Train Schedule Optimization in Public Rail Transport

T. Lindner* and U.T. Zimmermann

Department of Mathematical Optimization, Braunschweig University of Technology, Pockelsstraße 14, D-38106 Braunschweig, Germany

Abstract. Attractive train schedules are quite important for the success of public rail transport. We consider schedules which are repeated after some fixed time period. Periodicity is a well accepted, convenient attribute of all major railroad systems. Since trains share the public railroad network, we have to handle many conflicting demands. Increasing competition from transportation alternatives as well as the politically motivated process of deregulation forces railroad companies to reduce cost and resource utilization. In this situation, better decision tools are quite welcome. Here, we discuss models and corresponding methods for cost optimal train scheduling in real world railroad networks.

1 Introduction

The train schedule summarizes the arrival and departure times of trains at certain points of the railroad network. These data form the backbone of public rail transport planning. Different purposes require information on different levels of aggregation. In particular, data points may include stations (high degree of aggregation) or switches and important signal points (low degree of aggregation). For the German railroad network depending on the degree of aggregation, we have to handle 8000 to 27000 data points.

Former schedules for long distance trains were highly irregular. Since usually only one train per day was scheduled for a specific connection, a periodical schedule made no sense. However, in highly congested urban areas, periodical schedules were used almost from the very beginning, e.g. subway trains in London 1863, Paris 1900, Berlin 1902.

Customers prefer periodical schedules that are much easier to memorize. When introducing periodical or, as they were called, fixed interval schedules for long distance traffic in 1939, the Dutch railroad company marked a new epoch. Other European countries followed much later: Denmark in 1974, Switzerland in 1984, Belgium and Austria in 1991. In Germany, an hourly schedule for InterCity trains was started in 1979. Beginning in 1985/86, schedules for InterRegio trains were based on a two hour period. Finally, from 1992/93, regional trains were scheduled with some fixed time intervals. For a more detailed historical introduction, we refer to [11].

* Corresponding author. Email: T.Lindner@tu-bs.de

Traditionally, schedules are visualized by time space diagrams, cf. Fig. 1. For a particular route of the network, a time space diagram contains lines representing trains serving that route. Critical points or potential resource conflicts are quite obvious: the train's speed corresponds to the gradient of its line, a crossing of two lines shows that one train meets or passes another train.

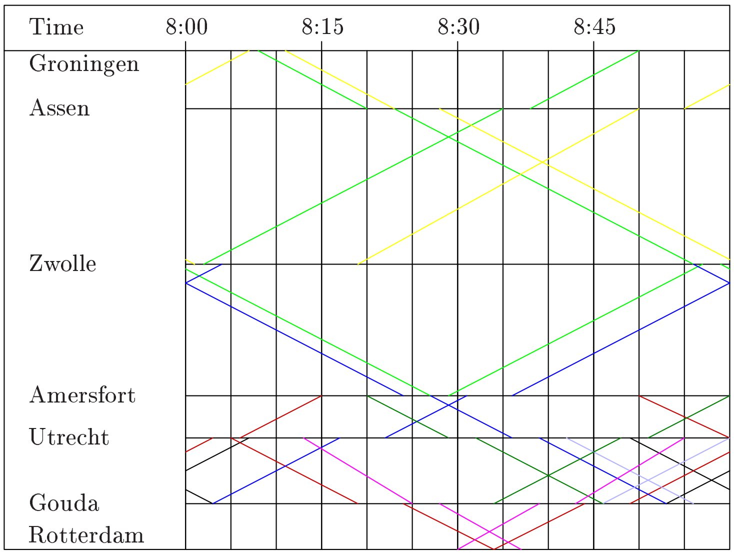

Fig. 1. Time space diagram for a schedule on the route Groningen-Rotterdam

Nowadays, software tools offer effective support for the construction of schedules. Information on track topology, engine and coach properties as well as available crews are stored in databases. Graphical user interfaces enable schedule planners to build and edit schedules interactively based on time space diagrams, e.g. Fig. 1. Conflicts are automatically indicated on the screen. However, the computer-automated generation and optimization of practicable schedules still remains quite time-consuming. In particular, currently implemented algorithms are simply too slow for networks of real world size.

Different departments of the railroad company formulate various goals. The sales and marketing departments give priority to travelers' requests, e.g. short travel time, direct connections and, if necessary, switching of trains on the same platform with short waiting time. The controlling, management and logistics departments pay much more attention to cost related aspects and ask for efficient management of rolling stock and personnel resources. Rolling stock is a scarce resource, rules implied by contracts and legal requirements set tight limitations for planning. High traffic load at critical points in the network and security requirements add a lot of operational constraints. Goals and requirements, cf. Fig. 2, are obviously in conflict. Moreover, external factors like political decisions influence the planning process.

Of course, planning of train schedules is only a small part of the general planning process of traffic systems (cf. [3]) which includes in particular line planning as well as management of rolling stock and crews. Interaction between all parts of the hierarchically decomposed planning process has to be handled. For example, an optimal train schedule is based on the line plan and will be used as input to the subsequent crew planning. Therefore, a proposed train schedule may need adaptation or redesign in order to meet certain requirements from other subproblems.

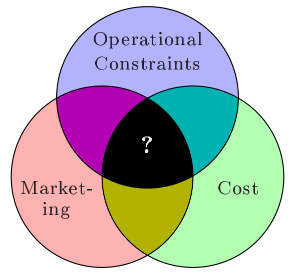

Fig. 2. Aspects of a schedule

2 Model

A railroad network is usually modeled by a graph $G = (V, E)$, where V denotes the set of nodes and E denotes the set of edges. Nodes represent stations or important network points like switches, and edges represent railroad tracks connecting these points. Train scheduling is based on a known line plan that defines the lines, i.e. the paths in the network that have to be served by trains with some fixed period T. This set of lines is denoted by $\mathcal{R}$. For a fixed line plan, we have to construct a train schedule which assigns feasible departure and arrival times to the nodes of all lines, cf. Fig. 3.

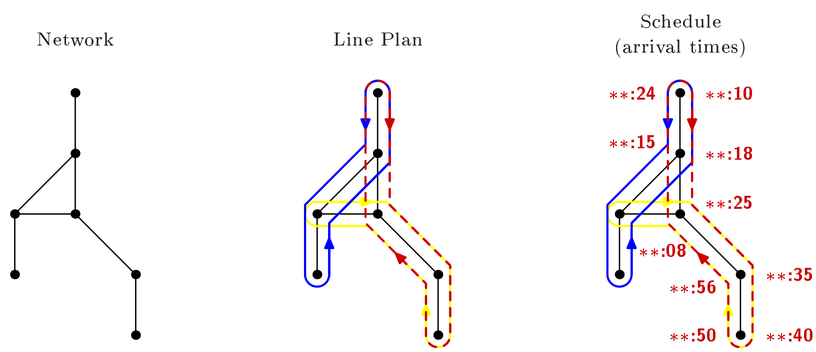

Fig. 3. Network, line plan and schedule for one line

Many requirements on train schedules can be modeled by so-called "periodical interval constraints". Here is a short example. At some station travelers want to change from line 1 to line 2. Therefore, in a feasible train schedule, the difference between the arrival of a train of line 1 and the departure of the corresponding train of line 2 has to stay within a certain interval. If the difference is too small, travelers may fail to reach the train of line 2. On the other hand, if the difference is too large, the waiting time at the station will be inconvenient.

Fig. 4 shows this situation for a period of $T = 60$ minutes. If 8–15 minutes are considered to be suitable, then 68–75 minutes are suitable as well, since arrivals and departures repeat every hour.

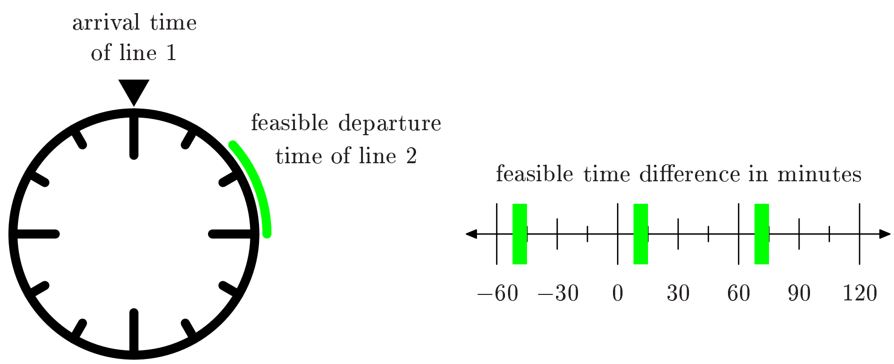

Fig. 4. Suitable differences between arrival of line 1 and departure of line 2

We model such a "periodical interval constraint" in the following way. If a_1^S is the arrival time of some train of line 1 at station S then we know that the arrival time of any train of line 1 at station S is $a_1^S + z \cdot 60$ for some $z \in \mathbb{Z}$. Similarly, the departure time d_2^S of some train of line 2 at station S fixes the departure times of all trains of line 2. Therefore, the time difference requirement for all trains can be modeled by

$$8 \leq d_2^S - a_1^S - z \cdot 60 \leq 15, \quad z \in \mathbb{Z}. \tag{1}$$

Many other requirements can be modeled in a similar way, e.g. headway times between trains on the same track and travel times of trains between stations. The problem of finding a feasible schedule, i.e. a schedule satisfying a set of periodical interval constraints is called PESP (Periodical Event Scheduling Problem) and was introduced in Serafini and Ukovich [14] in 1989. There are several different algorithmic approaches for solving PESPs, e.g. implicit enumeration methods in [13,14] and a cutting plane method in Odijk [12]. Voorhoeve's algorithm [15] is based on a constraint propagation.

For feasible train schedules, several objectives have been discussed in the literature. Minimizing total travel time is important in order to attract travelers. If train delays are essential, robustness of connections may well be as important. An obviously different point of view asks for the maximization of profits or the minimization of costs. Cordeau et al. [6] survey optimization models for train routing and scheduling.

Here, we propose a new model combining the PESP with the minimization of certain operational costs that were introduced by Claessens [4,5]. We suppose that the railroad company has the choice of assigning certain train types to the lines. A train type is characterized by speed, coach capacity, and bounds on the number of coaches, and cost. Here, investment cost per coach and engine as well as mileage dependent cost are included. Faster trains decrease the connection times along a line. If choosing some cheap train type for a line leads to an unavoidable conflict, a more expensive choice may be necessary to find feasible schedules. In our model, we minimize the cost of train type assignment for feasible schedules.

The operational costs of a schedule depend on the number of required engines and coaches as well as on the covered distance. We assume that a train cycles on one line only. The average speed of a train and the length of a line provide an estimate on the time required for one cycle around the line. If t denotes this cycle time, then $\lceil t/T \rceil$ trains are required to serve the line. The number of coaches of a train assigned to a line must be large enough to carry all the travelers on the corresponding paths.

In mathematical terms, we propose a mixed integer linear programming (MIP) model containing a large number of variables, i.e.

- $x_{r,\tau}$ train type $\tau \in \mathcal{T}_r$ is used for line $r \in \mathcal{R}$ (binary)
- $w_{r,\tau}$ number of coaches of type τ for trains of line $r \in \mathcal{R}$ (integer)
- $a^v_{r,\mu}$ arrival time of one train of line r, with direction μ, at station v
- $d^v_{r,\mu}$ departure time of one train of line r, with direction μ, at station v
- $\boldsymbol{z}$ vector of integers for the PESP constraints

where $\mathcal{T}_r$ denotes the set of train types that can be assigned to a line $r \in \mathcal{R}$. The direction μ can either be 0 or 1, corresponding to the two directions of each line.

The complete model, referred to as *minimum cost scheduling problem (MCSP)*, is shown in Fig. 5. We will explain its parts in some detail. The objective function sums all costs. Its left part contains the fix costs. The estimated cycle time of train type τ along line r is denoted by $\hat{t}_{r,\tau}$. C^{fix}_τ denotes the fix cost for one engine of type τ, C^{fixC}_τ the fix cost for one coach. The right part of the sum contains the mileage cost (per kilometer). Here, d_r denotes the length of line r.

With the first constraint class, we assure that the capacity of trains is sufficiently large to carry all travelers. The capacity of one train is the capacity of one coach, denoted by $\mathfrak{C}_\tau$, multiplied by the number of coaches $w_{r,\tau}$ of the train. Summation over all trains describes the available capacity.

minimize

$$\sum_{r \in \mathcal{R}} \sum_{\tau \in \mathcal{T}_r} \lceil \hat{t}_{r,\tau}/T \rceil \cdot (x_{r,\tau} \cdot C_\tau^{\text{fix}} + w_{r,\tau} \cdot C_\tau^{\text{fixC}}) + d_r \cdot (x_{r,\tau} \cdot C_\tau^{\text{km}} + w_{r,\tau} \cdot C_\tau^{\text{kmC}})$$

subject to

$$\begin{array}{lr}
\displaystyle\sum_{r \in \mathcal{R}, r \ni e} \sum_{\tau \in \mathcal{T}_r} \mathfrak{C}_\tau \cdot w_{r,\tau} \geq N_e & \text{for all } e \in E \\
\underline{W}_\tau \cdot x_{r,\tau} \quad \leq \quad w_{r,\tau} \leq \overline{W}_\tau \cdot x_{r,\tau} & \text{for all } r \in \mathcal{R} \text{ and } \tau \in \mathcal{T}_r \\
\displaystyle\sum_{\tau \in \mathcal{T}_r} x_{r,\tau} = 1 & \text{for all } r \in \mathcal{R}
\end{array}$$

and to PESP (many periodical interval constraints),

$$\begin{array}{lr}
\text{and to } x_{r,\tau} \in \{0,1\} & \text{for all } r \in \mathcal{R} \text{ and } \tau \in \mathcal{T}_r \\
w_{r,\tau} \in \mathbb{Z} & \text{for all } \tau \in \mathcal{R} \text{ and } \tau \in \mathcal{T}_r \\
a_{r,\mu}^v \in \mathbb{R} & \text{for all } r \in \mathcal{R},\, v \in r,\, \mu \\
d_{r,\mu}^v \in \mathbb{R} & \text{for all } r \in \mathcal{R},\, v \in r,\, \mu \\
\boldsymbol{z} \text{ vector of integer variables for PESP} &
\end{array}$$

Fig. 5. The minimum cost scheduling problem (MCSP)

With the second constraint class, we assure that the number of coaches for a train is feasible. The respective lower and upper bounds are denoted by $\underline{W}_\tau$ and by $\overline{W}_\tau$. In particular, if no train of type τ is assigned to the line r, i.e. $x_{r,\tau} = 0$, then no coach for this train can be chosen, i.e. $w_{r,\tau} = 0$.

With the third class of constraints, we assure that exactly one train type is chosen for each line. All three classes of constraints involve only train types and numbers of the coaches of these trains.

For fixed assignment of train types to lines, all other constraints are periodical interval constraints that we already mentioned before (cf. Fig. 4) describing a PESP. In particular, the arrival and departure times of all trains assigned to lines are subject to constraints similar to inequality (1). For each such constraint an integer variable models the periodicity of the schedule as in inequality (1). All these and all further integer variables in the remaining PESP constraints (e.g. waiting times at stations, headway etc.) are collected in a huge vector $\boldsymbol{z}$.

The MCSP contains subproblems that are very difficult to solve. The optimal choice of train types and the optimal choice of numbers of coaches are both NP-hard problems, and the generation of a feasible schedule (the PESP) is NP-complete (cf. [7]). On the other hand, the MIP model precisely states which train schedule we consider as feasible and which cost we want to minimize. Moreover, the MIP formulation enables easy addition of other requirements, e.g. available number of engines and coaches of a certain type, without destroying the structure of the model. The high flexibility of MIP models is very helpful when modeling and solving real world applications.

3 Solution Algorithm

The MCSP model is intended for strategic and tactical railroad planning, i.e. for long or medium term decisions rather than for day-by-day operations. In order to provide a helpful evaluation tool for different network or train type scenarios, only short computation times, say some minutes, are acceptable when solving the model for real world MCSP instances. Furthermore, the planner prefers to use standard computer equipment. For our test data like the InterCity or the InterRegio network of Germany and the Netherlands, the direct application of some commercial MIP solver resulted in several hours or even days of computation time on a 400 MHz Pentium II PC. For some instances, its main memory of 256 MB was not sufficient. Obviously, one has to find a more sophisticated approach for solving MCSP.

Our approach is based on decomposition. We consider the matrix of all objective function coefficients and all constraint coefficients of the MCSP in Fig. 6. Gray colored blocks will contain nonzero entries for the respective variables and the corresponding class of constraints.

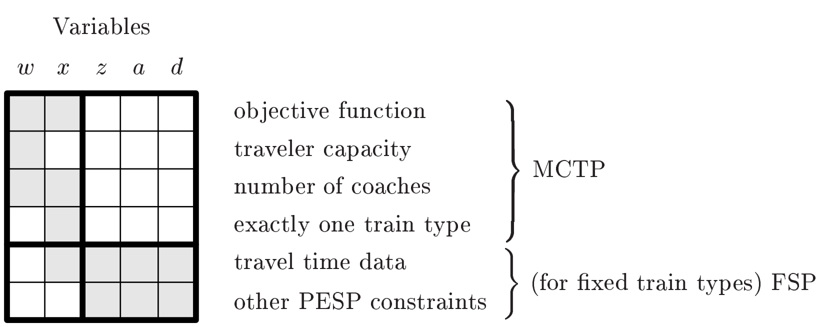

Fig. 6. Structure of objective function and constraints of the MCSP

With the exception of the x-variables in the travel time data constraints, the matrix can be divided into a two-block diagonal matrix. The first block represents the problem of minimizing costs subject to requirements on the trains that can be assigned to lines. This subproblem will be called *minimum cost train problem (MCTP)*. Its optimal solution assigns train types to the lines and defines the number of coaches in each assigned train or, shortly, it assigns trains to lines. At this point, it is not clear whether a feasible train schedule for these trains exists at all. For its existence, the number of coaches of the train plays no role. Therefore, the w-variables do not occur in the constraints describing feasible schedules. However, *for some fixed train type assignment*, these constraints form the second subproblem. This *feasible schedule problem (FSP)* is a PESP.

For solving MCSP, we propose a branch-and-bound method based on these subproblems. In particular, the MCTP is used as relaxation of the MCSP.

Branch-and-Bound Method

The method is shown in Fig. 7. Each branch-and-bound node represents an MCSP for which only a certain subset of train type assignments is allowed. In each node, we solve the corresponding MCTP relaxation. If the costs for the optimal MCTP train assignment are less than the costs of the currently best known MCSP train assignment, we try to find a feasible schedule of the MCTP train assignment. In other words, we solve the FSP for the fixed train type assignment defined by the optimal train assignment of the MCTP. If we find such a feasible train schedule, then we have a better MCSP train assignment. Otherwise, we branch on conflicting lines. For infeasible FSP, we determine a set of lines that cause infeasibility. For each of these lines, we generate a new MCSP for which the train type assigned to this line in the optimal solution of the previous MCTP is forbidden. For details, we refer to [8] and [10].

Solving MCTPs

Unfortunately, when directly applied to MCTP of real world size, commercial solvers may still need many hours of computation time. For its repeated use within a branch-and-bound scheme, we propose a reformulation with the following binary variables

$$w_{r,\tau,c} \quad \text{line } r \text{ uses train type } \tau \text{ with } c \text{ coaches}$$

Due to the definition of the binary variables, the constraints for the numbers of coaches are implicitly handled. Therefore the resulting binary MIP model, cf. Fig. 8, has less constraints. Though the number of variables is considerably increased in comparison with the previous model, we are able to solve this binary model much faster since the corresponding LP-relaxation is provably better.

Some further acceleration is possible using suitable classes of cutting planes. As a simple example for such a constraint, we consider a network edge with N_e travelers which is served by exactly two lines, say line 1 and line 2. Then, for each $d \in \{1, \dots, N_e\}$ either line 1 carries at least d travelers

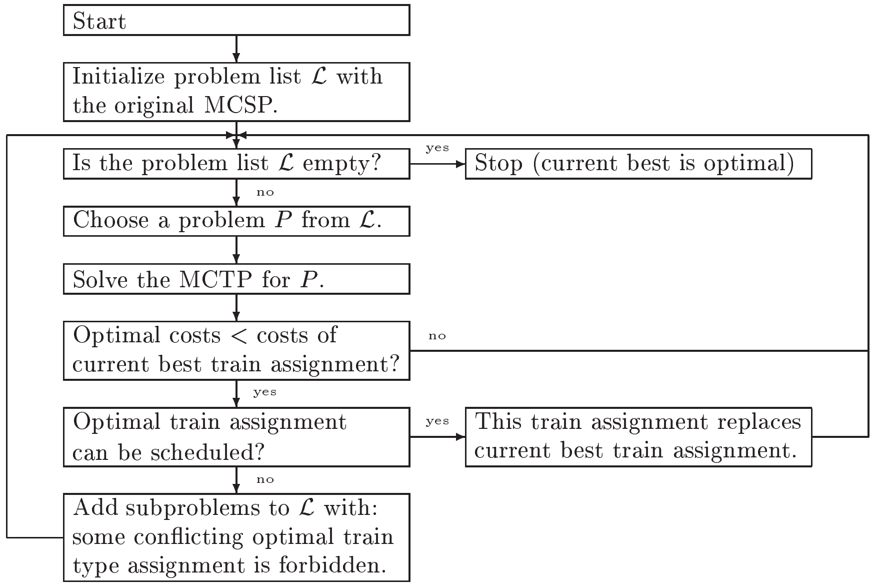

Fig. 7. Branch-and-bound method for MCSP solution

or line 2 carries at least $N_e - d$ travelers. Therefore, there has to be at least one feasible "train and coach" assignment with sufficient capacity to one of these lines. In terms of the "train and coach" assignment variables $w_{r,\tau,c}$, we may describe this implied requirement by the linear inequality

$$\sum_{\substack{(\tau,c) \text{ feasible for line 1} \\ c\cdot\mathfrak{C}_\tau \geq d}} w_{1,\tau,c} + \sum_{\substack{(\tau,c) \text{ feasible for line 2} \\ c\cdot\mathfrak{C}_\tau \geq N_e - d}} w_{2,\tau,c} \geq 1. \qquad (2)$$

Solving FSPs

Solving an FSP is equivalent to finding an integral solution to a set of linear inequalities. Again, the straightforward application of some commercial MIP solver leads to unacceptably high computation times. On the other hand, we already observed that FSPs are PESPs. There are several combinatorial algorithms for solving a PESP, e.g. the algorithm of Serafini and Ukovich. An enumeration scheme for all possible values for the z-variables forms the crucial part of this algorithm. The enumeration scheme proposed by Serafini and Ukovich is only suitable for small PESPs. We have developed a more effective scheme so that instances of the size of our FSPs can be solved in an acceptable amount of time. For details and more information on PESP algorithms, we refer to [9].

In view of the use of FSPs in the branch-and-bound method, it is important to analyze infeasible FSPs. Infeasibility leads to a branching step

$$\min \sum_{r\in\mathcal{R}} \sum_{\tau\in\mathcal{T}_r} \sum_{c=\underline{W}_\tau}^{\overline{W}_\tau} \Big(\lceil \hat{t}_{r,\tau}/T\rceil \cdot (C_\tau^{\mathrm{fix}} + c\cdot C_\tau^{\mathrm{fixC}}) + d_r \cdot (C_\tau^{\mathrm{km}} + c \cdot C_\tau^{\mathrm{kmC}})\Big)\, w_{r,\tau,c}$$

$$\sum_{r\in\mathcal{R}, r\ni e} \sum_{\tau\in\mathcal{T}_r} \sum_{c=\underline{W}_\tau}^{\overline{W}_\tau} \mathfrak{C}_\tau \cdot c \cdot w_{r,\tau,c} \geq N_e \text{ for each } e \in E$$

$$\sum_{\tau\in\mathcal{T}_r} \sum_{c=\underline{W}_\tau}^{\overline{W}_\tau} w_{r,\tau,c} = 1 \quad \text{for each } r \in \mathcal{R}$$

Fig. 8. Binary variable formulation of the MCTP

with the generation of new subproblems. In order to keep the size of the branch-and-bound tree as small as possible, it seems to be preferable to generate only few subproblems. In the algorithm of Serafini and Ukovich, one can sometimes easily identify a small set of train assignments to lines which imply infeasibility.

4 Results

Test data for our algorithms are real networks from the German railroad company *Deutsche Bahn (DB)* and from the railroad company of the Netherlands *Nederlandse Spoorwegen (NS)*. For the German railroad, we used network data, line plans, origin destination matrices and cost data for the InterCity (IC) and InterRegio (IR) networks. For the Dutch railroad, we obtained the respective data for the InterRegio (IR), InterCity (IC) and AggloRegio (AR) supply networks.

Some characteristics of these instances are displayed in Table 1. For all instances, 4 different train types have been considered.

	DB-IC	DB-IR	NS-IC	NS-IR	NS-AR
Number of Nodes	90	297	36	38	122
Number of Edges	107	384	48	40	134
Number of Lines	31	89	25	21	117
Average number of edges per line	7.5	5.9	5.0	5.8	4.2

Table 1. Some characteristics of the railroad networks

As an example, the InterCity network of the Netherlands is shown in Fig. 9. Cost optimal lines for this and other Dutch networks were generated in [3].

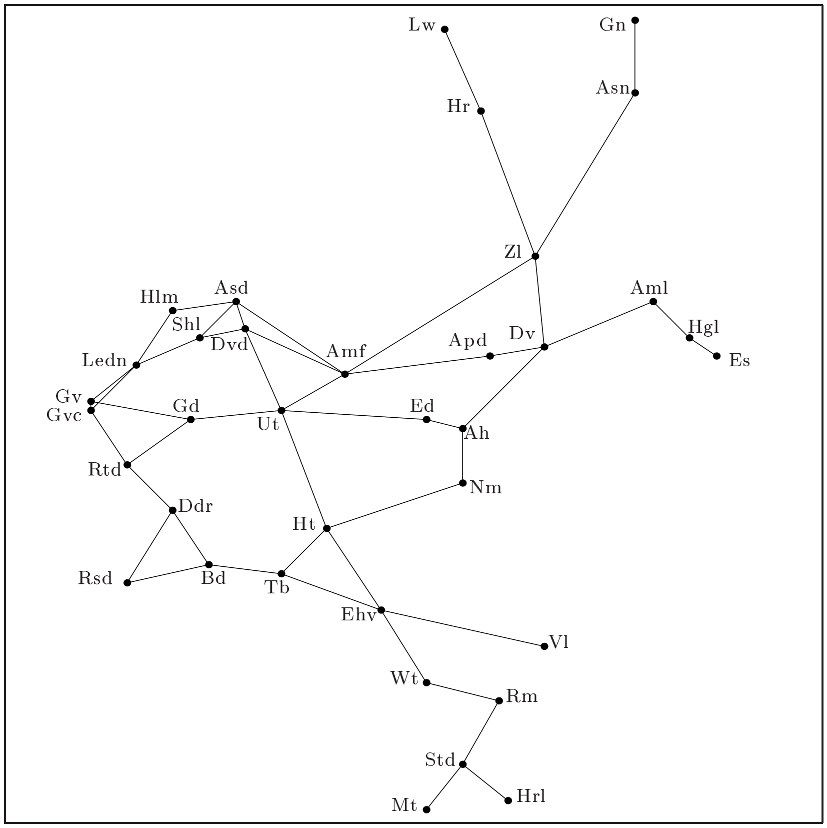

Fig. 9. InterCity network of the Netherlands

In order to guarantee suitable arrival and departure times for travelers changing trains, we studied the influence of respective periodical interval constraints at important stations, cf. Table 2.

With the exception of the Dutch AggloRegio network, which is a very large network, our branch-and-bound approach is very fast. Even when we enforce arrivals and departures suitable for travelers changing trains (by adding 40 periodical interval constraints), we generate nearly optimal train schedules within 5 min.

For larger networks, regional decomposition may be promising. At first, the network is partitioned into several regional networks which allow fast optimization. Then, the corresponding partial train schedules are combined to a train schedule of the complete network. The combination step will require some interactive corrections which may be supported by computerized tools.

Instance	DB-IC	DB-IR	NS-IC	NS-IR	NS-AR
Number of added constraints	0	0	0	0	0
Verified optimum found in	219 s	4 s	30 s	33 s	0:47 h
Optimality gap after 5 minutes	0%	0%	0%	0%	–*
Instance	DB-IC	DB-IR	NS-IC	NS-IR	NS-AR
Number of added constraints	40	40	40	40	40
Verified optimum found in	9:31 h	122 s	14:00 h	1:20 h	1:24 h
Optimality gap after 5 minutes	0.10%	0%	0.39%	0.27%	–*

*no solution in 5 minutes

Table 2. Results for test instances without and with added constraints

In manual schedule planning, regional decomposition is a well established approach.

5 Conclusions

With our method for minimizing costs of train schedules, traffic planners may interactively generate or evaluate different scenarios. We implemented a prototype that visualizes trains in the network, cf. Fig. 10. In particular, we display the current positions of all trains in the network. Trains are represented by squares numbered with the corresponding line. Different colors of squares indicate different train types.

As already mentioned, schedule planning is only one subproblem of traffic planning. Traditionally and with good reasons, the highly complex planning process is hierarchically decomposed, cf. Fig. 11. An survey on the application of mathematical programming methods for these subproblems is given in [1]. The solution of the line planning problem has already been the subject of a former BMBF-funded project (see [2]).

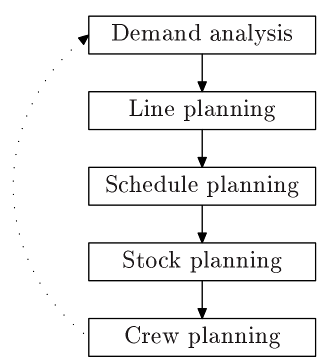

Fig. 11. Hierarchical traffic planning

Of course, it is well known that decisions on different levels are not at all independent. For example, any change in the transportation service, e.g. a new line plan or a new schedule, will influence the travelers' behavior. Travel demand data used in the development of a schedule will not match travel demand after the introduction of that schedule. In order to handle such difficulties, schedule planners try to simulate the travelers' behavior and to estimate the actual effect of a new schedule. They iteratively generate a schedule based on demand data, then estimate the new

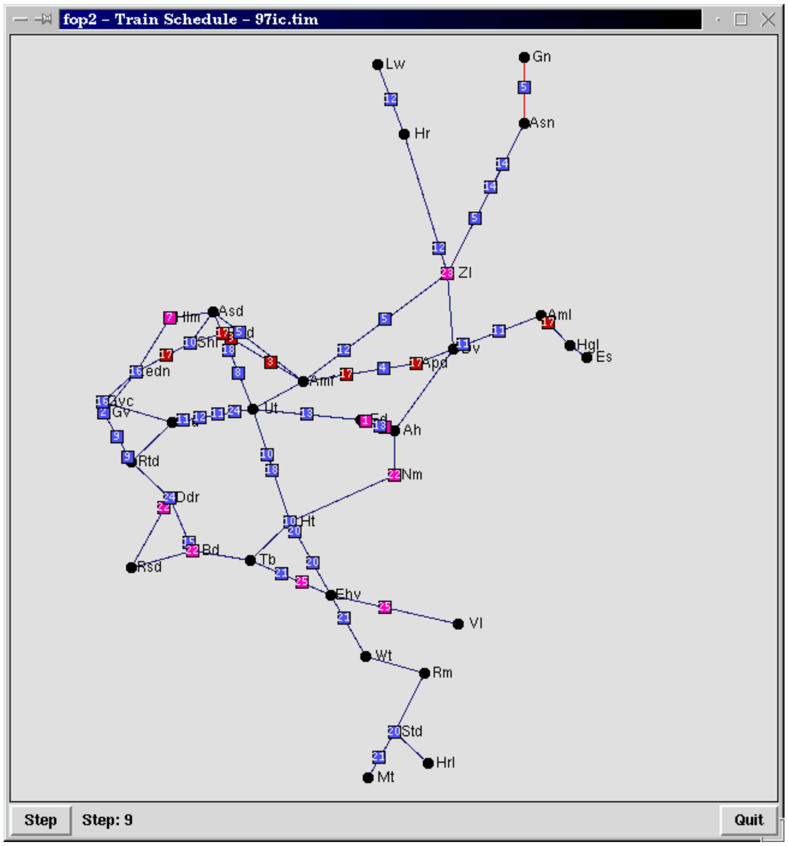

Fig. 10. Visualization of a schedule (screenshot)

demand resulting from the schedule etc. until changes in schedule and demand data become insignificant. Unfortunately, no mathematical argument is known that guarantees some kind of convergence for such an approach.

Acknowledgements We would like to express our thanks for encouragement, support and data made available by Adtranz Signal (Braunschweig), Dr. M. Krista (Lineas, Braunschweig) and Prof. Dr. L. Kroon (Railned, the Netherlands).

References

1. M.R. Bussieck, T. Winter and U.T. Zimmermann. Discrete optimization in public rail transport. *Math. Programming*, 79 (1–3) 415–444, 1997.

2. M.R. Bussieck, M. Krista, K.-D. Wiegand and U.T. Zimmermann. Linienoptimierung – Modellierung und praktischer Einsatz, in K.-H. Hoffmann, W. Jäger, T. Lohmann and H. Schnuck, editors. *Mathematik – Schlüsseltechnologie für die Zukunft*, Springer, 1997.
3. M.R. Bussieck. Optimal Lines in Public Rail Transport. PhD thesis, Technische Universität Braunschweig, 1998.
4. M.T. Claessens. De kost-lijnvoering. Master's thesis, University of Amsterdam, 1994.
5. M.T. Claessens, N.M. van Dijk and P.J. Zwaneveld. Cost optimal allocation of rail passenger lines. *European J. Oper. Res.*, 110 (3) 474–489, 1998.
6. J.-F. Cordeau, P. Toth and D. Vigo. A Survey of Optimization Models for Train Routing and Scheduling. *Transportation Science*, 32 (4), 380–404, 1998.
7. M.R. Garey and D.S. Johnson. *Computers and Intractability — A Guide to the Theory of NP-Completeness*. Freeman and Company, 1979.
8. T. Lindner. Train Schedule Optimization in Public Rail Transport. PhD thesis, Technische Universität Braunschweig, in preparation, 2000.
9. T. Lindner and K. Nachtigall. Periodical Event Scheduling. Technical Report, DLR, Braunschweig, in preparation, 2000.
10. T. Lindner and U.T. Zimmermann. Cost Optimal Train Scheduling, in preparation, 2000.
11. K. Nachtigall. Periodic Network Optimization and Fixed Interval Timetables. Habilitationsschrift. Universität Hildesheim, 1998.
12. M.A. Odijk. A constraint generation algorithm for the construction of periodic railway timetables. *Transportation Research B*, 30 (6), 1996.
13. A. Schrijver and A. Steenbeek. Dienstregelingontwikkeling voor Railned. Technical Report, Centrum voor Wiskunde en Informatica, 1994.
14. P. Serafini and W. Ukovich. A mathematical model for periodic scheduling problems. *SIAM J. Disc. Math.* 2 (4), 1989.
15. M. Voorhoeve. Rail Scheduling with Discrete Sets. Unpublished report, Eindhoven University of Technology, The Netherlands, 1993.