixby, R. E. (2002, January). Solving real-world linear programs: A decade and more of progress. *Operations Research* 50(1), 3–15.
- Borndörfer, R., M. Grötschel, S. Lukac, K. Mitusch, T. Schlechte, S. Sören, and A. Tanner (2006). An auctioning approach to railway slot allocation. *Competition and Regulation in Network Industries* 1(2), 163–196.
- Borndörfer, R. and T. Schlechte (2007). Models for railway track allocation. In C. Liebchen, R. K. Ahuja, and J. A. Mesa (Eds.), *ATMOS 2007 - 7th Workshop on Algorithmic Approaches for Transportation Modeling, Optimization, and Systems*, Volume 07001 of *Dagstuhl Seminar Proceedings*. Internationales Begegnungs- und Forschungszentrum für Informatik (IBFI), Schloss Dagstuhl, Germany.
- Brännlund, U., P. Lindberg, A. Nou, and J.-E. Nilsson (1998). Railway timetabling using langangian relaxation. *Transportation Science* 32(4), 358–369.
- Bureau of Transportation Statistics (2009). National transportation statistics. Technical report, U.S. Department of Transportation, Washington D.C.
- Bussieck, M., T. Winter, and U. Zimmerman (1997). Discrete optimization in public rail transport. *Mathematical Programming* 79, 415–444.
- Cacchiani, V., A. Caprara, and P. Toth (2010a). Non-cyclic train timetabling and comparability graphs. *Operations Research Letters* 38(3), 179–184.
- Cacchiani, V., A. Caprara, and P. Toth (2010b). Scheduling extra freight trains on railway networks. *Transportation Research Part B: Methodological* 44(2), 215 – 231.
- Cacchiani, V. and P. Toth (2012). Nominal and robust train timetabling problems. *European Journal of Operational Research* 219(3), 727–737.
- Caprara, A., M. Fischetti, and P. Toth (2002). Modeling and solving the train timetabling problem. *Operations Research* 50(5), 851–861.
- Caprara, A., M. Monaci, P. Toth, and P. L. Guida (2006). A lagrangian heuristic algorithm for a real-world train timetabling problem. *Discrete Appl. Math.* 154(5), 738–753.
- CER (2008, February). Rail freight quality progress report 2007/2008. Technical report, Community of European Railway and Infrastructure Companies, Bruxelles. Accessed on January 21, 2010 at http://www.cer.be/force-download.php?file=/media/Publications_2/cer_freight_quality_report_2008.pdf.
- Cordeau, J.-F., P. Toth, and D. Vigo (1998). A survey of optimization models for train routing and scheduling. *Transportation Science* 32(4), 380–404.
- Emery, D. (2008). Enhanced ETCS level 3 train control system. In A. Tomii, J. Allan, E. Arias, C. Brebbia, C. Goodman, A. Rumsey, and G. Sciutto (Eds.), *Computers in Railways XI*. WIT Press.
- Erol, B., M. Klemenz, T. Schlechte, S. Schultz, and A. Tanner (2008). TTPlib 2008 - A library for train timetabling problems. In A. Tomii, J. Allan, E. Arias, C. Brebbia, C. Goodman, A. Rumsey, and G. Sciutto (Eds.), *Computers in Railways XI*. WIT Press.

- Flier, H., T. Graffagnino, and M. Nunkesser (2009). Scheduling additional trains on dense corridors. In J. Vahrenhold (Ed.), *Experimental Algorithms*, Volume 5526 of *Lecture Notes in Computer Science*, pp. 149–160. Springer Berlin Heidelberg.
- Gallo, G., G. Longo, S. Nguyen, and S. Pallottino (1993). Directed hypergraphs and applications. *Discrete Applied Mathematics* 42, 177–201.
- Hansen, P. (2001). Positive train control. *Trains* 61(1), 68–77.
- Harrod, S. (2011). Modeling network transition constraints with hypergraphs. *Transportation Science* 45(1), 81–97.
- Harrod, S. (2012, July). A tutorial on fundamental model structures for railway timetable optimization. *Surveys in Operations Research and Management Science* 17(2), 85–96.
- Holowaty, M. C., R. U. Cogswell, and N. E. Moyer (2004, January). Technical monograph: Transportation planning for the richmond - charlotte railroad corridor volume ii, fra/rdv-04/02. Technical report, Federal Railroad Administration, Department of Transportation, Washington DC.
- International Union of Railways (2008). International railway statistics, full year 2008. Technical report, International Union of Railways (UIC), Paris. Retrieved on February 4, 2010 at url: <http://www.uic.org/training/spip.php?article1348>.
- Kraft, E. R. (1988). Analytical models for rail line capacity analysis. *Journal of the Transportation Research Forum* 29(1), 153–162.
- Kraft, E. R. and M. E. Smith (2001). Lipstick on a pig? *Railway Age* 202(9), 12–.
- Lusby, R., J. Larsen, M. Ehrgott, and D. Ryan (2011, October). Railway track allocation: models and methods. *OR Spectrum* 33(4), 843–883.
- Mills, R., S. Perkins, and P. Pudney (1991). Dynamic rescheduling of long-haul trains for improved timekeeping and energy conservation. *Asia-Pacific Journal of Operational Research* 8, 146–165.
- Richards, C. W. (2006). Map of the month: Multiple-track main lines. *Trains* 66(1), 54–55.
- Schlechte, T. (2012). *Railway Track Allocation - Models and Algorithms*. Ph. D. thesis, Technische Universität Berlin.
- Törnquist, J. (2006). Computer-based decision support for railway traffic scheduling and dispatching: A review of models and algorithms. In *5th Workshop on Algorithmic Methods and Models for Optimization of Railways*, Dagstuhl, Germany. Internationales Begegnungs- und Forschungszentrum für Informatik (IBFI), Schloss Dagstuhl, Germany. <http://drops.dagstuhl.de/opus/volltexte/2006/659>.
- Transportation Research Forum (2004, March). *NAJPTC Program, Panel 1A: Positive Train Control*. Transportation Research Forum. 45th Annual Forum, Evanston, IL.
- TTPLib (2008). TTPLib-Homepage. <http://ttplib.zib.de>.
- Vassallo, J. M. and M. Fagan (2007). Nature or nurture: Why do railroads carry greater freight share in the united states than in europe? *Transportation* 34, 177–193.

- Welch, N. and J. Gussow (1986). Expansion of Canadian National Railway's line capacity. *Interfaces* 16(1), 51–64.
- Wendler, E. (2009). Influence of ETCS on the capacity of lines. In *Compendium on ERTMS: European Rail Traffic Management System*,. Eurailpress.

A Detailed Formulation of RCHF

An example of a hypergraph formulation is presented below. It differs from Harrod (2011) in that the objective and coefficients have been revised to correspond with those of the Borndörfer et al. formulation.

The variables, parameters, and sets of the formulation are provided in Table 8 and Table 9.

(Hypergraph)

$$\begin{aligned}
\max \quad & \sum_{\substack{r \in R \\ (p_o^r, j, u, v) \in \Psi^r}} [c_p^r - c_{eo}^r \max(0, p_{po}^r - u) - c_{lo}^r \max(0, u - p_{po}^r)] x_{p_o^r, j, u, v}^r \\
& - \sum_{\substack{r \in R \\ (p_d^r, \mathbf{e}^r, u, v) \in \Psi^r}} [c_{ed}^r \max(0, p_{pd}^r - v) + c_{ld}^r \max(0, v - p_{pd}^r)] x_{p_d^r, \mathbf{e}^r, u, v}^r \\
& - \sum_{\substack{r \in R, s \in \Phi^r \\ (i, s, u, v) \in \Psi^r | i \neq s}} [c_{eo}^s \max(0, s_{po} - v) + c_{lo}^s \max(0, v - s_{po})] x_{i, s, u, v}^r \\
& - \sum_{\substack{r \in R, s \in \Phi^r \\ (s, j, u, v) \in \Psi^r | j \neq s}} [c_{ed}^s \max(0, s_{pd} - v) + c_{ld}^s \max(0, v - s_{pd})] x_{s, j, u, v}^r \\
& + \sum_{\substack{r \in R \\ (i, j, u, v) \in \Psi^r | i=j}} c_w^r x_{i, j, u, v}^r \quad (1a)
\end{aligned}$$

Linear Network Constraints

$$\sum_{(p_o^r, j, u, v) \in \Psi^r} x_{p_o^r, j, u, v}^r \leq 1 \quad \forall r \in R \quad (1b)$$

$$\sum_{(a, i, u, t) \in \Psi^r} x_{a, i, u, t}^r = \sum_{(i, j, t, v) \in \Psi^r} x_{i, j, t, v}^r \quad \forall r \in R, \{i \in B | i \neq p_o^r\}, t \in T \quad (1c)$$

$$\sum_{(p_d^r, \mathbf{e}^r, u, v) \in \Psi^r} x_{p_d^r, \mathbf{e}^r, u, v}^r \leq 1 \quad \forall r \in R \quad (1d)$$

$$x \in \{0, 1\} \quad (1e)$$

Table 8: Variables, Parameters, and Nodes of Problem (P)

Component	Type	Description
$x_{i,j,u,v}^r$	binary variable	Occupancy arc representing the possession of node i at time u and the exit into node j at time v of train r
c_p^r	parameter	Fixed value (or “profit”) of train r completing its journey
c_w^r	parameter	Cost per unit time of enroute waiting (delay, other than requested stops) of train r
p_o^r	parameter	Origin of train r
p_{eo}^r	parameter	Earliest allowed origination time of train r
p_{po}^r	parameter	Preferred origination time of train r
p_{lo}^r	parameter	Latest allowed origination time of train r
c_{eo}^r	parameter	Penalty per time unit of early deviation from preferred origination time of train r
c_{lo}^r	parameter	Penalty per time unit of late deviation from preferred origination time of train r
p_d^r	parameter	Destination of train r
p_{ed}^r	parameter	Earliest allowed terminating time of train r
p_{pd}^r	parameter	Preferred terminating time of train r
p_{ld}^r	parameter	Latest allowed terminating time of train r
c_{ed}^r	parameter	Penalty per time unit of early deviation from preferred terminating time of train r
c_{ld}^r	parameter	Penalty per time unit of late deviation from preferred terminating time of train r
s	parameter	Specified service stop of train r from set Φ^r
s_{eo}	parameter	Earliest allowed arrival time at stop s
s_{po}	parameter	Preferred arrival time at stop s
s_{lo}	parameter	Latest allowed arrival time at stop s
c_{eo}^s	parameter	Penalty per time unit of early deviation from preferred arrival time at stop s
c_{lo}^s	parameter	Penalty per time unit of late deviation from preferred arrival time at stop s
s_{ed}	parameter	Earliest allowed departure time at stop s
s_{pd}	parameter	Preferred departure time at stop s
s_{ld}	parameter	Latest allowed departure time at stop s
c_{ed}^s	parameter	Penalty per time unit of early deviation from preferred departure time at stop s
c_{ld}^s	parameter	Penalty per time unit of late deviation from preferred departure time at stop s
\mathbf{e}^r	artificial node	Artificial sink node designating train r is off the network
b_t^i	parameter	Multiplicity (capacity, count of trains) of block i at time t
v_t^i	parameter	Multiplicity (capacity, count of trains) of cell i at time t
ϵ	parameter	Dimension of leading transition window
δ	parameter	Dimension of lagging transition window

Table 9: Sets of Problem (P)

Set	Description
T	The discrete time horizon, ordered with starting value $t=1$
R	The set of all trains
B	The set of all track blocks, ordered by a common reference direction of travel, such as “north” or “south”
R^N	The set of trains, $R^N \subset R$, traveling in the direction defined by increasing track block index
R^S	The set of trains, $R^S \subset R$, traveling in the opposite direction of trains in set R^N , $R^N \cup R^S \equiv R$
Φ^r	The set of requested (required) timetable stops for train r
Ψ^r	The set of feasible path arcs (i, j, u, v) for train r supplied from pre-processing
A	the set feasible hyper arcs $(u, v, t, t + d)$ for train i (supplied from pre-processing)
Υ	The set of transition nodes between track segments

Side Constraints

$$\sum_{\substack{r \in R \\ (i,j,u,v) \in \Psi^r | u \leq t < v}} x_{i,j,u,v}^r \leq b_t^i \quad \forall i \in B, t \in T \quad (2a)$$

$$\sum_{\substack{r \in R^N \\ v \in \{t+1-\epsilon, \dots, t+1+\delta\} \\ (i,j,u,v) \in \Psi^r | j=a+1, j \neq i}} x_{i,j,u,v}^r + \sum_{\substack{r \in R^S \\ v \in \{t+1-\epsilon, \dots, t+1+\delta\} \\ (i,j,u,v) \in \Psi^r | j=a, j \neq i}} x_{i,j,u,v}^r \leq v_t^a \quad \forall (a, t) \in \Upsilon \quad (2b)$$

$$\sum_{\substack{r \in R^N | h^r \geq 1 \\ a \in \{i-h, \dots, i-1\} \\ (a,j,u,v) \in \Psi^r | u \leq t < v, a \neq j}} x_{a,j,u,v}^r + \sum_{\substack{r \in R^N \\ (i,j,u,v) \in \Psi^r | u \leq t < v}} x_{i,j,u,v}^r \leq b_t^i \quad \forall i \in B, t \in T \quad (2c)$$

$$\sum_{\substack{r \in R^S | h^r \geq 1 \\ a \in \{i+1, \dots, i+h\} \\ (a,j,u,v) \in \Psi^r | u \leq t < v, a \neq j}} x_{a,j,u,v}^r + \sum_{\substack{r \in R^S \\ (i,j,u,v) \in \Psi^r | u \leq t < v}} x_{i,j,u,v}^r \leq b_t^i \quad \forall i \in B, t \in T \quad (2d)$$

Objective and constraints (1) define a set of independent single commodity flows. Side constraints (2) bind these flows together and regulate them as a multicommodity flow. Constraint (1b) enforces a single departure for

each train, and because it is an inequality also offers the option of removing the train from the schedule if it is not productive, while constraint (1c) enforces conservation of mass at the nodes (enforcing a single train path) and sinking constraint (1d) ensures a single terminal arrival.

Side constraint (2a) enforces a common block occupancy limit, but instead of enumerating every track segment with a capacity of 1, parallel segments such as sidings or multiple through track are defined as a single block with capacity $b_t^i \geq 0$. Other variations of this constraint, such as enumerating every track segment, are possible. Constraint (2b) enforces the limits on transitions between track segments. The sensitivity of these limits are adjustable through the “window” parameters. Finally, constraints (2c) and (2d) limit the headways or follow on spacing of trains. All of these features are presented in detail in Harrod (2011).