# Single Car Routing in Rail Freight Transport


Armin Fügenschuh[1], Henning Homfeld[2], and Hanno Schülldorf[2,3]

[1] Zuse Institute Berlin
Department Optimization
Takustraße 7, 14195 Berlin, Germany
fuegenschuh@zib.de

[2] Technische Universität Darmstadt
Schlossgartenstraße 7, 64289 Darmstadt, Germany
{homfeld,schuelldorf}@mathematik.tu-darmstadt.de

[3] Deutsche Bahn Mobility Logistics AG
Verkehrsmodelle und Simulation (GSU 1)
Idsteiner Straße 16a, 60326 Frankfurt am Main, Germany
hanno.schuelldorf@deutschebahn.com



**Abstract.** Cars in rail freight service follow prescribed routes from their origin via intermediate shunting yards to their destination. The goal in designing such routes is to reduce the number of trains and their travel distances. Various real-world hard constraints make the problem difficult to formulate and also to solve. We present integer programming formulations for this car routing problem arising at the largest European railway company and discuss their pros and cons.



**Keywords.** Single Car Rail Freight Transport, Mixed-Integer Linear Programming, Modelling, Nonlinear Constraints, Lifting.


## 1 Introduction

In 2006, around 500,000 Mil. ton kilometers freight was transported in Germany. The vast majority of freight is transported on the street (75%), a small share of 5% by ships, and the remaining 20% by train. From these 100,000 Mil. ton kilometers, inland and cross-border traffic have an equal share of 45% each, leaving 10% for transit traffic. Deutsche Bahn, the largest German railway company, offers mainly two products to commercial and industrial customers that want to transport goods via rail.

Typically large customers order whole trains of about 20 to 40 cars. In this case, Deutsche Bahn as the operator can pull such a train by a locomotive from origin to destination. Small customers on the other hand order only 1 to 5 cars. In such case it is too expensive to pull this group of cars by a single locomotive through the network. Instead the cars are only pulled to the next classification yard. There they are grouped with the cars from other customers, and then as new trains pulled to the next classification yards. There the trains are disassembled, and the cars are again re-grouped with others. This procedure is


Dagstuhl Seminar Proceedings 09261
Models and Algorithms for Optimization in Logistics
http://drops.dagstuhl.de/opus/volltexte/2009/2184

now repeated until each car has reached its final destination. The central question emerging here for our research is the following: What are the "best" paths for all cars that fulfill certain operative restrictions. Whatever "best" means in this context will be specified in the sequel.

This problem is challenging from a computational point of view due to the enormous complexity inherent in a large railway company. Deutsche Bahn operates a network of a total length of about 38,000 km. This network is traversed each day by approximately 5,000 freight trains with together 150,000 cars. The origins and destinations of the customers' orders are 2,200 terminal stations. Re-classification is carried out at intermediate classification yards of three different sizes: 200 are considered as small ("Satellitenbahnhöfe"), 30 as medium ("Knotenbahnhöfe"), and 11 are large ("Rangierbahnhöfe"). The differences among these will be explained in the sequel.

Classification yards schematically are made of three parts: entry tracks, sorting tracks, and exit tracks. When a train enters the yard, it is parked on an entry track. There the train is disassembled, and the individual cars are pushed over the hump, entering the sorting tracks behind. Each sorting track is assigned to a unique station (another yard or a terminal). Only by gravitational force the cars then roll into the right sorting tracks. As soon as enough cars are gathered on one sorting track, these cars are connected. This new train is then pulled into the exit group, where it waits, until it can leave the yard and re-enter the network. The operations within a classification yard by themselves give rise to interesting and difficult optimization problems. For our model, however, we can treat each yard as a black-box, and only have to take the respective capacity restrictions into account.

The problem of routing single cars through a railway network with intermediate classification operations emerged in the OR literature in the 1960s [1]. Computational studies mainly cover scenarios from US and Canadian railway freight companies [2,3,4,5,6,7,8,9,10,11], where different operational rules compared to Deutsche Bahn are applied. There the problem is known as the *blocking problem*, since it involves the decision which cars should stay together as a block (more than a car, less than a train) during their journey through the network. We will explain these differences in the next section.

## 2 Aspects of Single Car Scheduling at Deutsche Bahn

There are mainly two aspects that distinguish the single car scheduling at Deutsche Bahn from some of the approaches found in the literature. First, the main cost factor are trains and the sum of all train travel distances (also called train kilometers, for short), the second most important is the use of infrastructure, and only the least important are car kilometers. Hence cars can make detours, as long as they do not violate some maximum travel time constraints. That means that one is interested in those car paths that bundle as much orders as possible.

### 2.1 The Bundling Effect

Consider the graph in Figure 1. In this example a graph with 5 nodes $V := \{A, B, C, D, E\}$ is given. The respective costs for traversing an edge are $c_{AB} = 100, c_{AE} = 40, c_{BF} = 10, c_{CD} = 100, c_{CE} = 30, c_{DF} = 30, c_{EF} = 60$. There are two orders, one emerging in $A$ with destination $B$ (blue node), and the other has source $C$ and destination $D$ (red node).

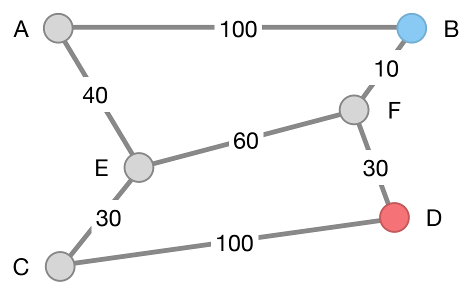


**Fig. 1.** An example for the bundling effect.

If the car travel costs would be the dominating factor then the optimal solution to this transportation problem is to assemble one train for each order. This gives car travel costs of $c_{AB} + c_{CD} = 200$ and similar costs for the trains.

On the other hand, if the train costs are dominating (as they do at our project partner), then the optimal solution is to transport the two orders to the intermediate yards $E$, assemble a train with these two orders that delivers them to $F$, where they are split up again and are taken in individual trains to their respective destination. For the car costs that means an increase from 200 to $c_{AE} + c_{CE} + 2 \cdot c_{EF} + c_{BF} + c_{DF} = 230$, but the train costs decrease from 200 to $c_{AE} + c_{CE} + c_{EF} + c_{BF} + c_{DF} = 170$.

### 2.2 Modes of Operating a Classification Yard

Generally speaking there are three main modes of operating a classification yard, called *individual car routing*, *blocking of orders*, and the rule of the *unique successor* (within Deutsche Bahn also called *Leitwege*). We will present all three in the sequel and discuss their differences. From an abstract point of view, we want to describe for each yard $s$ a function $f_s$ that yields the next station where the cars have to be sent to. The three operation modes differ in the input parameters that are necessary to determine this successor station.

Individual car routing gives the greatest flexibility of operation. It means that each car of the order receives its individual path through the network, that can differ from all other cars, even those cars from the same order. Speaking in terms of our function from above, it means that $f_s(i, j, l)$ depends on the origin $i$, the destination $j$, and car $l$ among all orders with cars from $i$ to $j$. From the organizational point of view that would mean not only to compute all these paths, but also attach the information to the cars so that the workers handling the cars can read it and put the car on the right sorting track. Although this system would offer great flexibility in practice, it is not implemented yet, and probably will not be implemented in the future, as long as humans are involved in the manual sorting process of the cars: The error sensitivity is simply to high so that it would most likely overcompensate the potential savings. However, if we leave the human factor out of our considerations we obtain car paths that give at least a theoretical lower bound on the actual costs.

Blocking cars that belong to the same order is similar to individual car routing, but now all cars from one order have to follow the same path. The cars can of course travel in different trains. At each intermediate station $s$ the next station $f_s(i,j)$ depends on the origin $i$ and the destination $j$ of the orders, but no longer on the number of the car $l$. For the manual sorting process that means that one source of potential errors is reduced and the entire freight system can be operated more stable.

The unique successor rule states that cars with the same destination share their remaining path to the destination from the station where they first meet. That is, at an intermediate station $s$ the next station $f_s(j)$ only depends on the destination $j$, and not even on the origin anymore. This yields a very simple rule and hence reduces the human error potential again by some order of magnitude. The price is a lower flexibility when designing car paths.

## 3 Model Formulation

We give a formulation of the problem as a linear mixed-integer programming problem.

Denote by $V$ the set of all stations (yards and terminals), and by $A \subset V \times V$ the set of precedence relations. The set of customer orders is denoted by $K$.

We introduce three families of integer variables. For the routes of the cars we introduce decision variables $x^k_{i,j} \in \mathbb{B}$ for all $(i,j) \in A$ and $k \in K$. If $x^k_{i,j} = 1$ then station $j$ is directly after station $i$ in the path for the cars belonging to order $k$. Another decision of the model concerns the number of trains from station $i$ to $j$, for which we introduce the integer variable $y_{i,j} \in \mathbb{N}$. Finally the model has to decide on the number of sorting tracks $n_{i,j} \in \mathbb{N}$ at station $i$ on which trains are assembled exclusively in direction of station $j$.

Similar to the variables there are also three main groups constraints: one for the orders, one for the trains, one for the yards.

The central constraints for the orders are the usual multi-commodity flow conservation constraints:

$$\forall i \in V, k \in K : \sum_{j:(i,j)\in A} x^k_{i,j} - \sum_{j:(j,i)\in A} x^k_{j,i} = \begin{cases} 1, \text{ if } i = o(k), \\ -1, \text{ if } i = d(k), \\ 0, \text{ otherwise}, \end{cases} \tag{1}$$

stating that each order $k$ has to start at its origin $o(k)$ and end at its destination $d(k)$, and is not lost in between.

There is an upper bound restriction $T_k$ on the time that the cars of order $k$ are allowed to travel in the network. Typically the freight company offers so-called A-C-relations (delivery within 48 hours) and A-B-relations (express delivery within 24 hours). We denote by $t_{i,j}$ the travel time from station $i$ to $j$ and by $u_i$ the time spend in station $i$. For the moment both are given input parameters, which is justified for $t_{i,j}$ but questionable for $u_i$. We will come back to this in Section 4.4. For now the time limit constraint is the following:

$$\forall k \in K : \sum_{(i,j)\in A} (u_i + t_{i,j}) \cdot x^k_{i,j} \leq T_k. \tag{2}$$

Each yard $i$ has a maximal capacity of cars $H_i$ that it can handle per time period. For large yards with a hump the limiting factor is the respective hump capacity. For small yards without a hump, the sorting is carried out by small locomotives, which is even more a limiting factor. If we denote by $v_k$ the number of cars belonging to order $k$ we can formulate this capacity restriction as the following constraint:

$$\forall i \in V : \sum_{k\in K, j:(i,j)\in A} v_k \cdot x^k_{i,j} \leq H_i. \tag{3}$$

The number of trains that can be assembled per time period on a single sorting track is a station dependent parameter $N_i$ for yard $i$. Typically this value is between 3 and 6. We thus have the constraint

$$\forall (i,j) \in A : n_{i,j} \leq N_i \cdot y_{i,j}. \tag{4}$$

This means, the number of tracks in station $i$ on which trains in direction of $j$ can be assembled equals the number of trains from $i$ to $j$ divided by the number trains per track and time period.

Each yard $i$ has a total number of sorting tracks $Y_i$. Hence the total number of assignments must not exceed this capacity restriction:

$$\forall i \in V : \sum_{j:(i,j)\in A} y_{i,j} \leq Y_i. \tag{5}$$

The trains have an upper limit on the total length and the weight of the cars. It is not allowed to assemble trains longer than 700m. This is because freight trains have to wait in the network at certain bypass tracks so that faster passenger trains can overtake them. Since these bypasses are 700m long (speaking of the German railway network, in Denmark, for example, this is 850m), no longer freight train is allowed to travel in the network. If $l_k$ denotes the length in meters of all cars belonging to order $k$ and $L_{i,j}$ the maximum length of trains between $i$ and $j$, then

$$\forall (i,j) \in A : \sum_{k\in K} l_k \cdot x^k_{i,j} \leq L_{i,j} \cdot n_{i,j}. \tag{6}$$

There is an upper bound weight restriction $W_{i,j}$ that reflects the strength of DB's locomotives and the track properties between $i$ and $j$. If $w_k$ denotes the weight in tons of all cars belonging to order $k$, then

$$\forall (i,j) \in A : \sum_{k \in K} w_k \cdot x^k_{i,j} \leq W_{i,j} \cdot n_{i,j}. \tag{7}$$

Finally we have the "Leitweg" unique successor constraint: The successive station only depends on the destination, not on the origin, and also not on the order. That means, if two different orders with the same destination meet at some station in the network, it is not allowed to send them further to different stations, which is modelled by the following inequalities:

$$\forall k,l \in K, l \neq k, d(k) = d(l), i \in V, (i,j_1),(i,j_2) \in A, j_1 \neq j_2 : \\ x^k_{i,j_1} + x^l_{i,j_2} \leq 1. \tag{8}$$

The overall objective is to find the most economical car paths. The most crucial cost component is the number of train kilometers, second is the amount of used infrastructure, which we assume is proportional to the number of sorting tracks. Of least importance is the number of car kilometers. Let $\delta_{i,j}$ denote the distance (in the network) between station $i$ and $j$, then we have the following objective function:

$$\alpha_1 \cdot \sum_{(i,j) \in A} \delta_{i,j} \cdot n_{i,j} + \alpha_2 \cdot \sum_{(i,j) \in A} y_{i,j} + \alpha_3 \cdot \sum_{k \in K, (i,j) \in A} \delta_{i,j} \cdot x^k_{i,j} \rightarrow \min, \tag{9}$$

where $\alpha_1, \alpha_2, \alpha_3$ are suitably scaled weight parameters.

## 4 Improving the Formulation

We describe two exact methods and one heuristic method to improve the model formulation from the last section. The goal is to achieve a formulation that a MILP solver is able to solve faster.

### 4.1 Lifting

It is possible to combine the unique successor constraints (8) with the flow conservation constraints (1). Consider two orders $k, l \in K, l \neq k$ with the same destination $d(k) = d(l)$ and a station $i \in V$. By (8) it is required that $k$ and $l$ must not use different arcs $(i,j_1),(i,j_2)$ with $j_1 \neq j_2$. Because of (1) it is required that exactly one outgoing arc for the cars of $k$ and $l$ is selected. Thus we can consider any partition of $V$ into disjoint sets $V_1, V_2 \subset V$ with $V_1, V_2 \neq \emptyset, V_1 \cap V_2 = \emptyset$, and $V_1 \cup V_2 = V$ (shortly $V_1 \dot{\cup} V_2 = V$), and lift further coefficients into (8) that were previously set to zero:

$$\forall k,l \in K, k \neq l, d(k) = d(l), i \in V, V_1 \dot{\cup} V_2 = V : \\ \sum_{j_1 : (i,j_1) \in A} x^k_{i,j_1} + \sum_{j_2 : (i,j_2) \in A} x^l_{i,j_2} \leq 1. \tag{10}$$

We remark that there are too many partitions to include all these inequalities right from the beginning. Instead we separate them on demand.

### 4.2 Tree Structures

Another way to formulate the unique successor constraints (8) or (10) is to consider all orders with the same destination simultaneously, and to use arborescences rooted at the common destination. This leads to a formulation with less constraints but more variables. We will later discuss whether the trade-off is positive.

Denote by $\mathcal{K} := \{d(k) : k \in K\}$ the set of all destinations. We introduce new binary variables $z^{\kappa}_{i,j} \in \mathbb{B}$ for all $(i,j) \in A$ and $\kappa \in \mathcal{K}$. If $z^{\kappa}_{i,j} = 1$ then all cars with destination $\kappa$ in station $i$ are sent to station $j$ next. The uniqueness of the successor station is guaranteed by the following constraints:

$$\forall i \in V, \kappa \in \mathcal{K} : \sum_{j:(i,j)\in A} z^{\kappa}_{i,j} \leq 1. \tag{11}$$

The cars are only allowed to use those arcs that were selected by the trees before:

$$\forall (i,j) \in A, k \in K : x^{k}_{i,j} \leq z^{d(k)}_{i,j}. \tag{12}$$

Equations (11) and (12) replace now equations (8) and (10).

### 4.3 Heuristic Cuts: Hierarchy Constraints

The previous two model improvements were exact in nature, that is, their inclusion does not chance the set of feasible solutions to the problem. We will now present a further idea that is of great help in finding solutions, but it will cut away parts of the feasible solutions. This is what we call *heuristic cuts* [12].

Valid inequalities are typically discovered by a polyhedral analysis of the convex hull of feasible solutions. Heuristic cuts, however, emerge from a different source. First, one solves a number of (small) test instances to proven optimality. Then, one has to analyse the structure of the solutions, in order to deduce certain structures or patterns within them. In doing so, new knowledge about optimal or near optimal solutions is gathered. If possible this new knowledge is formulated as (linear) inequalities and included in the model formulation, so that our anticipation of the solution is already in the model from the very beginning. Or, to put it in other words, the implicit description of an optimal solution is made more explicit.

In general, the cuts one discovers in that way are too strong from a polyhedral point of view. But, on the other hand, they can speed up the solution process drastically. If one was careful enough they only sacrify little optimality, and they might be able to streamline or polish the solution, so that minor peculiarities that cannot be explained, and might even be the result of weak input data, do not show up in the end.

After this brief introduction to heuristic cuts we will discuss what that means to our car routing model. From the solution of small instances we observed that cars are somehow attracted by large yards. A typical paths through the network starts at a terminal station, seeks it way to a satellite yards, and then to the medium and large yards where the cars travel the longest distances through the network. When they come closer to their destination they leave this hierarchy level. Then they are transported to a satellite and finally to the terminal. The vast majority of all paths has such a hump structure: the cars first climb up in the hierarchy from small to large yards, then they stay a while on that level, and finally they climb down from large to small again. Only very few paths had intermediate jumps, from high to low and back to high again.

Making use of this observation is further justified by an analysis of historical data. About 98% of all paths had a hump structure.

In order to separate the hump-structured paths from the zig-zag paths, we introduce the following heuristic cuts:

$$\forall j \in V, l \in K : \sum_{i:(i,j)\in A, h_i \leq h_j} x^l_{i,j} + \sum_{k:(j,k)\in A, h_j \geq h_k} x^l_{j,k} \leq 1. \tag{13}$$

Here $h_i$ is the hierarchy level of station $i \in V$: The larger the station, the smaller $h_i$.

### 4.4 Refining the Model: Turnover Times

The latter includes waiting times on entry and exit tracks, sorting times, and times for coupling and decoupling the trains, which all can be well estimated for each station $i$ from historical data. One component however cannot be taken as a constant, namely the waiting time on a sorting track, until enough cars are gathered to start the assembly of a new train. This time depends on the number of trains per time period that are created, and which is the variable $n_{i,j}$ in our model. Hence a more precise model would not take $u_i$ as a constant, but rather $u_{i,j}(n_{i,j})$ as a variable (or a function) depending on the number of trains:

$$\forall (i,j) \in A : u_{i,j} = \frac{24}{n_{i,j}}. \tag{14}$$

Due to these constraints our problem has turned into a nonlinear (nonconvex) mixed-integer problem. Since it is a combinatorial problem with otherwise linear structure we rather try to reformulate (14) as a linear constraint instead of using nonlinear solution techniques that are faster on the nonlinear part but then struggle with the integer decisions.

We made experiments with several linearization techniques [13]. Even in the best formulation we found the solution times increase by a large factor (10-100 times slower).

## 5 Conclusions and Future Work

We presented a linear mixed-integer model for a strategic freight car scheduling problem that arises at Deutsche Bahn. We gave some ways to reformulate the model in order to achieve better results faster when using a numerical solver for the resulting MIPs. An outlook to further refinements of the model that lead to nonlinear problems was given. Currently we test our implementations of the models and variants presented in this extended abstract. Computational results will thus be given in a forthcoming full article.

Our future work aims at an improvement of the lower bound, since we believe that the upper bound (feasible solution) is already quite satisfying. To this end, we are going to use a path based approach in a branch-and-price (column generation) way.

**Acknowledgements:** We gratefully acknowledge the funding of our joint research project OVERSYS by the German Federal Ministry for Science and Education (BMBF), research grant “Mathematics for Innovations in Industry and Services”, 5th funding period (2007-2010). We thank our industrial partners at Deutsche Bahn, Department GSU1, Gerald Pfau and Jörg Wolfner, and at DB Schencker Rail, Alexander Below, Christian Liebchen, and Thomas Klingler for their collaboration. We thank our academic partners within OVERSYS: Uwe Zimmermann and Ronny Hansmann (Technische Universität Braunschweig), Uwe Clausen, Annette Chmielewski, and Jens Baudach (Fraunhofer Institut für Materialfluss und Logistik), and Christoph Helmberg and Frank Fischer (Technische Universität Chemnitz). Last not least we thank Andreas Bley (Technische Universität Berlin) for pointing out the similarity of rail freight and telecommunication network planning (see Section 4.2).

## References

1. Assad, A.: Analytical models in rail transportation: An annotated bibliography. INFOR **19** (1981) 59 – 80
2. Ahuja, R.: Solving real-life railroad blocking problems. Interfaces (2007)
3. Assad, A.: Modelling of rail networks: toward a routing/makeup model. Transportation Research B **14** (1980) 101 – 114
4. Assad, A.: Analysis of rail classication policies. INFOR **21** (1983) 293 – 314
5. Barnhart, C., Hong, J., Vance, P.: Railroad blocking: A network design application. Operations Research **48** (2000) 603 – 614
6. Bodin, L., Golden, B., Schuster, A.: A model for the blocking of trains. Transportation Science B **14** (1980) 115 – 120
7. Crainic, T., Ferland, J.A., Rousseau, J.M.: A tactical planning model for rail freight transportation. Transportation Science **18** (1984) 165 – 184
8. Crainic, T., Rousseau, J.M.: Multicommodity, multimode freight transportation: a general modeling and algorithmic framework for the service network design problem. Transportation Research B **20** (1986) 225 – 242
9. Keaton, M.: Designing optimal railroad operating plans: Lagrangian relaxation and heuristic approaches. Transportation Research B **23** (1992) 330 – 345

10. Newton, H.: Network Design Under Budget Constraints with Application to the Railroad Blocking Problem. PhD thesis, Auburn University, Auburn, Alabama, USA (1997)
11. Newton, H., Barnhart, C., Vance, P.: Constructing railroad blocking plans to minimize handling costs. Transportation Science **32** (1998) 330 – 345
12. Eisenblätter, A., Fledderus, E., Fügenschuh, A., Geerdes, H.F., Heideck, B., Junglas, D., Koch, T., Kürner, T., Martin, A.: Mathematical methods for automatic optimization of umts radio networks. Technical report, Zuse Institut Berlin (ZIB), Berlin, Germany (2003) Online available at URL `http://momentum.zib.de`.
13. Jost, F.: Approximation einer Hyperbel in der diskreten Optimierung. PhD thesis, Technische Universität Darmstadt, Darmstadt, Germany (2009) In German.