Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 47 (2020) 19–26

Transportation Research Procedia

www.elsevier.com/locate/procedia

22nd EURO Working Group on Transportation Meeting, EWGT 2019, 18-20 September 2019, Barcelona, Spain

Block planning for intermodal rail: Methodology and case study

Gianluca Morganti[a,b], Teodor Gabriel Crainic[a,c,*], Emma Frejinger[a,b], Nicolettta Ricciardi[a,d]

[a]CIRRELT Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation, P.O. Box 6128, Station Centre-ville, Montréal, H3C 3J7, Canada

[b]Dept. of Computer Science and Operations Research, Université de Montréal, P.O. Box 6079, Station Centre-ville, Montréal, H3C 3A7, Canada

[c]School of Management, Université du Québec à Montréal, P.O. Box 8888, Station Centre-Ville, Montréal, H3C 3P8, Canada

[d]Dept. of Statistical Sciences, Sapienza Università di Roma, Piazzale Aldo Moro 5, 00185 Rome, Italy

Abstract

Blocking constitutes an important rail freight transport operation, by which cars with potentially different origins and destinations are grouped to be moved and handled as a single unit, yielding economies of scale. We address the tactical block-planning problem for intermodal railroads, which has been little studied so far. We propose a new block service network design model explicitly considering the specificity of intermodal rail, and which may be solved using commercial software for realistic sizes. We illustrate the performance and interest of the proposed method through an extensive case study of a major North American railroad.

Keywords: Intermodal rail; Blocking; Service network design model; Case study

1. Introduction

Railroads are core elements of intermodal transportation, moving containers loaded with a broad variety of commodities over long distances in a cost-effective and environmentally sustainable manner. Efficient and profitable railroads require adequate planning of operations and resources. We focus on the tactical block-planning problem arising in intermodal rail transportation, which has been little, if at all, addressed until now.

A block is a group of cars with possibly different origins and destinations, which moves as a single unit between a pair of yards, without cars being handled individually at intermediate yards. Blocking aims to take advantage of economies of scale and reduce the cost of handling cars. A block is moved by a sequence of trains, while a car can be moved by one or a sequence of blocks between its origin and destination (OD). The classical train blocking problem consists in selecting the blocks to build and assigning cars with given OD pairs to blocks (e.g., Ahuja et al., 2007,

* Corresponding author. Tel.: +01-514-343-7143
E-mail address: TeodorGabriel.Crainic@CIRRELT.net

10.1016/j.trpro.2020.03.068

2005; Jha et al., 2008; Newton et al., 1998; Barnhart et al., 2000; Bodin et al., 1980). Blocking has also been addressed in the larger setting of tactically planning the service and schedule network of railroads (e.g., Crainic et al., 1984; Crainic and Rousseau, 1986; Zhu et al., 2014). Existing methods assume, however, that the demand of the blocking model is given a numbers of cars to be blocked. Consequently, they do not account for the important characteristics particular to intermodal rail, the most obvious one being demand expressed in numbers of containers of different types rather than in numbers of loaded cars. This means that container assignment to and loading on cars must be explicitly included in the blocking model. The multitude of car and container types results in many ways to load containers that must respect different loading rules, particularly when double stacking is allowed. Representing in a computationally efficient way the assignment of containers to cars within a tactical blocking model is particularly challenging. Our research objective is to contribute filling these gaps in knowledge and decision-support instruments, and propose a blocking model that considers several types of containers and cars and integrates the container-to-car assignment.

Section 2 describes the tactical block-planning problem for intermodal rail. The modelling framework is presented in Section 3. Section 4 is dedicated to the numerical experiments performed, and the presentation of the performance results and associated managerial insights. Finally, Section 5 concludes.

2. Problem setting

Railroads operate according to two levels of consolidation. First, cars are grouped together within blocks to be moved together, as a unique entity, from the origin of the block, the terminal where it is constructed, to its destination terminal, where it is dismantled. The cars making up a block do not necessarily have the same origins or destinations, and these are not necessarily the same as those of the block they are grouped into. They share, however, a part of their respective trips. The second consolidation level puts together blocks to make up trains. Consequently, a block travels from its origin to its destination either on a unique train or on a sequence of trains. In the latter case, the block is transferred as a unit between two consecutive trains at an intermediate terminal. Trains are made up either of a single block or a sequence of blocks. A train is made up at its origin terminal and dismantled at its destination terminal, where its blocks are either transferred to other trains or dismantled as well. The route of a train gives its origin, destination, and intermediate terminals where it stops to drop or pick up blocks, while its schedule gives the departure and arrival times at all the terminals, origin, destination, intermediate, on its route.

The trains and blocks make up a service network in space and time aimed at moving the cars holding the freight from the origin to the destination (OD) of the respective demand. Cars are of different types. In the general, non-intermodal, case, the selection of the car types assigned to each customer is part of the commercial transaction between the railroad and its customer. Consequently, the assignment decision, as well as the activities related to the loading of the cars, is outside the scope of the blocking problem. The latter is thus concerned with cars only, not the freight they carry. Cars carrying freight for a particular OD travel on a sequence of blocks and trains. At the origin terminal, cars undergo classification, a sorting operation to put them to the track where the block to which they are assigned is being built. When the block arrives at its destination and is dismantled, the cars it contained are either delivered to customers, if at their own destination (how this operation is performed is beyond the scope of this paper) or, in the case of non-intermodal traffic, are directed to the classification part of the terminal to be sorted and attached to their new block. The blocking problem aims to select the blocks to build, including the train sequence, and the classification strategy, that is the assignment of cars to blocks. It belongs, together with the selection of train services and resource (e.g., locomotives and cars) management strategies, to the tactical planning set of decisions, yielding the operation plan for the next "season" (Crainic and Kim, 2007). It also belongs to shorter-term planning decisions (e.g., the week) when plans need to be adjusted due to, e.g., incidents and unforeseen delays or demand variations.

Intermodal rail transportation follows the same general canvas, with a number of important characteristics that sets it apart and requires particular methodological developments. Reclassification of cars at intermediary terminals is seldom performed to avoid additional delays. Indeed, is not included in the blocking plan of the major North American railroad we collaborate with. The most important difference, however, concerns the need to explicitly consider the selecting the type and number of cars assigned to an OD demand, given the particular number and type of the corresponding containers. Intermodal rail transportation thus implies a third consolidation operation, of containers on multi-platform cars, which makes the corresponding blocking problem different and much harder than for regular traffic.

There is a large variety of container types (e.g., 40- and 53-feet long) and railroads use fleets of cars of various types, each with one or several platforms and slots on the platforms. Single- and double-stack platforms have one and two slots, respectively. The matching/loading of containers to slots given the car type is an important but complex issue since 1) not all combinations are legal or suitable, and 2) a very large number of loading alternatives exist (Mantovani et al., 2017). How to reflect the differentiation of car and container types and the large number of loading alternatives in a way appropriate for tactical planning is a challenge that we address in this paper.

We conclude this section by emphasizing that in the blocking literature referred to above, block planning is performed first, the train make up and selection coming after. The problem we address is different. Trains and their schedules are selected first, in a separate planning phase, and are part of the input to the blocking problem. The model presented in the following section addresses this problem setting. We also analyse a number of problem variants reflecting practice, e.g., the existence of a list of preferred blocks for each train and the possibility to not move the complete demand by paying a penalty cost or to split the demand among several blocks.

3. Modelling

The railroad operates over an infrastructure network, where a number of terminals $\theta \in \Theta$ are dedicated, totally or partially, to the intermodal traffic. The plan is built for a *cyclic schedule* of given *length* (e.g., a week), assumed to be repeated over the tactical planning horizon. A set of train services $\Sigma = \{\sigma\}$ is given with their respective origin, destination and intermediary-stop terminals, and schedules. We propose a *block service network design model* based on a *three-layer cyclic space-time network* $\mathcal{G} = \{\mathcal{N}, \mathcal{A}\}$ (Zhu et al., 2014). Different from most service network design literature (Crainic and Kim, 2007), however, we do not discretize time. We rather use a continuous time representation where time moments correspond to the arrival/departure times services at terminals. Demand is defined in this context by its origin terminal, time at the origin terminal, destination terminal, due-date at destination, container type, number of containers of that type, and a penalty cost for late arrival. The container types in this version of the problem are the 40- and 53-feet container boxes; 20-feet containers are represented by transforming the corresponding demand into 40-feet containers; similarly, 45 and 48-feet containers are treated as 53-feet ones due to their loading mode. This hypothesis simplifies the presentation of the container-to-railcar assignment without impact on the generality of the formulation.

The network is schematically illustrated in Figure 1, for a terminal and time moments corresponding to two trains, one being at one of its intermediary stop, the other being at its origin. We first describe briefly each layer, nodes, and arcs, and then present the model formulation.

The known arrival and departure times of each scheduled train at the terminals on its route yield the corresponding arrival and departure nodes, TIN and TOUT in the *train layer*, defining the time structure of the entire network. We model train activities through *TrainHandling* arcs representing the time trains spend at terminals and *TrainMoving* arcs between terminals. The train features, e.g., maximum length, provide the capacity u_a of the moving arcs.

The *block (& car) layer* is illustrated below the train layer. Let $\mathcal{B} = \{b\}$ represent the set of *potential blocks* generated by the set of trains. A *block* is defined by its unique route and schedule from its origin to its destination, made up of movements on TrainMoving arcs and activities at terminals. We model building new blocks and assigning them to specific trains, transferring blocks from one train to another, and dismantling blocks at their destinations through the *Block2Train*, *BlockTransfer*, and *Block2Dismantle* intra-layer arcs, respectively. Each TIN node thus generates a block-drop-off node BD n the block layer. Similarly, each TOUT node generates a BOT node in the block layer, where the new blocks just formed and the blocks transferred from arriving trains, all sharing the same first/next train on their respective routes, are collected before being attached to the train. Each block is thus characterized by its: *route* and *schedule*, given by the set of trains moving it from its origin terminal o_b to its destination d_b, through (possibly) a sequence of intermediary terminals; *capacity* u_b in terms of total block length; and *fixed cost* f_b, representing the total cost of building it and transferring at intermediate terminals, as well as costs for the total transfer and idle (i.e., the train carrying it stops at a terminal for activities that do not involve it) times.

Demand enters and exists the system through the *container layer*. Demand $k \in \mathcal{K}$ is defined by its *origin terminal* o_k, *time at* the *origin* terminal α_k, *destination terminal* d_k, *due-date* at destination β_k, *container type* $\tau_k \in \mathcal{T}$, and *volume*, i.e., number of containers of that type, v_k.

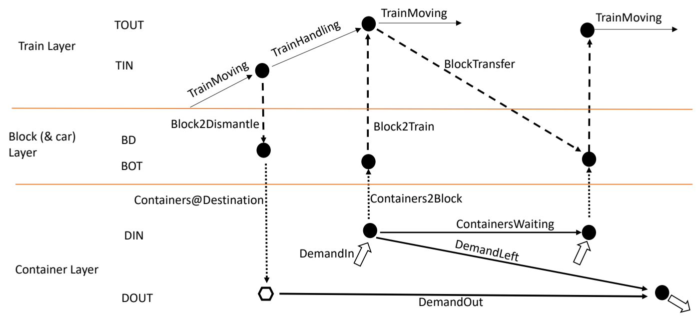

Fig. 1. Three-Layer Time-Space Network Representation

Let $\mathcal{B}_k$ be the set of *possible blocks* which may be used to transport demand k. A unit *transportation cost* c_{bk} is associated to each block $b \in \mathcal{B}_k$. The particular items included in c_{bk} vary according to the particular application. To better connect the selection of the blocks and the objectives of the railroad for a high-level of in-movement equipment (i.e., not idle), we include in the unit transportation costs c_{bk} measures related to the cost of hauling the cars, the train-to-train transfer time, the total number of transfers, and the idle time.

Upon arrival, containers wait on *ContainersWaiting* arcs until the selected block and departure time. The inter-layer *Container2Block* arcs support the container-to-car-to-block assignment model we propose, which is based on the relations among the lengths of blocks, platforms, and the two main container types (40 and 53 feet), and is described below. The *Containers@Destination* inter-layer arcs support the flow of containers at destination, unloaded from cars out of the dismantled blocks. This flow then moves on the *DemandOut* arcs, paying possibly late-arrival penalty costs p^{LATE}_{bk}, until a super-sink node.

Two variants of the problem setting are of interest in relation to demand. First, whether the containers of a particular demand may be separated and delivered by several itineraries, using different blocks (and trains), possibly at different times. When such *splitting* is allowed, one may still wish to limit the number of selected blocks, e.g., through penalty-like itinerary cost p^{SPLIT}_{k}. The second variant concerns the case when the total train capacity is less than the total demand. In such cases, an artificial arc, indicated as *DemandLeft* in Figure 1, is added between the arrival node of each demand and the sinkt node, with a penalty-like unit cost of p^{LEFT}_{k} for the flow that cannot be delivered by the given rail network.

Recall that, different from most other commodities hauled by railroads, intermodal demand must be loaded onto cars at the origin terminal and must be unloaded at destination. We now describe the approach we propose to integrate the container-to-car assignment into the tactical-planning model. Recall that *length* is a major constraining feature for trains, as well as for blocks both in terminals and when put on trains, and that it is the capacity measure used in this model. Furthermore, cars come in various multi-platform configurations, a platform providing two slots, one on the bottom and one on top, for containers of a given length, 40 and 53 feet in our case. The length of a car is then determined for the most part by its number of platforms and, consequently, so is the length of the block. We therefore use the *platform*, of a given type, as our loading unit making up the block.

Let $\Gamma = \{\gamma\}$ be the set of platform types. Consistent with the container types, two platform types are considered in this paper, the 40- and 53-feet long (λ^{p40} and λ^{p53} represent their respective lengths). The basic loading rules considering these container and platform types are:

- *40-feet platform*: 1) single 40-feet container in bottom slot; 2) 2 40-feet containers in the two slots; 3) 1 40-feet container in bottom slot and a 53-feet one in the top slot;
- *53-feet platform*: 1) single 53-feet container in bottom slot; 2) 2 53-feet containers in the two slots.

Because the overall goal of the loading problem is to maximize the number of forty-foot equivalent units per unit of train length, it is always best to use as many 40-feet platforms as possible. Under this hypothesis, when nb_{53}, the number of 53-feet containers, is greater than or equal to nb_{40}, the number 40-feet containers, then the 40-feet containers should be placed in bottom slots and the 53-feet ones on top, as much as possible. The number of 53-feet platforms, nbp_{53}, is hence given by Equation (1), which easily yields the number of 40-feet platforms, nbp_{40}, through Equation (2)

$$nbp_{53} = \max\{0, \lceil (nb_{53} - nb_{40})/2 \rceil\}, \tag{1}$$

$$nbp_{40} = \lceil (nb_{53} + nb_{40})/2) \rceil - nbp_{53} \tag{2}$$

We propose a mixed integer linear programming (MILP) formulation with decision variables:

- $y_b = 1$, if block $b \in \mathcal{B}$ is selected, 0 otherwise;
- $y_{bk} = 1$, if block $b \in \mathcal{B}_k,\ k \in \mathcal{K}$, is selected, 0 otherwise;
- x_{ak}, volume of demand $k \in \mathcal{K}$ on arc $a \in \mathcal{A}$, with x_{bk}, the volume of demand k on block $b \in \mathcal{B}_k$;
- x_k, volume of demand k left at the origin terminal, i.e., the volume on the corresponding artificial arc;
- x_b^{c40} and x_b^{c53}: number of 40- and 53-feet containers, respectively, assigned to block $b \in \mathcal{B}$;
- $x_b^{\pi 40}$ and $x_b^{\pi 53}$ number of 40- and 53-feet foot platforms in block $b \in \mathcal{B}$, respectively.

The integer programming formulation for the general problem setting, where the total train capacity might not be sufficient and the split of demand is permitted, may then be written as

$$\text{Minimize} \sum_{b\in\mathcal{B}} f_b y_b + \sum_{k\in\mathcal{K}} \left(\sum_{b\in\mathcal{B}_k} (c_{rk} + p_{rk}^{\text{LATE}}) x_{bk} + \sum_{b\in\mathcal{B}_k} p_k^{\text{SPLIT}} (y_{bk} - 1) \right) + p_k^{\text{LEFT}} x_k \tag{3}$$

subject to the following constraints

$$\sum_{b\in\mathcal{B}_k} x_{bk} + x_{ak} = v_k,\ k \in \mathcal{K}, \tag{4}$$

$$\sum_{b\in\mathcal{B}_k} y_{rk} \geq 1,\ k \in \mathcal{K}, \tag{5}$$

$$x_{bk} \leq y_{rk} u_{b_r},\ b \in \mathcal{B}_k,\ k \in \mathcal{K}, \tag{6}$$

$$x_{a_i \in \mathcal{A}_\theta^{\text{DIN}} k} + x_{a_i \in \mathcal{A}_\theta^{\text{DWT}} k} + x_{ak} = v_k,\ \ \theta = o_k,\ i : t(i) = \alpha_k,\ k \in \mathcal{K}, \tag{7}$$

$$x_{a_i \in \mathcal{A}_\theta^{\text{DIN}} k} = \sum_{b\in\mathcal{B}_{ki}} x_{bk},\ \theta = o_k,\ i \in \mathcal{N}_\theta^{\text{DIN}},\ i : t(i) = \alpha_k,\ k \in \mathcal{K}, \tag{8}$$

$$x_{a_{i-1} \in \mathcal{A}_\theta^{\text{DWT}} k} + x_{a_i \in \mathcal{A}_\theta^{\text{DIN}} k} = x_{a_i \in \mathcal{A}_\theta^{\text{DWT}} k},\ \ \theta = o_k,\ \forall i : t(i) > \alpha_k,\ k \in \mathcal{K}, \tag{9}$$

$$x_{a_i \in \mathcal{A}_\theta^{\text{DIN}} k} = \sum_{b\in\mathcal{B}_{ki}} x_{bk},\ \theta = o_k,\ i \in \mathcal{N}_\theta^{\text{DIN}},\ \forall i : t(i) > \alpha_k, ,\ k \in \mathcal{K}, \tag{10}$$

$$x_{a_i \in \mathcal{A}^{\text{D2D}} k} = \sum_{b\in\mathcal{B}_k : \beta_b = t(i)} x_{bk},\ \ \theta = d_k,\ i \in \mathcal{N}_\theta^{\text{CD}},\ k \in \mathcal{K} \tag{11}$$

$$\sum_{i\in\mathcal{N}_\theta^{\text{CD}}} x_{a_i \in \mathcal{A}^{\text{D2D}} k} + x_k = v_k,\ \ \theta = d_k,\ k \in \mathcal{K}, \tag{12}$$

$$x_b^{\pi 53} = \max \left[0, \left\lceil \frac{1}{2} \left(\sum_{k:\tau_k = p53} x_{bk} - \sum_{k:\tau_k = p40} x_{bk} \right) \right\rceil \right],\ b \in \mathcal{B}, \tag{13}$$

$$x_b^{\pi 40} = \left\lceil \frac{1}{2} \left(\sum_k x_{bk} \right) \right\rceil - x_b^{\pi 53},\ b \in \mathcal{B}, \tag{14}$$

$$\lambda_b = \lambda^{p40} \sum_{k:b\in\mathcal{B}_k} x_b^{\pi 40} + \lambda^{p53} \sum_{k:b\in\mathcal{B}_k} x_b^{\pi 53},\ b \in \mathcal{B}, \tag{15}$$

$$\lambda_b \leq u_b,\ b \in \mathcal{B}, \tag{16}$$

$$\sum_{b\in\mathcal{B}_a} \lambda_b \leq u_a\ \forall a \in \mathcal{A}_\sigma^{\text{TM}},\ \sigma \in \Sigma, \tag{17}$$

$$y_b, y_{rk} \in \{0,1\}, x_{bk}, x_{ak}, x_k \leq 0,\ k \in K,\ r \in \mathcal{R}_k,\ b \in \mathcal{B},\ a \in \mathcal{A}. \tag{18}$$

The objective function (3) minimizes the total cost composed of the block selection and operation costs, plus the costs associated to handling demand: costs for moving and waiting, penalty for late delivery, splitting costs, and penalty for not delivering the full volume. Recall that the loading procedure automatically enforces the maximum utilization of the most desired platforms.

Constraints (4) ensure that all the volume of each demand is shipped out either on itineraries or on the artificial arc, while Constraints (5) specify that at least one itinerary must be selected for each demand. The latter constraints enforce equality to 1 when demand splitting is not allowed. The linking constraints (6) enforce the rule that flow is shipped out only on selected itineraries and that it is not larger than the capacity of the corresponding block.

Flow conservation at the origin of the demand is taken care of by Constraints (7) - (8), the latter link the number of containers to be sent at that time to the blocks they are assigned to. The next two, (9) - (10), perform the same role at the subsequent DIN time moments. Constraints (11) - (12) complete the flow-conservation task at the destination of demand. Constraints (11) link the flow of commodity k on the exiting arc to volumes carried by the blocks that brought it to the destination terminal. Constraints (12) then make sure all containers are delivered at destination.

Constraints (13) - (14) compute (Equations (1) - (2)) for each potential block the number of loaded platforms of each type corresponding to the commodity traffic assigned to them. Constraints (15) - (16) compute the block length and enforce the block capacity constraint, while Constraints (17) enforce the train capacity on each of the respective moving links. Constraints (18) define the feasible domain of the formulation.

4. Experimental Results & Analyses

Our objectives for the experimental campaign were to 1) evaluate the efficiency of addressing the blocking model for realistic settings using off-the-shelf commercial software; 2) study the impact of a number of important problem characteristics on the blocking plan generated by the model, and 3) explore the potential of the methodology as what-if analysis tool by addressing a number of problem variants.

The model was implemented as a mixed integer linear program and was addressed using Java and CPLEX 12.7.0 (single thread). Experiments were performed on a computer Intel(R) Xeon(R) CPU E5-2609 v2 @ 2.50GHz with 8 CPUs and 128 GB de RAM.

An extensive set of instances were generated based on data of a large North American railroad, reflecting realistic problem characteristics and actual practice. The instances include 192 terminals, 519 trains, 5264 demands (OD commodities), 20-, 40- and 53-feet containers, and 40- and 53-feet platforms. 20-feet containers were transformed into 40-feet ones, while 53-feet containers included 43- and 45-feet ones as the same loading rules apply to the three types. The schedule length covers seven days (10080 minutes). We varied the dimensions of the set of potential blocks: a **Complete** (as identified in the following tables) set of 16654 blocks, corresponding to all feasible possibilities (a single restriction only: not more than 24-hour transfer delays); a **Constrained** set of 1929 blocks, where each train has a list of preferred blocks; two sets of intermediate dimensions, **Inter1** and **Inter2**, with 3906 and 7023 blocks, respectively, obtained by restricting the permitted number of transfers or the maximum transfer delay. We preformed experiments for the **No-split** demand case, as well as for demand **Split** without any penalty, with **Low penalty** and **High penalty** costs.

Table 1 displays the results obtained in terms of optimality gaps (in %) after 3 hours and, in between parentheses, 24 hours of CPU time, for the four block sets and the demand-handling scenarios. Not surprisingly, the larger the number lof potential blocks, the more difficult it gets to solve the scheduled block service network design problem, pointing to future developments of tailor-made solutions, including dynamic block generation and matheuristics, to address even larger instances. Yet, remarkably good solutions are obtained relatively fast, particularly when practice-oriented problem settings are considered. This is also an indication that, in actual operations, it is not worth letting the software to run for very long as proving convergence, which may take a very long time, is not necessarily what is sought after in practice. The next experimental results are displayed for the cases of 3 hours CPU time only.

With respect to the handling of demand, one notices that allowing splitting the demand provides significant efficiency increases in all cases. The more realistic settings are solved to optimality, and dramatic efficiency gains are observed for the largest block set considered. Splitting is beneficial even when low to moderate penalties have to be paid, which is an interesting managerial insight for future negotiations with customers. Very high penalties greatly deteriorate efficiency; it is better in this case, to use the model with the no-split option.

Table 1. Efficiency - Optimality Gaps

Block set	No split	Split - No penalty	Split - Low penalty	Split - High penalty
Complete (16.6k)	16 (15)	3 (3)	7 (6)	42 (34)
Constrained (2k)	1 (1)	0 (0)	0 (0)	3 (3)
Inter1 (4k)	2 (1)	0 (0)	0 (0)	3 (3)
Inter2 (7k)	10 (9)	2 (2)	5 (4)	30 (21)

Table 2. Block Service Network Structure

Block set	No split	Split - No penalty	Split - Low penalty	Split - High penalty
Complete (16.6k)	905 (1.6)	939 (0)	948 (0.3)	932 (2.9)
Constrained (2k)	721 (12.3)	733 (11.5)	728 (11.5)	731 (12.6)
Inter1 (4k)	727 (12)	744(11)	736 (0)	742 (12.2)
Inter2 (7k)	848 (5.1)	881 (4)	862 (4)	877 (6.3)

Table 3. Costs Performance

Block set	No split	Split - No penalty	Split - Low penalty	Split - High penalty
Complete (16.6k)	4.5	3.9	4.0	5.7
Constrained (2k)	8.9	8.7	8.7	9.0
Inter1 (4k)	8.7	8.5	8.5	8.8
Inter2 (7k)	5.8	5.3	5.5	7.5

Table 2 displays the results in terms of block service network structure and performance with respect to demand fulfilment. The figures indicate for each setting (the same as previously), the number of selected blocks and, in parentheses, the percentage of undelivered demand. One observes that, as expected, starting with a large set of potential blocks yields better results in terms of system performance compared to more restrained sets. A larger, and thus more diversified, initial set means a larger selection of blocks with similar cost values, providing the opportunity to better use the available train capacity and, thus, to more largely spread out the demand among these alternative paths and deliver more. Indeed, the percentage of un-delivered demand is significantly lower for the largest set of potential blocks than for the others. As a managerial insight, this result indicates that even when sets of preferred blocks are attached to trains, these sets should be larger, even though the computational effort will also be larger. The results also show the interest of the proposed model as an analysis tool to evaluate train-selection strategies. In all cases, though, the selected block service network is rather stable, the variations in the numbers of blocks selected being significantly smaller than those within the numbers of potential blocks.

The results with respect to the handling of demand are similar to those of the first set of results. Splitting increases the performance in terms of servicing the demand, as larger blocks service networks and lower percentages of undelivered demand are observed. High penalties are to be avoided, low-medium-valued ones providing control, i.e., good results, particularly when the number of potential blocks is large, with little performance degradation.

Finally, Table 3 displays results relative to the cost of the block service network provided by the model. Costs are in millions of Canadian dollars, and were obtained after 3 hours of CPU time. The respective figures after 24 hours of CPU time, not displayed, are a little lower, stating again that proving convergence is a very slow process. One observes significant cost reductions when the model may choose among a large number of potential blocks and even larger when demand splitting is permitted. The cost-related results confirm all the previous conclusions relative to the split/no-split and low/high penalty costs policies.

5. Conclusions

We addressed the tactical block planning problem of intermodal rail and proposed a new continuous-time, three-layer service network design model that explicitly addresses the challenging characteristics of intermodal transportation. We performed an extensive experimentation studying the case of a major North American railroad. The results showed the interest of the methodology as one can address realistic instances with commercial software, obtaining good solutions within acceptable computing times. The analysis also provided several managerial insights, e.g., the possibility to split the demand significantly increases the efficiency of the procedure and the quality of the solution with respect to total cost and undelivered demand; high penalties for splitting demand greatly deteriorate efficiency and solution quality, while low-medium ones provides control with little performance degradation; the larger the number of possible blocks, the more blocks are selected and, then, the better trains are used and demand is serviced for a better economic performance of the system.

Acknowledgements

We gratefully acknowledge the funding provided by the CN Chair in Optimization of Railway Operations, Université de Montréal, by the Natural Sciences and Engineering Council of Canada (NSERC), through its Cooperative Research, Discovery, and Acceleration Grants programs, and by the Ministero dell'Istruzione, dell'Università e della Ricerca (MIUR) of Italy, through its Research Projects of Relevant National Interest (PRIN) program. We also gratefully acknowledge the support of Fonds de recherche du Québec through their strategic infrastructure grants. While working on the project, the second author was Adjunct Professor, Department of Computer Science and Operations Research, Université de Montréal.

References

Ahuja, R.K., Cunha, C.B., Şahin, G., 2005. Network Models in Railroad Planning and Scheduling, in: Tutorial in Operations Research INFORMS 2005. INFORMS, pp. 54–101. Published online: 14 Oct 2014; http://dx.doi.org/10.1287/educ.1053.0013.

Ahuja, R.K., Jha, K.C., Liu, J., 2007. Solving Real-Life Railroad Blocking Problems. Interfaces 37, 404–419.

Barnhart, C., Jin, H., Vance, P.H., 2000. Railroad Blocking: A Network Design Application. Operations Research 48, 603–614.

Bodin, LD., Golden, B.L., Schuster, A.D., Romig, W., 1980. A Model for the Blocking of Trains. Transportation Research Part B: Methodological 14, 115–120.

Crainic, T.G., Ferland, J.A., Rousseau, J.M., 1984. A Tactical Planning Model for Rail Freight Transportation. Transportation Science 18, 165–184.

Crainic, T.G., Kim, K.H., 2007. Intermodal Transportation, in: Barnhart, C., Laporte, G. (Eds.), Transportation. North-Holland, Amsterdam. volume 14 of *Handbooks in Operations Research and Management Science*. chapter 8, pp. 467–537.

Crainic, T.G., Rousseau, J.M., 1986. Multicommodity, Multimode Freight Transportation: A General Modeling and Algorithmic Framework for the Service Network Design Problem. Transportation Research Part B: Methodological 20, 225–242.

Jha, K.C., Ahuja, R.K., Şahin, G., 2008. New Approaches for Solving the Block-to-Train Assignment Problem. Networks 51, 480–62.

Mantovani, S., Morganti, G., Umang, N., Crainic, T.G., Frejinger, E., Larsen, E., 2017. The Load Planning Problem for Double-Stack Intermodal Trains. European Journal of Operational Research 267, 107–119.

Newton, H.N., Barnhart, C., Vance, P.H., 1998. Constructing Railroad Blocking Plans to Minimize Handling Costs. Transportation Science 32, 330–345.

Zhu, E., Crainic, T.G., Gendreau, M., 2014. Scheduled Service Network Design for Freight Rail Transportations. Operations Research 62, 383–400.