CETRA 2014 28–30 April 2014, Split, Croatia
3rd International Conference on Road and Rail Infrastructure

A Metaheuristic Approach for Integrated Timetable Based Design of Railway Infrastructure

Igor Grujičić[1], Günther Raidl[1], Andreas Schöbel[2], Gerhard Besau[3]
[1] *Institute of Computer Graphics and Algorithms, Algorithm and Data Structures Group, Vienna Technical University, Favoritenstraße 9/1861, 1040 Vienna, Austria*
[2] *Institute of Transportation, Research Centre for Railway Engineering, Vienna University of Technology, Karlsplatz 13/230-2, 1040 Vienna, Austria*
[3] *ÖBB Infrastruktur AG, Praterstern 3, 1020 Vienna, Austria*

Abstract

The design of new railway infrastructure is a complex planning process in most countries today due to a multitude of requirements. From an operational point of view new infrastructure basically has to fulfill the needs defined by customers. To this end passenger traffic is often organized in an integrated timetable with well defined arrival and departure times at major hub stations. So far there is no automated tool available to help in determining a minimum cost infrastructure fulfilling all the requirements defined by a timetable and the operation of the railway system. Instead, this task is typically carried out manually, based on graphical design, human experience, and also intuition. In our work we model this planning task as a combinatorial network optimization problem, capturing the most essential aspects. We then present a constructive heuristic algorithm that makes use of a dynamic programming procedure for realizing individual commercial stops. Computational experiments on instances derived from real scenarios indicate that the suggested approach is promising and the analysis of obtained results gives useful hints for future work in this area.

Keywords: Railway Infrastructure Design, Integrated Timetables, Combinatorial Optimization, Dynamic Programming, Heuristics

1 Introduction

The design of new railway infrastructure is nowadays strongly guided by pre-specified integrated timetables that have been derived from expected traffic to be served [2]. Integrated timetables synchronize the traffic in major nodes (hubs, e.g., main railway stations in major cities) at regular time intervals, ensure connectivity between different lines with minimum waiting times and allow passengers to remember easily the regular departure and arrival times. In many

European countries integrated timetables have been successfully introduced in the last years and could prove their substantial advantages.

Implementing the concept of integrated timetables, however, imposes major challenges and constraints, see e.g. [1]. In fact, the almost simultaneous arrival of the most relevant trains at a station and the strongly regulated travel times between stations, which must be multiples of a basic cycle interval, frequently demand extensions of existing railway infrastructure.

So far there is no systematic, automated tool available to aid the design of minimum cost infrastructure that fulfills all the requirements defined by the timetable and the operation of the railway system. Instead, this task is typically carried out manually based on graphical design, human experience, and also intuition, see e.g. [3]. In this paper we present a concrete combinatorial approach for modeling the basic problem. It considers already existing railway infrastructure as well as various extension possibilities in a fine-grained way. We then suggest a constructive heuristic algorithm for approximately solving this problem, which makes use of a dynamic programming procedure for locally optimal realizing individual commercial stops.

The following section presents the formal optimization model, which is based on the model we already introduced in [4] but refined in several details. Our solution method is described in Section 3. Section 4 summarizes experimental results obtained on some benchmark instances that were derived from real scenarios in Austria and have been validated by simulation of railway operation, e.g. OpenTrack or RailSys. Finally, Section 5 concludes the paper and provides thoughts on future work.

2 Combinatorial optimization model

We define the *Integrated Timetable Based Design of Railway Infrastructure* (TTBDRI) as a combinatorial optimization problem, trying to consider the most relevant real-world aspects. We are given the following input data.

- An undirected graph $G=(V,E)$ represents the existing railway infrastructure plus all possible extensions on a detailed level. The node set V contains different types of nodes, first of all the following infrastructure nodes corresponding to real objects:
    - *track segment nodes* representing physical, simple track segments of a certain length, they always have at most degree two;
    - *signal position nodes* representing signaling stations; they again always have degree two;
    - *crossing nodes* representing crossings of two lines; their degree always is four;
    - *switch nodes* representing classical switches; they have degree three (or possibly higher if more complex switches are modeled by single nodes);

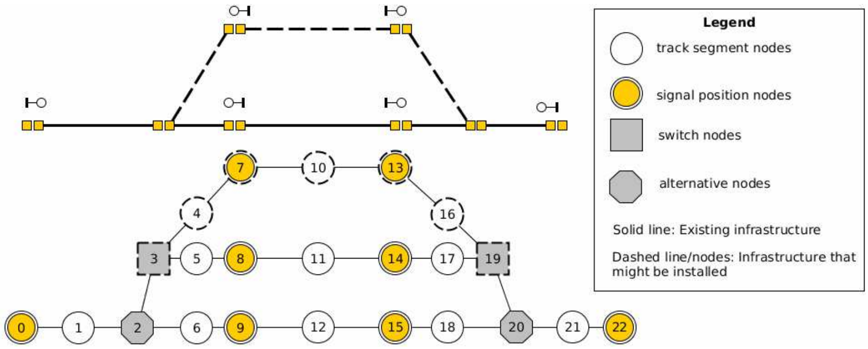

Figure 1. Example for a graph G modeling the existing infrastructure and possible extensions.

To model mutually exclusive alternatives for infrastructure extensions, we further use:

- o *alternative nodes*, which have degree $k+1$ for k mutually exclusive options.

Edges E represent the corresponding connections of the respective nodes. Multiple parallel tracks are always modeled by multiple paths. In order to avoid parallel edges and thus the need of a multigraph, it might occasionally be necessary to include:

- o *virtual nodes*; they always have degree two and might be considered as track segment nodes of length zero, i.e., they are just connecting two adjacent objects.

Figure 1 shows an example of infrastructure modeling.

- Let $R \subseteq V$ be the set of signal position nodes. Paths starting and ending at such nodes and otherwise containing only nodes from $V \setminus R$ are called (compound) routes ("Fahrstraßen"). Once a compound route is reserved for a train, no other train is allowed to enter any part of this route before the train has left and the route is released again.
- Let the subgraph $G^0 = (V^0, E^0)$, with $V^0 \subseteq V$ and $E^0 \subset E$, correspond to the already existing infrastructure and the graph $G' = (V', E')$ with $V' = V \setminus V^0$, $E' = E \setminus E^0$ represent the additionally possible infrastructure by which the existing infrastructure may be extended. Alternative nodes are considered to be part of V^0 iff one of the modeled options corresponds to an existing infrastructure, virtual nodes are part of V^0 if both adjacent nodes are also in V^0. All nodes $v \in V$ have associated costs $c_v \geq 0$ and lengths $l_v \geq 0$ with $c_v = 0$ for alternative nodes, virtual nodes, and all nodes in $v \in V^0$, $l_v = 0$ for signal position nodes, alternative nodes and virtual nodes.

- Set S represents the major railway stations considered in the integrated timetable. Each railway station $s \in S$ has associated a set of simple track segment nodes $V(s) \subset V$ corresponding to the tracks at platforms for boarding/disembarking trains in station s.
- Let $G^{\mathrm{D}} = (V, A)$ be the directed version of graph G, where we have for each edge $(u,v) \in E$ two corresponding oppositely directed arcs $(u,v),(v,u) \in A$.
- An integrated timetable specifies a set of commercial stops $C = \{C_1, \ldots, C_{|C|}\}$ to be realized, where a commercial stop $C_\ell \in C$ is a tuple $(s_\ell^{\mathrm{start}}, s_\ell^{\mathrm{end}}, T_\ell^{\mathrm{start}}, T_\ell^{\mathrm{end}}, G_\ell^{\mathrm{D}}, train_\ell, l_\ell)$ with $s_\ell^{\mathrm{start}}, s_\ell^{\mathrm{end}} \in S$ being start and destination stations and T_ℓ^{start} and T_ℓ^{end} the times when the train may leave station s_ℓ^{start} and has to arrive at station s_ℓ^{end} latest, respectively. The commercial stop has to be realized by a path in a given subgraph $G_\ell^{\mathrm{D}} = (V_\ell, A_\ell)$ with $V_\ell \subseteq V$ and $A_\ell \subseteq A$. It can safely be assumed that G_ℓ^{D} is acyclic. Finally, $train_\ell$ indicates the used train's ID. Typically, a train is used for a series of commercial stops. Let $l(train_\ell)$ refer to the train's length.
- Values $maxspeed_{\ell,v} \geq 0$ indicate the maximum allowed average speed by which the train realizing commercial stop $C_\ell \in C$ may go over node $v \in V_\ell$.

A solution consists of:

- a subgraph $G'' = (V'', E'')$ with $V'' \subset V$ and $E'' \subseteq E'$ indicating the infrastructure to be installed.
  Let $G^{\mathrm{e}} = (V^{\mathrm{e}}, E^{\mathrm{e}})$ represent the complete augmented infrastructure, i.e., $V^{\mathrm{e}} = V^0 \cup V''$ and $E^{\mathrm{e}} = E^0 \cup E''$.
- for each commercial stop $C_\ell \in C$ a directed path $P_\ell \subseteq A_\ell$ starting at a node from $V(s_\ell^{\mathrm{start}})$ and ending at a node from $V(s_\ell^{\mathrm{end}})$. Let $V(P_\ell) \subseteq V_\ell$ be the set of all nodes on this path. Considering the signal position nodes R as separators, P_ℓ can be partitioned into the ordered list of compound routes $L_\ell = (P_{\ell,1}, \ldots, P_{\ell,\lambda_\ell})$ with corresponding node sets $V(P_{\ell,1}), \ldots, V(P_{\ell,\lambda_\ell})$. The length of route $P_{\ell,i}, i = 1, \ldots, \lambda_\ell$, is $l(P_{\ell,i}) = \sum_{v \in V(P_{\ell,i})} l_v$.
- for each (infrastructure) node $v \in V(P_\ell), C_\ell \in C$, an (average) speed $speed_{\ell,v}$ that does not exceed the limit $maxspeed_{\ell,v}$. Consequently, the train takes time $T_{\ell,v} = l_v / speed_{\ell,v}$ for passing node v.
- for each route $P_{\ell,i}, i = 1, \ldots, \lambda_\ell$, $C_\ell \in C$, a reservation time slot $(T_{\ell,i}^{\mathrm{enter}}, T_{\ell,i}^{\mathrm{exit}})$ in which the train will safely be able to pass this route.

To be feasible, a solution must satisfy:

- For each commercial stop $C_\ell \in C : \forall (u,v) \in P_\ell \rightarrow u,v \in V^{\mathrm{e}} \wedge (u,v) \in E^{\mathrm{e}}$, i.e., the infrastructure used in the chosen paths must exist or be installed.
- All constraints for realizing possible extensions (e.g., mutual exclusivity of some alternatives) must be adhered.
- The time slots of consecutive routes of a commercial stop overlap exactly by the corresponding safety margins.
- For each commercial stop $C_\ell \in C$, the earliest start and latest arrival times T_ℓ^{start} and T_ℓ^{end} are adhered, respectively.
- At each time, each node $v \in V^{\mathrm{e}} \setminus R$ (i.e., except signal position nodes) may only be part of at most one reserved route.
- If the same train is used for two successive commercial stops, its arrival node at the station's track must be the same as the node where it leaves from later.

The objective is to find a feasible solution with minimum total costs $\sum_{v \in V^{\mathrm{e}}} c_v$.

The main difference between this formal model and the one presented in [4] is the introduction of different types of nodes in graph G . This allows for more flexibility and a more precise modeling, e.g., alternative nodes are used to distinguish between cases with or without building switches, depending whether the extension is applied (see Figure 1). In the previous model this was not possible.

3 Constructive heuristic solution approach

Our heuristic solution approach for TTBDRI consists of a construction framework in which an exact dynamic programming (DP) procedure is embedded for realizing the individual commercial stops. In the following subsection we present the DP, while Section 3.2 describes the construction framework.

3.1 Dynamic programming

The main idea for our DP is to use it for finding an optimal solution for just one given commercial stop $C_\ell \in C$. It will be applied iteratively until all commercial stops are realized and, possibly, a complete locally optimal solution is found. Therefore, from this point on, we will concentrate only on one given commercial stop $C_\ell \in C$ for which we want to find a cost-minimal realization.

To cover the aspect that a train may possibly start from and end at different platforms, we introduce artificial start and end nodes σ and τ , respectively, to the set V_ℓ of the commercial stop C_ℓ , i.e., $V_\ell' = V_\ell \cup \{\sigma, \tau\}$, and we set their costs and lengths to zero and maximum speed to one. Furthermore, we augment the arc set to $A_\ell' = A_\ell \cup \{(\sigma, s) \mid \forall s \in V(s_\ell^{\mathrm{start}})\} \cup \{(s, \tau) \mid \forall s \in V(s_\ell^{\mathrm{end}})\}$.

For every node $v \in V_\ell \setminus V_\ell^{\mathrm{R}}$ we define a set Y_v of time intervals in which it may be possible to reserve node v for the train to pass it. Every such time interval of Y_v has length at least equal to the minimum reservation time needed for node v. This minimum reservation time is the sum of the time needed for travelling through node v, the minimum time needed for travelling through any possible predecessor and the minimum time needed for travelling through any possible successor of node v.

For general principle of DP see e.g. [6]. Our DP stores labels $(c, T_{\mathrm{R}}^{\mathrm{start}}, T_{\mathrm{R}}^{\mathrm{end}}, t, \pi)$ for reached nodes, where

- π represents the preceding node;
- c represents the accumulated costs for the path from σ to v including c_v;
- $T_{\mathrm{R}}^{\mathrm{start}}$ represents the earliest time from which the reservation of the node v may start;
- $T_{\mathrm{R}}^{\mathrm{end}}$ represents the latest time until which the reservation of the node v may last and
- t represents the earliest arrival time at node v in time interval $[T_{\mathrm{R}}^{\mathrm{start}}, T_{\mathrm{R}}^{\mathrm{end}}]$.

The initial label for node σ is $(0, T_\ell^{\mathrm{start}}, T_\ell^{\mathrm{end}}, T_\ell^{\mathrm{start}}, \mathrm{null})$.

The extend function $(c, T_{\mathrm{R}}^{\mathrm{start}}, T_{\mathrm{R}}^{\mathrm{end}}, t, \pi) \rightarrow (\bar{c}, \bar{T}_{\mathrm{R}}^{\mathrm{start}}, \bar{T}_{\mathrm{R}}^{\mathrm{end}}, \bar{t}, \bar{\pi})$ for considering as next step to go from node u to node v is: $\bar{c} = c + c_v$, $\bar{\pi} = u$ and for the calculation of $\bar{t}$, $\bar{T}_{\mathrm{R}}^{\mathrm{start}}$ and $\bar{T}_{\mathrm{R}}^{\mathrm{end}}$ we need to distinguish the following cases:

- when $v \in V(s_\ell^{\mathrm{start}})$, then for every $[T^{\mathrm{low}}, T^{\mathrm{up}}] \in Y_v$ the extend function returns a label with $[\bar{T}_{\mathrm{R}}^{\mathrm{start}}, \bar{T}_{\mathrm{R}}^{\mathrm{end}}] = [T^{\mathrm{low}}, T^{\mathrm{up}}]$ and $\bar{t} = T^{\mathrm{low}}$;
- when $v \in V_\ell^{\mathrm{R}} \cup \{\tau\}$, then the extend function returns the label with $[\bar{T}_{\mathrm{R}}^{\mathrm{start}}, \bar{T}_{\mathrm{R}}^{\mathrm{end}}] = [T_{\mathrm{R}}^{\mathrm{start}}, T_{\mathrm{R}}^{\mathrm{end}}]$ and $\bar{t} = t + l_u / maxspeed_u$;
- when $v \in V_\ell \setminus (V_\ell^{\mathrm{R}} \cup V(s_\ell^{\mathrm{start}}))$ we distinguish the following:
    1. when $u \in V_\ell^{\mathrm{R}}$, then for every $[T^{\mathrm{low}}, T^{\mathrm{up}}] \in Y_v$ having a nonempty intersection with $[t, T_{\mathrm{R}}^{\mathrm{end}}]$ the extend function returns a label with $[\bar{T}_{\mathrm{R}}^{\mathrm{start}}, \bar{T}_{\mathrm{R}}^{\mathrm{end}}] = [T^{\mathrm{low}}, T^{\mathrm{up}}] \cap [t, T_{\mathrm{R}}^{\mathrm{end}}]$ and $\bar{t} = t$;
    2. when $u \in V_\ell \setminus V_\ell^{\mathrm{R}}$, then for every $[T^{\mathrm{low}}, T^{\mathrm{up}}] \in Y_v$ having a nonempty intersection with $[T_{\mathrm{R}}^{\mathrm{start}}, T_{\mathrm{R}}^{\mathrm{end}}]$ the extend function returns a label with $[\bar{T}_{\mathrm{R}}^{\mathrm{start}}, \bar{T}_{\mathrm{R}}^{\mathrm{end}}] = [T^{\mathrm{low}}, T^{\mathrm{up}}] \cap [T_{\mathrm{R}}^{\mathrm{start}}, T_{\mathrm{R}}^{\mathrm{end}}]$ and if $T_{\mathrm{R}}^{\mathrm{start}} \geq \bar{T}_{\mathrm{R}}^{\mathrm{start}}$, $\bar{t} = t + l_u / maxspeed_u$, else $\bar{t} = \bar{T}_{\mathrm{R}}^{\mathrm{start}} + \left[\left(t + l_u / maxspeed_u\right) - T_{\mathrm{R}}^{\mathrm{start}}\right]$.

An extension is feasible iff the following two conditions hold. (a) The actual arrival time at node v has to be feasible, i.e., $\bar{t} \in [\bar{T}_{\mathrm{R}}^{\mathrm{start}}, \bar{T}_{\mathrm{R}}^{\mathrm{end}}]$. (b) A time exists at which the train can pass from previous route to the current one. This is expressed as

$\overline{T}_{\mathrm{R}}^{\mathrm{start}} \in [T_{\mathrm{W}}^{\mathrm{start}}, T_{\mathrm{W}}^{\mathrm{end}}]$ where $[T_{\mathrm{W}}^{\mathrm{start}}, T_{\mathrm{W}}^{\mathrm{end}}] = [t, T_{\mathrm{R}}^{\mathrm{end}}]$ of the last signal position node on a path from σ to v.

Labels that are dominated by others can be removed. A label $l_1 = (c, T_{\mathrm{R}}^{\mathrm{start}}, T_{\mathrm{R}}^{\mathrm{end}}, t, \pi)$ dominates a label $l_2 = (\overline{c}, \overline{T}_{\mathrm{R}}^{\mathrm{start}}, \overline{T}_{\mathrm{R}}^{\mathrm{end}}, \bar{t}, \overline{\pi})$ iff the reservation time interval $[T_{\mathrm{R}}^{\mathrm{start}}, T_{\mathrm{R}}^{\mathrm{end}}]$ of label l_1 contains the reservation time interval $[\overline{T}_{\mathrm{R}}^{\mathrm{start}}, \overline{T}_{\mathrm{R}}^{\mathrm{end}}]$ of label l_2 and $c \le \overline{c}$ as well as $t \le \bar{t}$ with at least one of the latter two inequalities being strictly fulfilled.

Once when we have reached artificial end node τ actual solution is obtained by going backward until artificial start node σ is not reached. In every backward step we calculate the reservation time interval for visited node as well as appropriate speed used for travelling through it.

3.2 Construction heuristic

Our construction heuristic can be described by the following pseudo-code:

ConstructionHeuristic(S, C, i)

Given: partial solution S - a list of solutions for individual commercial stops; set C of not visited commercial stops; first not jet visited commercial stop i ;

Output: complete solution S if there is such, incomplete solution otherwise;

```
for all c ∈ C do
        if DP succeeded to find solution for the commercial stop c then
                S[i] = found solution;
                if i < the total number of given commercial stops then
                        ConstructionHeuristic( S, C \ {c}, i + 1 );
                else
                        return;   //complete solution obtained
                endif
        end if
end for
```

In the first call of above function we set S to be an empty set and C to be the whole set of the given commercial stops.

4 Experimental results

All experiments were carried out on an Intel Core i7-860 processor on 2.80GHz with 8GB of RAM. The algorithm has been implemented in C++.

Test instances model existing infrastructure between Feldkirch in Austria and Buchs in Switzerland with all intermediate stations in Austria, Liechtenstein and Switzerland. F_B_scenario1 represents an infrastructure with a possible flying crossing extension at Nendeln station and two trains of type RailJet. F_B_scenario_2 consider possible extensions at Schaanwald, Nendeln and Tisis and use four trains, two S-Bahns and two RailJets. RailJet trains have only two stops, the start and the end station. S-Bahn trains, however, stop at every intermediate station between their start and end stations with the minimum dwell time of 30 seconds. Thus, for every RailJet we have one commercial stop, while for every S-Bahn we have 8 commercial stops in this particular case.

Table 1. Summary of the experimental results on a set of real-world instances.

Instance	|V|	|E|	Number of		Objective value [in Millions of Euro]	Execution time [s]
			trains	commercial stops		
F_B_scenario1	171	176	2	2	24.890	0.079
F_B_scenario2	210	215	4	18	39.090	20.64

Table 2. Predefined arrival and departure times of used trains.

Train type	Direction Feldkirch - Buchs		Direction Buchs - Feldkirch	
	Departure t.	Arrival t.	Departure t.	Arrival t.
RailJet	51’	06’	54’	09’
S-Bahn	48’	12’	48’	12’

5 Conclusions and future work

In this article we have presented a formal combinatorial optimization model for the integrated timetable-based design of railway infrastructure. We have then suggested a first heuristic approach for approximately solving this problem, which consists of a constructive framework in which an exact dynamic programming procedure is embedded for realizing individual commercial stops. Obtained results appear reasonable and encouraging but also indicate the need of further algorithmic improvements to solve more complex scenarios more effectively.

In future work we aim at applying more sophisticated hybrid metaheuristics to obtain better solutions with prolonged computations (see e.g. [5]), but also exact techniques based on mathematical programming methods like column generation and Benders’ decomposition.

References

[1] Caimi, G., Laumans, M., Schüpbach, K., Wörner, S. & Fuchsberger, M.: The periodic service intention as a conceptual framework for generating timetables with partial periodicity. Transportation Planning and Technology, 34(4), pp. 323–339, 2011.

[2] Lichtenegger, M.: Der Integrierte Taktfahrplan: Abbildung und Konstruktion mit Hilfe der Graphentheorie, Minimierung der Realisierungskosten. Eisenbahntechnische Rundschau, 40, pp. 171–175, 1991.

[3] Schöbel, A. & Besau, G.: Timetable Based Design of Railway Infrastructure. In EURO-ZEL 20th International Symposium, June 2012.

[4] Schöbel, A., Raidl, G. R., Grujicic, I., Besau, G. & Schuster, G.: An optimization model for integrated timetable based design of railway infrastructure. In Proceedings of the 5th International Seminar on Railway Operations Modelling and Analysis – RailCopenhagen 2013, pp. 765-774, IAROR, 2013.

[5] G. R. Raidl, J. Puchinger, and C. Blum. Metaheuristic hybrids. In M. Gendreau and J. Y. Potvin, editors, Handbook of Metaheuristics, volume 146 of International Series in Operations Research & Management Science, pages 469–496. Springer, 2nd edition, 2010.

[6] B. Richard, Dynamic programming, Princeton University Press, Princeton, New Jersey, 1957.