

Shunting of Passenger Train Units: an Integrated Approach

Alexander Schrijver¹, Ramon M. Lentink^{2,*},
Leo G. Kroon^{2,3}

¹ *Centrum voor Wiskunde en Informatica, the Netherlands*

² *Erasmus Center for Optimization in Public Transport (ECOPT),
Erasmus University Rotterdam, the Netherlands*

³ *NS Reizigers, Department of Logistics, the Netherlands*

* *Corresponding author, email: rlentink@fbk.eur.nl*

Abstract

In this paper, we describe a new model for the Train Unit Shunting Problem. This model is capable of solving the matching and parking subproblems in an integrated manner, requiring a reasonable amount of computation time. Furthermore, the model incorporates complicating details from practice, such as trains composed of several train units and tracks that can be approached from two sides. Computational results are presented for real-life cases of NS Reizigers, the main Dutch passenger railway operator.

1 Introduction

During the rush hours, the rolling stock of a passenger railway operator is typically operating the timetable or it is in maintenance. However, outside the rush hours, an operator usually has a surplus of rolling stock. The idle rolling stock can be parked at a shunt yard in order to be able to fully exploit the main railway infrastructure by other trains. Especially during the night, most rolling stock has to be parked, since usually there are just a few passenger night train services. In the Netherlands, mainly freight trains operate at night.

The Dutch railway network consists of approximately 3,000 kilometers of railway track and some 400 stations. This infrastructure is used on a daily basis by a total of 4,500 scheduled train services on lines ranging in length between 15 and over 200 kilometers. Approximately 1 million passengers travel more than 40 million kilometers in total each workday, which makes the Dutch network one of the most heavily utilized networks in the world.

In the Netherlands most train services are operated by train units, which are classified according to types and subtypes. Train units can move bi-directionally without the need for locomotives. Only train units of the same type can be combined to form longer train compositions, taking into account certain restrictions on the length of the resulting train. Subtypes belonging to the same type are discerned from each other by their numbers of carriages per train unit. The



Figure 1: An ICM train unit with 3 carriages (ICM 3) and one carriage.

different subtypes of train units have different characteristics such as seating capacity and length. The upper part of Figure 1 shows an example of a Dutch train unit with 3 carriages, while the lower part depicts one carriage. This particular type of train unit (ICM) consists of subtypes with 3 or 4 carriages, which are typically used for intercity services.

The process of parking train units at a shunt yard together with several related processes is called shunting and the corresponding planning problem is called the *Train Unit Shunting Problem* (TUSP). A major complicating issue is the fact that train units are strongly restricted in their movements by the railway infrastructure. Another characteristic is time as a restrictive resource for shunting. For example, for safety reasons it is mandatory to respect a certain minimum headway time between two train movements on the same track. Moreover, railway passenger services are typically described by a timetable with planned times and exact compositions of the arriving and departing train services. Finally, it is common to solve the TUSP station by station and for a 24-hour period.

During the night, the goal of shunting is to select the positions and compositions of the trains at the shunt yard in such a way that the operations in the next morning can start up as smoothly as possible. The created shunt plans should be robust, since it is almost sure that disruptions will occur in real time. When this happens, the plans should need as little changes as possible.

In general, train units of the same subtype can be used interchangeably. This flexibility implies that, given a timetable with times and exact compositions of the arriving and departing train services, a planner has to determine a matching of arriving and departing units at a station. A large part of this matching is already prescribed by the timetable, e.g. the matching of the arriving and departing train units of through train services, which continue passenger service after a short dwell time.

Figure 2 shows the layout of station Zwolle, which is a station in the north-eastern part of the Netherlands. The left side of the station is the A-side, while the right side is the B-side. The black areas in the figure represent the platforms, while next to the platform tracks and around those tracks several shunt tracks are located. Many of these tracks can be approached from both sides.

The choice to park a train unit on a particular shunt track has several implications. First, if train units on a track belong to different types or subtypes, then the order of the train units is important. In this case, it is important that no train unit is blocking the arrival or departure of another one at the same track. Second, in practice, certain routes and shunt tracks are preferred over others. Here, a track is less desired if it is used frequently for other purposes, e.g. for through train services or for temporary parking of train units. Third, not all tracks have catenary installed, where tracks without catenary can only

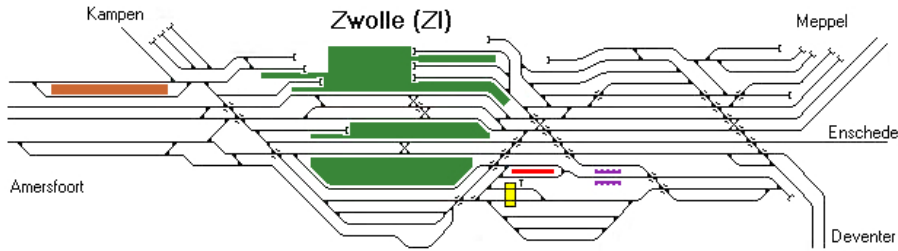


Figure 2: The layout of station Zwolle.

be used for parking diesel train units. Other aspects that also play a role, but are considered outside the scope of this paper, are the routing of train units between platforms and shunt tracks and the availability of crew to carry out the resulting shunt activities within certain time intervals.

The main contribution of this paper is that we provide a model for solving this shunting problem for general track configurations, where trains might consist of several units. In the model, the matching and parking subproblem are treated simultaneously.

This paper continues with an in-depth problem description, containing an overview of relevant literature. Then, we discuss a model for a specific configuration of the tracks. In addition, we investigate the influence of tracks with multiple subtypes of train units in this model and we extend the model with general track configurations. Computational results for several real-life cases precede the conclusions.

2 Problem Description

To start with, we introduce the terms *arriving (departing) shunt units* for train units that need to be parked at (supplied from) the shunt tracks. We say that a *crossing* occurs at a shunt track whenever a train unit i obstructs the arrival or departure of a train unit j . Then, we formally define the *Train Unit Shunting Problem* (TUSP) as follows:

Definition 1. Given

- a railway station,
- a shunt yard, usually geographically separate from the station,
- a timetable, with for each train service the arrival and / or departure time and platform at the involved station, and its composition,

the *Train Unit Shunting Problem* consists of (i) matching the arriving and departing shunt units, and (ii) parking these shunt units on the shunt tracks, such that the total shunting costs are minimal and no crossing occur.

A solution to the TUSP assigns arriving shunt units to departing ones and to a shunt track. Moreover, if such a track can be approached from both sides, the solution also describes the arrival and departure sides for each train unit parked at the track.

The cost of a shunt plan is determined by the number of splits of train units resulting from the same train service and the number of tracks with multiple subtypes of train units parked at it. The number of splits is a proxy for the amount of resource consumption at the station, such as crew and railway infrastructure. Indeed, if two train units from the same train service are split, this results in two routes to shunt tracks, which requires two drivers and a longer reservation of the infrastructure, compared to keeping the units together. The main characteristics of the shunting problem are the following:

- Arrivals and departures of train units may be mixed in time. This implies that, within the planning horizon, the first departure may take place before the last arrival has taken place.
- Shunt units might belong to different types or subtypes (and thus have different lengths). The type of a unit may restrict the set of shunt tracks where the unit can be parked. For example, electrical train units can only be parked at a track with catenary.
- Shunt tracks may have different types and lengths. The type of a track determines how a unit may approach the track. Some tracks may be approached from one side only. These tracks will be called *LIFO tracks*. Other tracks may be approached from both sides, which will be called *free tracks*. Note that for these track configurations, it is possible to arrive from one side and depart from the other, as well as arriving from and departing to the same side.

As opposed to similar versions of this problem already known in literature, see e.g. Blasum et al. [2000] and Winter [1999], subtype mismatches are not allowed in Dutch railway rolling stock deployment. An important reason for this restriction is the fact that previous planning processes already decided upon the exact configurations of the train services and local changes in these configurations are highly undesired.

2.1 Literature overview

Cordeau et al. [1998] provide a recent overview of the use of Operations Research in railway systems focusing on train routing and scheduling problems.

An elaborate introduction to the TUSP, including a solution approach and computational results, can be found in Freling et al. [2002]. Here, the matching and parking subproblems are solved separately, nevertheless resulting in solutions with high quality.

Some special cases of the TUSP have been dealt with by Winter and Zimmermann [2000] and Blasum et al. [2000] for dispatching trams in a depot. Winter [1999] theoretically extends this approach with length restrictions and mixed arrivals and departures. Moreover, he also discusses an application to a bus depot, including computational results.

Di Stefano and Kóci [2003] study the computational complexity of several variants of subproblems of the TUSP. Furthermore, they also present algorithms for solving some of these subproblems, including bounds on the objectives and on the complexity of the algorithms.

Furthermore, Gallo and di Miele [2001] discuss an application for dispatching buses in a depot, with an extension of their models to take into account mixed arrivals and departures. Another application of bus dispatching is described in Hamdouni et al. [2004]. Here, robust solutions are emphasized by having as little different types of buses as possible in one lane, and within one lane by grouping together the buses of the same type as much as possible.

Tomii et al. [1999] and Tomii and Zhou [2000] propose a genetic algorithm that takes into account some related processes of TUSP. However, their parking problem is of a less complex nature, since in their context at most one train unit can be parked on a shunt track at the same time.

Lentink et al. [2003] introduce the routing subproblem of the TUSP, accompanied by a solution methodology and computational results. Note that this routing subproblem lies outside the scope of the current paper.

Zwaneveld [1997] studied a routing problem for train units over railway infrastructure at a station, which is strongly related to the TUSP. In this problem, one is looking for a set of routes and platforms for train services in a one hour period where arrival and departure times are fixed. This approach is applied to a number of railway stations in the Netherlands.

Lübbecke and Zimmermann [2003] discuss a related problem that arises at an in-plant private freight railroad. In this problem, one assigns transportation requests to certain regions of the in-plant railroad and selects cars of specific types from a shunt track in this region for servicing a specific request. However, the authors only discuss LIFO tracks and assume that there is no prescribed order of different types of cars in a train. In addition, it is assumed that there are no limitations for the temporary parking of cars, when these are not servicing a request. In addition, the subject of the paper by He et al. [2000] is the separation of train units from arriving freight trains, sorting these according to their destination and finally combining them to form new departing trains, which resembles the matching of arriving and departing shunt units without parking. This is a typical problem in railfreight logistics.

Dahlhaus et al. [2000] discuss the related problem of rearranging carriages in a freight train to group them by destination. Their goal is to use a minimum number of tracks for this rearrangement and they show that this problem is \mathcal{NP} -hard.

3 A Basic Model with LIFO tracks

To start with, we discuss a model for the TUSP in the special case that all tracks are LIFO tracks. Before discussing this model, we will introduce some notation.

The set S contains all the shunt tracks where units can be parked. For each $s \in S$, c_s gives the length of track s . Each station has an A-side and a B-side. A shunt track s can be opened at the A-side, which is the side closest to the A-side of the station, the B-side or at both sides. This uniformly defines the side of a shunt track. In this section we assume that the tracks are only opened at the A-side.

We define $T = \{1, \dots, |T|\}$ as the set of train units that arrive or depart. For each train unit $t \in T$, we know whether it is arriving or departing. Therefore, we can partition the sequence T in a set of arriving train units, denoted by

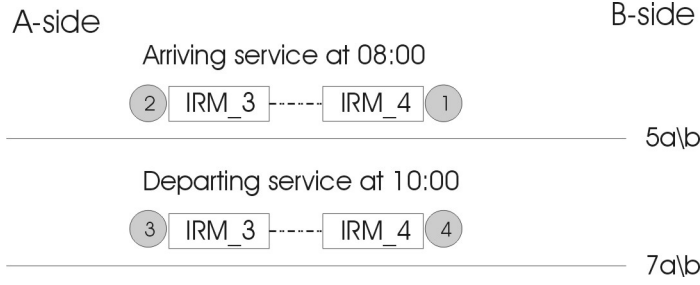


Figure 3: An example of the ordering of train units.

T_+ , and a set of departing train units, denoted by T_- , with $T = T_+ \cup T_-$ and $T_+ \cap T_- = \emptyset$. Moreover, we introduce Y as the set of subtypes of train units and the mapping τ_t , which maps a train unit $t \in T$ to its subtype $\tau \in Y$. We assume that the arrivals and departures are in balance for each subtype. Also, the mapping l_t maps a unit $t \in T$ to its length. For each $t \in T$, we know the train service i_t in which it arrives at or departs from the station.

As mentioned before, the timetable consists of a sequence of arriving and departing train services. For each service, we know the planned time of arrival or departure as well as the exact composition of the train operating this service. Similar to the A-side of a track, we introduce the A-side of a train as the side of the train which is closest to the A-side of the station, whenever the train is within the boundaries of the station. The set T is sorted according to the partial ordering $<$ on the train units. By definition $X_i < X_j$ if and only if one of the following conditions is satisfied:

1. Unit X_i arrives or departs in a train with an earlier planned time than the train to which unit X_j belongs.
2. Arriving units X_i and X_j result from the same train and X_j is closer to the A-side of the train than X_i .
3. Departing units X_i and X_j depart in the same train and X_i is closer to the A-side of the train than X_j .

Case 2 states that an arriving train composition $X_1 \cdots X_k$, where X_1 is closest to the A-side of the train and X_k is farthest from it, is ordered as $X_k < \dots < X_1$ in T , since this is the order in which the units arrive at a shunt track opened at the A-side. In Figure 3 the arriving train at 08:00 has composition IRM_3 IRM_4 and is ordered as IRM_4 $<$ IRM_3. Case 3 states that a departing train service $X_1 \cdots X_k$ is ordered as $X_1 < \dots < X_k$ in T , since the units leave the shunt track via the A-side in this order. Figure 3 gives an example of an arriving and a departing train service at the platforms 5a\b and 7a\b. For clarity, we numbered the train units ①,②,③,④ which results in $T = \{\textcircled{1}, \textcircled{2}, \textcircled{3}, \textcircled{4}\}$ with order $\textcircled{1} < \textcircled{2} < \textcircled{3} < \textcircled{4}$. The relation $X_i \leq X_j$ holds if and only if $X_i < X_j$ or $X_i = X_j$.

We define I as the set of pairs of train units $t, t+1 \in T$ that arrive or depart in the same train service, i.e. $i_t = i_{t+1}$. Finally, we define L as the set of pairs

of train units that can be matched:

$$L := \{(t, u) \mid t \in T_+, u \in T_-, t < u, \tau_t = \tau_u\}. \quad (1)$$

We assume that the conditions in the definition of the set L are sufficient for a feasible matching of two train units, although this is not necessarily the case. Indeed, even if the previous conditions are met, a train unit t cannot be matched to a train unit $u > t$ if no train unit of subtype τ_t is present at some point in time between t and u . The set X of potential crossings is defined as:

$$X := \{(t, u), (v, w) \mid ((t, u), (v, w)) \in L^2, t < v < u < w\}$$

Given these definitions, we introduce the following decision variables:

$$\begin{aligned} z_{t,s} &= \begin{cases} 1 & \text{if train unit } t \text{ is parked at or retrieved from track } s; \\ 0 & \text{otherwise.} \end{cases} \\ x_{t,u,s} &= \begin{cases} 1 & \text{if arriving train unit } t \text{ is matched to departing unit } u \\ & \text{and parked at track } s; \\ 0 & \text{otherwise.} \end{cases} \\ b_{t,s} &= \begin{cases} \text{the length of the train units at track } s \text{ after the arrival} \\ \text{or departure of unit } t. \end{cases} \\ d_t &= \begin{cases} 1 & \text{if units } t \text{ and } t+1 \text{ are related to the same train service} \\ & \text{and are parked at or retrieved from different tracks;} \\ 0 & \text{otherwise.} \end{cases} \\ m_{\tau,s} &= \begin{cases} 1 & \text{if at least one unit of subtype } \tau \text{ is parked at track } s; \\ 0 & \text{otherwise.} \end{cases} \\ n_s &= \text{the number of subtypes } \tau \text{ in excess of 1 parked at track } s. \end{aligned}$$

The penalty on the variables n_s models a preference for solutions with less different subtypes parked at a track, which adds to a more robust solution in practice. After introduction of the weights D for splitting of trains and N for tracks with multiple subtypes of units, the model reads:

$$\text{minimize } D \sum_{t \in T} d_t + N \sum_{s \in S} n_s \quad (2)$$

$$\text{subject to } \sum_{s \in S} z_{t,s} = 1 \quad \forall t \in T, \quad (3)$$

$$\sum_{u:(t,u) \in L} x_{t,u,s} = z_{t,s} \quad \forall t \in T_+, s \in S \quad (4)$$

$$\sum_{u:(u,t) \in L} x_{u,t,s} = z_{t,s} \quad \forall t \in T_-, s \in S \quad (5)$$

$$x_{t,u,s} + x_{v,w,s} \leq 1 \quad \forall s \in S, \{(t, u), (v, w)\} \in X \quad (6)$$

$$b_{t,s} = b_{t-1,s} + l_t z_{t,s} \quad \forall t \in T_+, s \in S \quad (7)$$

$$b_{t,s} = b_{t-1,s} - l_t z_{t,s} \quad \forall t \in T_-, s \in S \quad (8)$$

$$b_{t,s} \leq c_s \quad \forall t \in T_+, s \in S \quad (9)$$

$$d_t \geq z_{t,s} - z_{t+1,s} \quad \forall s \in S, (t, t+1) \in I \quad (10)$$

$$z_{t,s} \leq m_{\tau_t,s} \quad \forall s \in S, t \in T \quad (11)$$

$$\sum_{\tau \in Y} m_{\tau,s} \leq n_s + 1 \quad \forall s \in S \quad (12)$$

$$z_{t,s} \in \{0, 1\} \quad \forall t \in T, s \in S \quad (13)$$

$$x_{t,u,s} \in \{0, 1\} \quad \forall (t, u) \in L, s \in S \quad (14)$$

In this model, restrictions (3) state that each train unit needs to be parked at a track. Restrictions (4) state that each arriving train unit is parked at a track and matched to a departing unit. This also holds for restrictions (5) for departing train units. In addition, restrictions (6) prohibit crossings. Restrictions (7) are used for administrating the length of the units parked at a track at arrival of a train unit. Again, restrictions (8) are the same, but for departing units. Restrictions (9) ensure that the total length of the units parked at a track never exceeds the capacity of the track. Moreover, restrictions (10), (11), and (12) are used for determining the right values of the decision variables in the objective function (2). Finally, restrictions (13) and (14) are integrality restrictions on the z and x variables.

3.1 Improvements on the basic model

The bottleneck for this model is the large number of restrictions (6), which equals $|S| \times |X|$ and amounts to 339.739 restrictions in a particular real-life application at station Zwolle. We discuss two approaches to tackle this problem: aggregation of these crossing restrictions and improving and strengthening the possibly aggregated restrictions to *cliques*.

For a convenient discussion of the aggregation of crossing restrictions, we introduce the set Z as the set of pairs $(v, u) \in T^2$ such that there exist $(t, u), (v, w) \in L$ with $t < v < u < w$. Z describes an arrival, followed by a departure, which might be involved in a crossing. Given Z , we can replace restrictions (6) with:

$$\sum_{t < v: (t, u) \in L} x_{t,u,s} + \sum_{w > u: (v, w) \in L} x_{v,w,s} \leq 1 \quad \forall (v, u) \in Z, s \in S \quad (15)$$

These restrictions sharpen the restrictions (6) and are far less in number, 17.005 for the same real-life application. Given this replacement, we define Model 1 by (2)-(5), (15), (7)-(14). In the remainder of this section, we try to reduce the number of restrictions even further.

In fact, the restrictions (6) and (15) are special cases of the well-known clique inequalities. Before discussing these inequalities, we define the graph $G = (L, E)$. Here, L is defined in (1) and $E := \{(t, u), (v, w) \mid ((t, u), (v, w)) \in L^2, t \leq v < u \leq w\}$, which is a slight extension of X . In this paper, a clique is a subset K of L , such that the elements in K are pairwise adjacent in G . Now, we can improve the crossing restrictions (15), to the following ones:

$$\sum_{(t, u) \in K} x_{t,u,s} \leq 1 \quad \forall \text{ cliques } K, s \in S \quad (16)$$

These restrictions (16) are at least as strong as restrictions (15). Indeed, for each $(v, u) \in Z$, the set

$$\{(t, u) \in L \mid t < v\} \cup \{(v, w) \in L \mid w > u\}$$

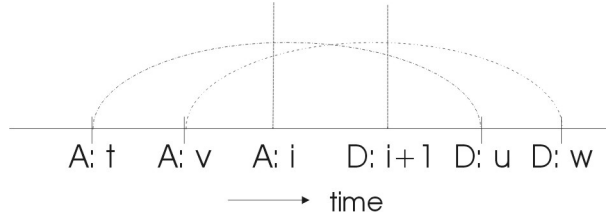


Figure 4: An example of a crossing at a specific LIFO track.

is a clique. Real-life applications require very many maximal cliques (77.868.374 for the previously mentioned application), which invalidates explicit enumeration of these restrictions in a solution approach. However, in order to reduce the number of restrictions, we can also find an implicit manner to describe all clique inequalities. Therefore, we define $T' := \{i \in T_+ \mid i+1 \in T_-\}$ as the set of pairs of an arriving train unit directly followed by a departing train unit. Given $i \in T'$, we define $L_i := \{(t, u) \in L \mid t \leq i < i+1 \leq u\}$, which is the set of possible matchings arriving no later than i , and not departing before $i+1$. In the following discussion we will need the well-known notion of *comparability graphs*. A graph H is a comparability graph if a partial order \preceq on the vertices of H exists, such that for each pair of vertices $(u, v) \in H$:

$$u \text{ and } v \text{ are adjacent in } H \Leftrightarrow u \preceq v \text{ or } v \preceq u. \quad (17)$$

Now, we can state the following two lemmas:

Lemma 1. *For each clique K there exists an $i \in T'$ such that $K \subset L_i$.*

Proof. Determine the largest v for which $(v, w) \in K$ for some w and the smallest u for which $(t, u) \in K$ for some t , according to the ordering $<$. Because $t \leq v < u \leq w$, we know that $v < u$. Moreover, $v \in T_+$ and $u \in T_-$. Therefore, an $i \in T'$ with $v \leq i < i+1 \leq u$ exists. ■

Lemma 2. *The subgraph G_i of G induced by L_i is a comparability graph for each $i \in T'$.*

Proof. Define the following partial order on L_i :

$$(t, u) \preceq (v, w) \Leftrightarrow t \leq v \text{ and } u \leq w$$

This order satisfies requirement (17). ■

The relation $(t, u) \prec (v, w)$ holds if and only if $(t, u) \preceq (v, w)$ and $(t \neq v \text{ or } u \neq w)$. If $(t, u) \prec (v, w)$, then a crossing occurs if both (t, u) and (v, w) are assigned to the same LIFO track, which is clarified in Figure 4. Since we only consider A-side open LIFO tracks in this section, the unit v is closest to the A-side after arrivals t and v . However, unit u , which is matched to unit t , needs to depart before unit w , which is matched to unit v . Therefore, the arriving unit v is obstructing the departure of unit u .

In order to describe the clique inequalities in the subgraph G_i in an efficient way, we introduce the variables $y_{v,w,s,i}$ for each $s \in S, i \in T'$, and $(v, w) \in L_i$:

$$y_{v,w,s,i} = \begin{cases} 1 & \text{if a conflicting matching } (t, u) \in L_i, (t, u) \prec (v, w), \\ & \text{is parked at track } s; \\ 0 & \text{otherwise.} \end{cases}$$

For each $s \in S, i \in T'$, and $(t, u), (v, w) \in L_i$, the following inequalities ensure that the $y_{v,w,s,i}$ variables are set to the appropriate values:

$$y_{v,w,s,i} \geq y_{t,u,s,i} + x_{t,u,s} \quad \text{if } (t, u) \prec (v, w) \quad (18)$$

$$y_{v,w,s,i} + x_{v,w,s} \leq 1 \quad (19)$$

Note that it suffices to restrict restrictions (18) to (v, w) directly subsequent to (t, u) according to \prec . Moreover, restrictions (19) are required only for pairs (t, u) without successor.

Lemma 3. *Restrictions (16) are equivalent to the combination of restrictions (18) and (19).*

Proof. For both parts, we start with a maximal clique K without loss of generality. According to Lemma 1, an $i \in T'$ exists such that $K \subset L_i$. In addition, we use the order \prec on the elements of K :

$$(t_1, u_1) \prec (t_2, u_2) \prec \dots \prec (t_k, u_k), \quad (20)$$

such that (t_{j+1}, u_{j+1}) is the successor of (t_j, u_j) and (t_k, u_k) has no successor. Now (18), (19) \Rightarrow (16) follows directly from:

$$\begin{aligned} \sum_{(t,u) \in K} x_{t,u,s} &= \sum_{j=1}^k x_{t_j, u_j, s} \leq \sum_{j=1}^{k-1} (y_{t_{j+1}, u_{j+1}, s, i} - y_{t_j, u_j, s, i}) + x_{t_k, u_k, s} \\ &= y_{t_k, u_k, s, i} - y_{t_1, u_1, s, i} + x_{t_k, u_k, s} \leq y_{t_k, u_k, s, i} + x_{t_k, u_k, s} \leq 1. \end{aligned}$$

Secondly, (16) \Rightarrow (18), (19). For t, u, s, i let the maximal clique be attained by clique K . Let $K' := K \cup \{t, u\}$. Then

$$y_{v,w,s,i} \geq \sum_{\substack{(a,b) \in K' \\ (a,b) \prec (v,w)}} x_{a,b,s} = \sum_{\substack{(a,b) \in K \\ (a,b) \prec (v,w)}} x_{a,b,s} + x_{t,u,s} = y_{t,u,s,i} + x_{t,u,s}.$$

This completes the proof. ■

We define this extension as Model 2. It consists of (2)-(5), (18)-(19), (7)-(14), where compared to Model 1 restrictions (6) have been replaced by restrictions (18)-(19). It requires far less constraints, 2.684 for our typical application, but far more variables, 79.002 for the same application.

4 Restricting the Number of Mixed Tracks

Suppose, one would know beforehand that only units of one subtype would be parked at a certain track. In this case, the aggregated crossing restrictions (15), or their equivalents (18)-(19), are superfluous for this track. Indeed, since train units of the same subtype can be used interchangeably, the order of the train units at such a track is irrelevant. This would add to the robustness of a solution, it being better able to handle disruptions in the operations.

However, we do not want to choose beforehand which tracks will be *mixed tracks*, containing several subtypes, and which ones will not. This can be achieved by introducing *virtual tracks* and by assigning these virtual tracks to physical tracks. We let S represent the set of virtual tracks. Moreover, we introduce the set P of physical tracks, with $|S| = |P|$. A matching from S to P assigns the virtual tracks to the physical ones. Let S' be the set of mixed virtual tracks and let $S'' = S \setminus S'$ be the set of non-mixed virtual tracks. Then it suffices to activate the crossing restrictions (15) only for the mixed tracks in S' instead of for all virtual tracks in S :

$$\sum_{t < v: (t,u) \in L} x_{t,u,s} + \sum_{w > u: (v,w) \in L} x_{v,w,s} \leq 1 \quad \forall (v,u) \in Z, s \in S' \quad (21)$$

Similar changes for restrictions (18) and (19) are required. Of course, only units of one subtype can be parked at a track in S'' , which results in this additional set of restrictions:

$$n_s = 0 \quad \forall s \in S'' \quad (22)$$

The matching can be described by the following decision variables:

$$r_{s,p} = \begin{cases} 1 & \text{if virtual track } s \in S \text{ is assigned to physical track } p \in P; \\ 0 & \text{otherwise.} \end{cases}$$

We replace the parameters c_s describing the length of a track $s \in S$ with c_p representing the length of the physical track $p \in P$. For these decision variables, the restrictions that make it a matching are:

$$\sum_{s \in S} r_{s,p} = 1 \quad \forall p \in P \quad (23)$$

$$\sum_{p \in P} r_{s,p} = 1 \quad \forall s \in S \quad (24)$$

$$r_{s,p} \in \{0, 1\} \quad \forall s \in S, p \in P \quad (25)$$

Given this matching, we rewrite the restrictions (9) on the track capacity as:

$$b_{t,s} \leq \sum_{p \in P} c_p r_{s,p} \quad \forall t \in T_+, p \in P \quad (26)$$

This results in Model 3 consisting of (2)-(5),(18)-(19),(7),(8),(21)-(26),(10)-(14). Note that (2)-(5),(18)-(19),(7),(8),(10)-(14) remain unchanged, and are defined in terms of virtual tracks.

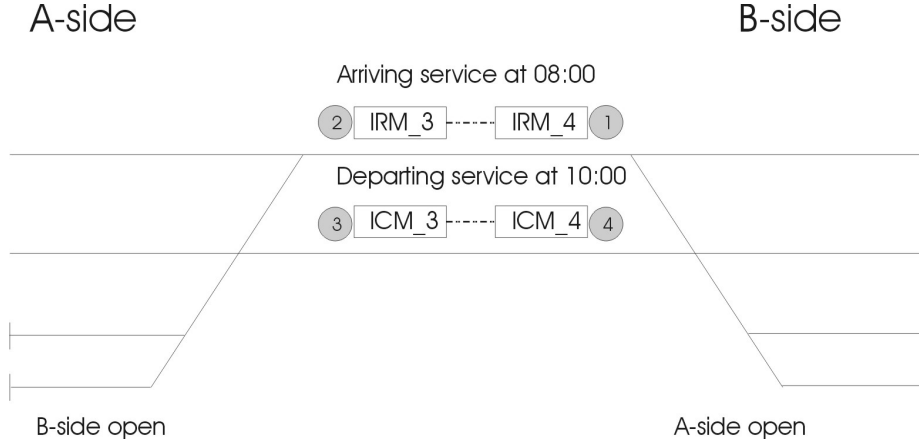


Figure 5: Tracks opened at different sides require a different ordering of train units within a train.

5 Extending the Basic Model to Free Tracks

Until now, we have assumed that the shunt tracks can only be approached from the A-side of the track. In this section, we extend our model in such a way that we can also consider shunt tracks which can be approached from the B-side, and even from both sides of the track.

Before describing these extensions of our model, we need to take a closer look at the ordering $<$ on the train units. In Figure 5, both the IRM.3 and IRM.4 as well as the ICM.3 and ICM.4 units arrive respectively depart in one train. The partial ordering $<$ results in the order ①,②,③,④. For a track, open at the B-side, this order should be: ②,①,④,③. This ordering is wrong for compositions with multiple train units and needs to be reversed in this case. This results in the partial ordering $<'$ of the train units. The relation $X_i <' X_j$ holds if and only if one of the following conditions is satisfied:

1. Unit X_i arrives or departs in a train service with an earlier planned time than the service to which unit X_j belongs.
2. Arriving unit X_i is closer to the A-side of the train than arriving unit X_j .
3. Departing unit X_j is closer to the A-side of the train than departing unit X_i .

Note that both orderings $<$ and $<'$ are the same for train units in different train compositions and only differ for train units in the same train composition.

We partition the set P into the set P_0 with tracks which can be approached from the A-side only, P_1 with tracks which can be approached from the B-side only, and $P \setminus \{P_0 \cup P_1\}$ with the tracks that can be approached from both sides.

As a straightforward extension, one could add two indices to the decision variables $x_{t,u,s}$, resulting in $x_{t,u,s,d,e}$ variables, where d describes the arrival side at track s for unit t , and e describes the departure side at track s for unit u . Crossing constraints, similar to restrictions (6), can be determined. Moreover, some restrictions for the sides of the tracks which can be used are required.

However, since the number of decision variables roughly increases by a factor 4, this approach will not result in a model that is able to solve real-life instances. Therefore, we discuss a different approach in the next paragraph.

5.1 The arrival and departure side as a decision variable

An alternative for modeling the free tracks is the introduction of one decision variable for each train unit, which indicates the side of a track via which the unit arrives or departs:

$$k_t = \begin{cases} 0 & \text{if train unit } t \text{ arrives or departs via the A-side of a track;} \\ 1 & \text{if train unit } t \text{ arrives or departs via the B-side of a track.} \end{cases}$$

Note that, compared to the straightforward extension, this results in a huge reduction in the number of decision variables, since the decision variables k_t make the additional indices d and e in the $x_{t,u,s,d,e}$ decision variables superfluous. Indeed, a variable k_t is used for all shunt tracks $s \in S$.

In order to extend the Models 1 and 2 to include free tracks, we need restrictions like $z_{t,s} - k_t \leq 0$ if track s is not accessible from the A-side, and $z_{t,s} + k_t \leq 1$ if track s is not accessible from the B-side. However, the model with virtual tracks requires the following restrictions:

$$z_{t,s} - k_t + \sum_{p \in P_1} r_{s,p} \leq 1 \quad \forall s \in S', t \in T \quad (27)$$

$$z_{t,s} + k_t + \sum_{p \in P_0} r_{s,p} \leq 2 \quad \forall s \in S', t \in T \quad (28)$$

Indeed, suppose that train unit t is parked at track s , resulting in $z_{t,s} = 1$ and virtual track s is assigned to a physical track $p \in P_1$, which can only be approached from the B-side. Then restrictions 27 ensure $k_t = 1$, ensuring that unit t will approach track s from the B-side. Restrictions (28) can be explained similarly.

For ease of discussion, we start with the crossing restrictions in the special case that each train consists of exactly one train unit. Note, that in this case both orders $<$ and $<'$ are the same. The resulting restrictions are generalized later on. For each $s \in S$, and $((t, u), (v, w)) \in L^2$, the conditions on k_t are:

$$\text{if } x_{t,u,s} = 1 \text{ and } x_{v,w,s} = 1 \text{ and } t < v < u < w, \text{ then } k_v \neq k_u \quad (29)$$

$$\text{if } x_{t,u,s} = 1 \text{ and } x_{v,w,s} = 1 \text{ and } v < t < u < w, \text{ then } k_t = k_u \quad (30)$$

Theorem 1. *If each train consists of exactly one unit, restrictions (29) and (30) are necessary and sufficient for describing the crossing constraints for $s \in S$, and $((t, u), (v, w)) \in L^2$.*

Proof. The theorem is equivalent to showing that for each (t, u) and (v, w) in L with $\{t, u\} \cap \{v, w\} = \emptyset$, there exist disjoint curves in $\mathbb{R} \times [0, 1]$ from (t, k_t) to (u, k_u) and from (v, k_v) to (w, k_w) if and only if restrictions (29) and (30) hold. A case where this is possible is represented in Figure 6.

Figure 4 gives the special case with $t < v < u < w$ and $k_v = k_u = 0$, where a crossing occurs indeed. Without loss of generality, we can assume $u < w$, and therefore $t < u < w$. Now, there are three possibilities for the position of

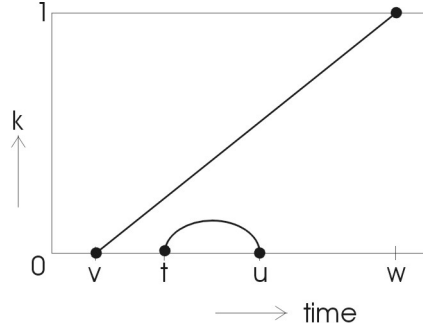


Figure 6: A case where (t, u) and (v, w) can be assigned to the same track

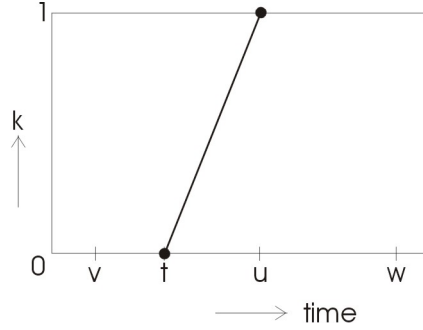


Figure 7: A graphical representation of restrictions (34).

v compared to t, v, w . If $u < v < w$ then such curves can be drawn, because the train units do not have a common time interval at track s . Moreover, if $t < v < u$, such curves exist only if $k_v \neq k_u$, conform Figure 4. In addition, if $v < t$, these curves exist only if $k_v = k_u$. Exactly these restrictions on the k variables are described by restrictions (29) and (30). Note that if $v > w$ then v arrives later than w , which contradicts $(v, w) \in L$. ■

We can rewrite the restrictions (29) and (30) in linear form for each $s \in S$, and $((t, u), (v, w)) \in L^2$, resulting in:

$$x_{t,u,s} + x_{v,w,s} \leq 3 - k_u - k_v \quad \text{if } t < v < u < w \quad (31)$$

$$x_{t,u,s} + x_{v,w,s} \leq 1 + k_u + k_v \quad \text{if } t < v < u < w \quad (32)$$

$$x_{t,u,s} + x_{v,w,s} \leq 2 - k_t + k_u \quad \text{if } v < t < u < w \quad (33)$$

$$x_{t,u,s} + x_{v,w,s} \leq 2 + k_t - k_u \quad \text{if } v < t < u < w \quad (34)$$

For example, restrictions (34) prohibit the situation depicted in Figure 7. When trains might consist of several train units, the restrictions (31)-(34) generalize to the following, again for each $s \in S$, and $((t, u), (v, w)) \in L^2$:

$$x_{t,u,s} + x_{v,w,s} \leq 1 + k_u + k_v \quad \text{if } t < v < u < w \quad (35)$$

$$x_{t,u,s} + x_{v,w,s} \leq 3 - k_u - k_v \quad \text{if } t < v < u < w \quad (36)$$

$$x_{t,u,s} + x_{v,w,s} \leq 2 - k_t + k_u \quad \text{if } v <' t, u < w \quad (37)$$

$$x_{t,u,s} + x_{v,w,s} \leq 2 + k_t - k_u \quad \text{if } v < t, u <' w \quad (38)$$

$$x_{t,u,s} + x_{v,w,s} \leq 3 + k_t - k_u - k_v \quad \text{if } t < v <' t, u <' w \quad (39)$$

$$x_{t,u,s} + x_{v,w,s} \leq 3 + k_w - k_u - k_v \quad \text{if } t <' v, u < w <' u \quad (40)$$

$$x_{t,u,s} + x_{v,w,s} \leq 2 - k_t + k_u + k_v \quad \text{if } v < t <' v, u < w \quad (41)$$

$$x_{t,u,s} + x_{v,w,s} \leq 2 - k_w + k_u + k_v \quad \text{if } t < v, u < w <' u \quad (42)$$

$$x_{t,u,s} + x_{v,w,s} \leq 3 + k_t - k_u - k_v + k_w \quad \text{if } t < v <' t, u < w <' u \quad (43)$$

Note that in restrictions (35) - (43) a relation like $u < w <' u$ implies that the train units u and w belong to the same train. These restrictions are explained by the following discussion on restrictions (40).

Restrictions (40) are only restrictive if $k_u = 1, k_v = 1$, and $k_w = 0$ since otherwise the righthand side is at least two, and the restriction is trivially fulfilled because $x_{t,u,s}$ and $x_{v,w,s}$ are binary. Moreover, we know that units u and w will leave in the same train. In addition, we know that $v <' u$ because otherwise $w <' u <' v$, which contradicts $(v, w) \in L$. The situation at track s after both arrivals t and v is given in Figure 8, which is independent of the value of k_t . In this figure, the arrow from (t, u) , resp. (v, w) , represents the fact that this unit needs to depart from the B-side, resp. A-side, of shunt track s .

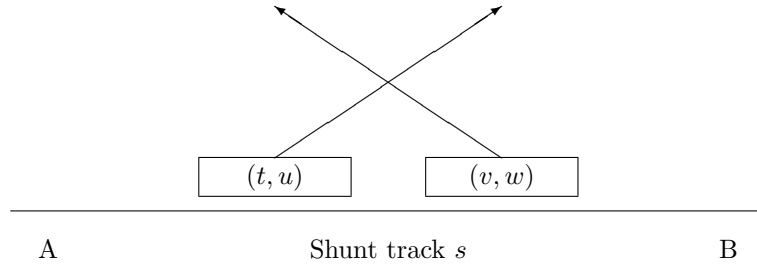


Figure 8: The situation at track s after both arrivals.

This results in a crossing because w needs to leave from the A-side of s , while u needs to leave from the B-side and the units are parked simultaneously at s .

This formulation assumes that train units arriving or departing in the same train service may end up on the same shunt track in a different order than in the train service. This occurs when, e.g. one train unit is routed via the A-side and the other via the B-side of the same shunt track. In practice, this never happens because of the additional complexity in the local operations. However, instances exist where only such impractical solutions are feasible.

6 Computational Results

1. Basic case: Zwolle, Enschede
2. The model described in the last section, with aggregation of the crossing constraints. Possibly restricting the free tracks to be open at their preferred side only (see also Trans. Sci. paper)

3. Restricting the number of mixed tracks (possibly with assigning subtypes to them before solving)

OPEN ISSUES:

1. HOW TO DESCRIBE REDUCTIONS IN RESTRICTIONS. REPORTING THE SIZES OF THE RESULTING MIXED INTEGER PROBLEMS IN THE COMPUTATIONAL RESULTS?
2. DESCRIBE RESTRICTING A TRACK TO A SPECIFIC TYPE BEFOREHAND.
3. WHICH MODELS AND DATA TO RUN FOR COMPUTATIONAL RESULTS? COMPARISON WITH TRANS. SCI. PAPER?
4. INTRODUCE RESTRICTIONS $p_t \geq z_{t+1,s} - z_{t,s}$ TO TIGHTEN THE FORMULATION?
5. ELABORATE ON REPLACEMENTS FOR RESTRICTIONS (35)-(43)?

7 Conclusions and Further Research

Contribution of this paper.

Train services have fixed arrival and departure times, but flexible arrival times at and departure times from the shunt tracks. For example, the departure time of an arriving shunt unit from a platform to a shunt track is flexible within a time interval starting at the arrival time of the unit at the platform and ending some time before the next arrival of another train service at the same platform.

References

- ULRICH BLASUM, MICHAEL R. BUSSIECK, WINFRIED HOCHSTÄTTLER, CHRISTOPH MOLL, HANS-HELMUT SCHEEL, AND THOMAS WINTER (2000). *Scheduling trams in the morning*. *Mathematical Methods of Operations Research*, 49(1):137–148.
- JEAN-FRANÇOIS CORDEAU, PAOLO TOTH, AND DANIELE VIGO (1998). *A survey of optimization models for train routing and scheduling*. *Transportation Science*, 32(4):380–404.
- ELIAS DAHLHAUS, PETER HORAK, MIRKA MILLER, AND JOSEPH F. RYAN (2000). *The train marshalling problem*. *Discrete Applied Mathematics*, 103(1–3):41–54.
- GABRIELE DI STEFANO AND MAGNUS LOVE KÖCI (2003). *A graph theoretical approach to the shunting problem*. *Electronic Notes in Theoretical Computer Science*, 92(1):16–33.
- RICHARD FRELING, RAMON M. LENTINK, LEO G. KROON, AND D. HUISMAN (2002). *Shunting of passenger train units in a railway station*. Technical Report ERS-2002-074-LIS, accepted for publication in *Transportation Science*, Erasmus University, Rotterdam, Netherlands.

- GIORGIO GALLO AND FEDERICO DI MIELE (2001). *Dispatching buses in parking depots*. *Transportation Science*, 35(3):322–330.
- MOHAMED HAMDOUNI, GUY DESAULNIERS, ODILE MARCOTTE, FRANÇOIS SOUMIS, AND MARIANNE VAN PUTTEN (2004). *Dispatching buses in a depot using block patterns*. Technical Report G-2004-051, GÉRAD, Montréal, Canada.
- SHIWEI HE, RUI SONG, AND SOHAIL S. CHAUDHRY (2000). *Fuzzy dispatching model and genetic algorithms for railyards operations*. *European Journal of Operations Research*, 124(2):307–331.
- RAMON M. LENTINK, PIETER-JAN FIOOLE, LEO G. KROON, AND COR VAN 'T WOUTD (2003). *Applying Operations Research techniques to planning of train shunting*. Technical Report ERS-2003-094-LIS, Erasmus University, Rotterdam, Netherlands.
- MARCO LÜBBECKE AND UWE T. ZIMMERMANN (2003). *Shunting minimal rail car allocation*. Technical report, Braunschweig University of Technology.
- NORIO TOMII AND LI J. ZHOU (2000). *Depot shunting scheduling using combined genetic algorithm and PERT*. In JEFF ALLAN, R.J. HILL, CARLOS A. BREBBIA, GIUSEPPE SCIUTTO, AND SATORU SONE (editors), *Proceedings of COMPRAIL 2000*, pages 437–446. Wit Press, Southampton.
- NORIO TOMII, LI J. ZHOU, AND NAOTO FUKUMARA (1999). *An algorithm for station shunting scheduling problems combining probabilistic local search and PERT*. In IBRAHIM IMAM, YVES KODRATOFF, AYMAN EL-DESSOUKI, AND MOONIS ALI (editors), *Multiple approaches to intelligent systems: 12th international conference on industrial and engineering applications of artificial intelligence and expert systems*, volume 1611 of *Lecture Notes in Artificial Intelligence*, pages 788–797. Springer-Verlag.
- THOMAS WINTER (1999). *Online and real-time dispatching problems*. Ph.D. thesis, Technical University, Braunschweig.
- THOMAS WINTER AND UWE T. ZIMMERMANN (2000). *Real-time dispatch of trams in storage yards*. *Annals of Operations Research*, 96:287–315.
- PETER J. ZWANEVELD (1997). *Railway planning - routing of trains and allocation of passenger lines*. Ph.D. thesis, Erasmus University, Rotterdam.