



# Robust model and solution algorithm for the railroad blocking problem under uncertainty

R. M. Hasany and Y. Shafahi\*

*Department of Civil Engineering, Sharif University of Technology, Tehran, P.O. Box 11155-8639, Iran.*

Received 30 December 2015; received in revised form 21 December 2016; accepted 4 March 2017

## KEYWORDS

Railroad blocking problem;  
Robust optimization;  
Branch-and-cut algorithm;  
Uncertainty;  
Railway planning.

**Abstract.** The railroad blocking problem emerges as an important issue at the tactical level of planning in freight rail transportation. This problem consists of determining the optimal paths for freight cars in a rail network. Often, demand and supply resource indicators are assumed certain; hence, the solution obtained from a certain model might not be optimal or even feasible in practice due to the stochastic nature of these parameters. To address this issue, this paper develops a robust model for this problem with uncertain demand and travel time as supply resource indicators. Since the model combines integer variables and nonlinear functions, a branch-and-cut algorithm is used to solve the linearized version of the robust model. The performance of the proposed algorithm in several instances is examined and discussed. The high efficiency and effectiveness of the proposed algorithm are demonstrated through a comparison with a well-known solver. Finally, this algorithm is applied to a blocking problem of the railways of Iran. The results show that, by ignoring approximately 10% of the optimal value of the deterministic model, we have an optimal solution that remains unchanged with a probability of more than 0.98.

© 2018 Sharif University of Technology. All rights reserved.

## 1. Introduction

### 1.1. Railroad blocking problem

The planning problems in railway transportation are classified into strategic, tactical, and operational levels [1,2]. The strategic level is concerned with long-term planning that requires a large amount of investment. High-level managers who design and execute strategic decisions to paint a picture of the long-term goals of the railway system make the decisions at this level. The tactical level supports the strategic level by translating strategic decisions into specific decisions relevant to a distinct area of railway planning. Indeed,

the tactical level is the connection between the strategic and operational levels. The operational level lies at the bottom of railway planning. The decisions at this level are focused on specific procedures and are made by frontline managers.

Finding the optimal paths for freight cars, i.e., the so-called blocking problem, is an important issue at the tactical planning level [3]. There are three approaches to sending each shipment (i.e., the number of cars with the same origin-destination pair) through a rail network. In the first approach, the cars that are associated with a shipment are assigned to the path with the least travel time, which sends them by a direct train that does not stop at any classification yard (which is the station with the ability to separate, sort, and assemble trains) until reaching its destination. If the number of cars associated with a shipment is approximately equal to the train's hauling capacity, the first approach can be reasonable; however, the number of cars associated with a shipment is usually less than

\*. Corresponding author. Tel.: +98 21 66164246;  
Fax: +98 21 66014828  
E-mail addresses: [rmhasany@gmail.com](mailto:rmhasany@gmail.com) (R. M. Hasany);  
[Shafahi@sharif.edu](mailto:Shafahi@sharif.edu) (Y. Shafahi)

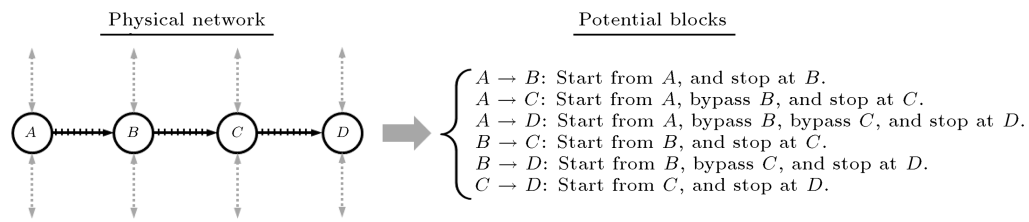


Figure 1. Potential blocks in a small example.

the hauling capacity of the outgoing train [4], and using this approach will lead to dispatching trains at their under-hauling capacities. The second approach focuses on dispatching trains at their near-hauling capacities by assembling the incoming cars in a classification yard as long as the number of cars does not exceed the hauling capacity of the train. Because the cars are not grouped at their origins based on their destinations, the cars can stop at several classification yards on the way to their destinations, causing delay in shipments. For instance, Barnhart et al. [5] reported that stopping at each classification yard would lead to a one-day delay; further, Bontekoning and Priemus [6] reported that classification operations could take 10 to 50% of train travel time. The third approach adopts the optimization techniques allowing a train to enter into some, but not all, classification yards to improve the performance of the rail system. In particular, this approach involves identifying the list of visited classification yards for each shipment to minimize the objective function (such as moving time and delay time experienced inside the yards) while satisfying railway restrictions. Note that the third approach can be stated as the general form, and the two earlier approaches are only special, extreme cases of the general form. The railroad blocking problem focuses on solving the general form to obtain the optimal path for each shipment.

Four important terms with respect to the railroad blocking problem should be defined. First, a *block*, consisting of a sequence of physical links, is a key element in the railroad blocking problem. If the cars are traveling across a block, they are not classified as long as they reach the end of that block. Every block is associated with two nodes: the start node, where the cars are attached to the departing train; and the end node, where the cars are released from the train. Second, the *blocking network* consists of a set of blocks and their corresponding nodes that are origin, destination, or classification yards. Third, the *blocking path* of a shipment is defined as a sequence of blocks that connects the origin to the destination for that shipment. Fourth, a feasible solution to the blocking problem is called the *blocking plan*, consisting of a set of blocking paths for all shipments.

To clarify the definitions, consider a small rail (physical) network with four nodes connected by three

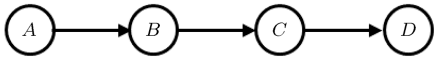
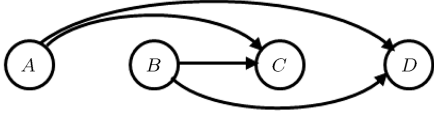


arcs, which can be a part of a large network. Let nodes A and B be the origins and nodes C and D be the destinations. The four shipments are shipment 1 (from A to C), shipment 2 (from A to D), shipment 3 (from B to C), and shipment 4 (from B to D). Nodes B and C operate as classification yards in this network. The network and the list of potential blocks are shown in Figure 1. As an example, consider block  $A \rightarrow D$ . Through this block, a train starts moving from node A, bypasses nodes B and C, and finally stops at node D. The potential blocks create four blocking plans, as shown in Table 1. Consider blocking plan 4 with three blocks. Shipments 1 and 2 move through block  $A \rightarrow C$ . At node C, shipment 1 reaches its destination, while shipment 2 is released from the train at node C. Shipments 3 and 4 move through block  $B \rightarrow C$ , and shipment 3 reaches its destination at node C, while shipment 4 is released from the train at node C. Finally, shipments 2 and 4 move together through block  $C \rightarrow D$  to reach their destination (node D).

### 1.2. Literature review of the railroad blocking problem

Several studies have examined the railroad blocking problem. Assad [3] initially introduced an integer linear program model for this problem. The objective functions of the model included the train costs (operating and delay costs through the rail tracks) and yard costs (classification and delay costs inside yards). The model included two constraints: each shipment is shipped from its origin to its destination, and the amount of flow moving through each block is bounded by block capacity. In addition to the Assad's objective function, other functions have been used frequently in the literature. Bodin et al. [7] derived a delay function in terms of car flow, which is added to train costs. Marin and Salmerón [8] considered investment to purchase new fleets, in addition to the train costs. The number of visited classification yards was introduced by Fügenschuh et al. [4] as a crucial component of objective function.

Many realistic limitations were not presented in Assad's work. Efforts to consider more limitations have been undertaken, including a maximum limit on the number of cars that can be classified inside the yards [9], a maximum limit on the number of cars that can move through each block [10], a maximum limit

**Table 1.** An illustration of blocking plans.

No.	Blocking plan	Blocks	Shipments	Blocking paths
1		$A \rightarrow B$ $B \rightarrow C$ $C \rightarrow D$	1	$A \rightarrow B \rightarrow C$
			2	$A \rightarrow B \rightarrow C \rightarrow D$
			3	$B \rightarrow C$
			4	$B \rightarrow C \rightarrow D$
2		$A \rightarrow C$ $A \rightarrow D$ $B \rightarrow C$ $B \rightarrow D$	1	$A \rightarrow C$
			2	$A \rightarrow D$
			3	$B \rightarrow C$
			4	$B \rightarrow D$
3		$A \rightarrow B$ $B \rightarrow C$ $B \rightarrow D$	1	$A \rightarrow B \rightarrow C$
			2	$A \rightarrow B \rightarrow D$
			3	$B \rightarrow C$
			4	$B \rightarrow D$
4		$A \rightarrow C$ $B \rightarrow C$ $C \rightarrow D$	1	$A \rightarrow C$
			2	$A \rightarrow C \rightarrow D$
			3	$B \rightarrow C$
			4	$B \rightarrow C \rightarrow D$

on the sum of the blocks that leave the yards [11], a maximum limit on the time spent reaching the destination for each origin-destination pair [12], and the number of blocking paths with a positive flow for each shipment could be more than one path [13].

Solution algorithms for the railroad blocking problem can be categorized in two ways. One way is to design a metaheuristic that possibly achieves a good solution to the problem with a short running time. Researchers have made great efforts to design metaheuristics for the railroad blocking problem. Some examples include neural networks [14], genetic algorithms [15], simulated annealing [9], ant colony optimization [16], taboo searches [17], and local searches for three major United States' railroads [18]. Metaheuristics iteratively search among the feasible regions to find the near-optimal solution with less computation time, compared to the exact algorithms for integer programming. However, metaheuristics might not converge to the optimal solution, and this flaw points to exact algorithms whose convergence to optimality is assured, which is categorized as the second way.

Bodin et al. started using exact algorithms for finding the optimal solution to the railroad blocking problem [7], and implemented a branch-and-bound algorithm using data from Norfolk and Western Railroad. The branch-and-price algorithm was applied to the railroad blocking problem for a major domestic railroad in the United States [10], the Deutsche Bahn railroad [19], and some hypothetical networks [20]. Keaton [11] initially adopted the Lagrangian relaxation

method to decompose the original problem into two independent sub-problems for a portion of the Conrail system in the United States. Barnhart et al. [5] also adopted the Lagrangian relaxation method and added inequalities to the relaxed Lagrangian problem to lift the lower bounds. They used this algorithm for realistic network data provided by CSX transportation. Fügenschuh et al. [4] presented a mixed-integer, non-linear program and described two exact techniques and one heuristic to improve the proposed formulation.

There has been much attention paid to the impact of uncertainty on the results of railway planning [21–24]. Nevertheless, the blocking problem under uncertainty has been ignored. The study of Jin [25] is currently the only one available in the literature addressing this problem. In this study, the author introduced uncertainty in the nature of a demand and supply resource indicator in a limited number of scenarios to solve the uncertain model. Indeed, if the assumed values of random parameters are not equal to their values after realization, then the optimal solution might not be optimal or even feasible anymore [26].

### 1.3. Uncertainty in the railroad blocking problem

At the tactical level of railroad planning, the quantity of input parameters is often assumed certain and known. While the time horizon of tactical planning varies from one to three months, disregarding uncertainty leads to a clear difference between the optimal solutions of the deterministic and stochastic

models [27]. In the deterministic model, the nominal quantity of uncertain parameters can be calculated using different methods [28–30]; however, it has been reported that the nominal quantity of demand is over- or under-estimated [31], and the whole system could face an enormous amount of unexpected costs [32]. Additionally, the travel time spent through the rail network is severely affected by such factors as changing weather, fleet availability, passing or overtaking, and moving of trains with higher priority (such as passenger trains) on the same track [6]. Thus, considering the demand and travel time as random parameters is much closer to reality.

*Sensitivity analysis* is a common tool used to address uncertainty in mathematical modeling. This method analyzes the optimal solution by changing the parameter values in a predefined range. More precisely, it is necessary initially to solve the model with nominal data; then, by changing only one parameter within its range, it is necessary to ascertain whether the optimal solution remains optimal or not. Sensitivity analysis is assumed as the post-analysis method which only enables us to verify possible changes to the optimal solution as an input [33].

*Stochastic programming* is another tool that provides a solution and takes advantage of probability distributions governing the data being known or able to be estimated. This approach has two main issues. First, the formulation of stochastic programming assumes that each uncertain parameter can be modeled as a random value with a valid probability distribution; therefore, this approach cannot be effective when the history data for the random parameter are not sufficient. The second issue relates to the difficulty of solving stochastic programming models, particularly in the presence of integer variables and nonlinear functions in the objective function or constraints.

*Robust optimization* is another approach to addressing uncertainty in operations research, which is the focus of this paper. Several researchers have emphasized the subject of robust optimization. Soyster [34] modeled the robust counterpart of a linear program. The proposed model was very conservative, indicating that the difference between the optimum values for a deterministic model and its robust counterpart is considerably high. In light of this issue, Ben-Tal and Nemirovski [35] proposed a robust model for a linear program with uncertainty, which is inherent to technological coefficients. The authors showed that the new robust model is less conservative than the model proposed by Soyster; in addition, the proposed model allowed for the control of the degree of conservatism.

In general, the robust approach has three advantages. Most importantly, the robust approach guarantees the protection of the obtained solution against constraint violations due to the uncertainty parameter.

Since the blocking parameters usually change from month to month, the deterministic blocking plan must be subsequently re-optimized to obtain a new blocking plan for new realizations of uncertain parameters. If the new blocking plan is different from the previous blocking plan, then the train schedule required to carry a new set of blocks could be different from the previous schedule, and the previous locomotive and crew assignments might no longer be valid [36]. Second, the application of the robust approach will result in a model that requires less effort to solve, compared to the stochastic model. Third, the robust approach only requires a range of the uncertain parameters, rather than having a probability function for each uncertain parameter. It was reported that, for a stochastic parameter, the range is estimated much easier than probability function [37].

Our contributions made to the literature are mainly in two aspects: modeling and offering a solution algorithm. The first contribution of this paper is that it proposes a robust blocking problem with uncertain demands and travel time as the supply resource indicators to produce a solution that is immunized against any realization of uncertain parameters in the ranges. The resulting model combines the difficulty of discrete variables with the challenge of nonlinear functions in constraints. It is necessary to solve this model for realistic networks; however, the current state-of-the-art software is not able to find the optimal solution or even produce a feasible solution in a reasonable amount of time. Therefore, the second contribution is devoted to the development of a specialized branch-and-cut algorithm. This algorithm involves running a branch-and-bound algorithm and using the process to generate a cut that is obtained by the linear approximation of the nonlinear terms in constraints.

The rest of the paper is organized as follows. Section 2 introduces the deterministic form of the blocking problem. Section 3 presents different sources of uncertainty in the blocking problem and the robust formulation of the deterministic model. Section 4 describes a branch-and-cut method for solving the robust formulation of blocking problem. Section 5 shows numerical examples and a case study. Finally, Section 6 presents the main conclusions of the presented work.

## 2. Deterministic model for the blocking problem

To explain uncertain parameters to the railroad blocking problem, the presentation of the deterministic model is required. The model proposed by Newton et al. [10], which is one of the most commonly used models in the railroad blocking problem with certain parameters, is used in this section. The model is as follows:

**Input parameters:**

$A$	The set of blocks indexed by $a$ ,
$K$	The set of shipments indexed by $k$ ,
$S$	The set of stations indexed by $i$ ,
$Q^k$	The set of candidate blocking paths for shipment $k \in K$ indexed by $q$ ,
$wa_i$	Time for separating and sorting each car inside station $i \in S$ (hr),
$t^a$	Traveling time for block $a \in A$ (hr),
$\delta_q^a$	1 if block $a \in A$ is on blocking path $q \in Q^k$ and 0 otherwise,
$\xi_i^a$	1 if station $i \in S$ is the start station of block $a \in A$ and 0 otherwise,
$d^k$	Demand value for shipment $k \in K$ during the planning horizon (car),
$\tau_q^k$	Travel time of path $q \in Q^k$ for shipment $k \in K$ , that is, $\sum_{a \in A} (t^a + \sum_{i \in S} \xi_i^a wa_i) \delta_q^a$ (hr),
$V_i$	The outflow car capacity of station $i \in S$ (car),
$W_i$	The maximum number of blocks directed out of station $i \in S$ ,
$u^a$	The upper limit on car flows in block $a \in A$ (car).

**Decision variables:**

$z^a$	1 if candidate block $a \in A$ is selected and 0 otherwise,
$v_q^k$	Car flows associated with blocking path $q \in Q^k$ for $k \in K$ (car).

Barnhart et al. [5] presented the blocking problem as follows:

$$(P) \quad \min_{v_q^k, z^a} \sum_{k \in K} \sum_{q \in Q^k} \tau_q^k v_q^k, \quad (1)$$

s.t.

$$\sum_{k \in K} \sum_{q \in Q^k} \delta_q^a v_q^k \leq u^a z^a, \quad \forall a \in A, \quad (2)$$

$$\sum_{q \in Q^k} v_q^k \geq d^k, \quad \forall k \in K, \quad (3)$$

$$\sum_{a \in A} \xi_i^a z^a \leq W_i, \quad \forall i \in S, \quad (4)$$

$$\sum_{a \in A} \sum_{k \in K} \sum_{q \in Q^k} \delta_q^a \xi_i^a v_q^k \leq V_i, \quad \forall i \in S, \quad (5)$$

$$v_q^k \geq 0, \quad z^a \in \{0, 1\}, \quad \forall k \in K; \quad \forall q \in Q^k; \quad \forall a \in A. \quad (6)$$

The objective function (1) minimizes the total travel time for shipments flowing across the rail network. Constraint (2) limits the sum of flows between many origins to many destinations through block  $a \in A$  to  $u^a$ . Additionally, this constraint ensures that if  $z^a = 0$ , then the blocking path  $q \in Q^k$  containing block  $a \in A$ , i.e.,  $\delta_q^a = 1$ , is not involved in the solution. Constraint (3) conserves the car flows of each shipment. Constraint (4) enforces the number of blocks directed out of any station to limit its block capacity. Similarly, Constraint (5) enforces the car flow directed out of any station to limit its volume capacity.

Above, Barnhart et al. [5] assumed that demand value and travel time are certain. In the rest of the paper, we present the robust formulation of the railroad blocking problem when these two parameters are uncertain.

**3. Robust formulation of the railroad blocking problem**

The aim of this section is to determine the optimal blocking paths for each shipment in the presence of uncertainty in demand value and travel time. Suppose that demand value,  $\tilde{d}^k$ , and travel time,  $\tilde{\tau}_q^k$ , are uncertain and belong to two bounded sets,  $U_{d^k}$ , and  $U_{\tau_q^k}$ . The definition of these two sets can be written as follows:

$$U_{d^k} = \left\{ d^k + \theta \hat{d}^k \mid -1 \leq \theta \leq 1 \right\}, \quad (7)$$

$$U_{\tau_q^k} = \left\{ \tau_q^k + \mu \hat{\tau}_q^k \mid -1 \leq \mu \leq 1 \right\}. \quad (8)$$

Let  $d^k$  and  $\tau_q^k$  be the nominal values of demand and travel time, respectively, and  $\hat{d}^k$  and  $\hat{\tau}_q^k$  be the maximum variations of demand value and travel time. The uncertain model for the blocking problem is given by the following:

$$(RP1) \quad \min_{v_q^k, z^a} \sum_{k \in K} \sum_{q \in Q^k} \tilde{\tau}_q^k v_q^k, \quad (9)$$

s.t.

$$\sum_{q \in Q^k} v_q^k \geq \tilde{d}^k \quad \forall k \in K, \quad (10)$$

Eqs. (2) and (4)-(6).

To develop the robust formulation of (RP1), the methodology initially proposed by Ben-Tal and Nemirovski [26] is adopted. Note that these researchers only addressed uncertainty in the standard form, in which the technological coefficients are the only source of uncertainty; however, the uncertain parameters in the blocking problem (i.e.,  $\tilde{d}^k$  and  $\tilde{\tau}_q^k$ ) appear in both the objective function (9) and right-hand side of

Constraint (10). To rewrite (RP1) in the standard format, two modifications are required [38]. The first modification adds the objective function to the constraint set. To do so, an unrestricted decision variable,  $\Psi$ , which is equal to and greater than the objective function (9) is introduced; then, this constraint is added to the constraint set. The second modification changes the role of the demand value in the technological coefficient in Constraint (10). To do so, a variable,  $p^k$ , is introduced for each  $k \in K$ , and  $\tilde{d}^k$  appears as a coefficient of  $p^k$  provided that  $p^k = 1$  always to obtain  $\tilde{d}^k p^k = \tilde{d}^k$ . Using these two modifications, model (RP1) can be rewritten in the standard format as follows:

$$(RP2) \quad \min_{v_q^k, z^a, \Psi, p^k} \Psi, \quad (11)$$

s.t.

$$-\Psi + \sum_{k \in K} \sum_{q \in Q^k} \tilde{\tau}_q^k v_q^k \leq 0, \quad (12)$$

$$-\sum_{q \in Q^k} v_q^k + \tilde{d}^k p^k \leq 0 \quad \forall k \in K, \quad (13)$$

$$\Psi \text{ is free}, \quad p^k = 1 \quad \forall k \in K, \quad (14)$$

Eqs. (2) and (4)-(6).

Based on the work of Ben-Tal and Nemirovski [26], the robust formulation of model (RP2) is derived as follows. For further details on the conversion of (RP2) to (RP3) see the appendix.

$$(RP3) \quad \min_{v_q^k, y_q^k, x_q^k, p^k, s^k, r^k, \Psi, z^a} \Psi, \quad (15)$$

s.t.

$$-\Psi + \sum_{k \in K} \sum_{q \in Q^k} \tau_q^k v_q^k + \sum_{k \in K} \sum_{q \in Q^k} \hat{\tau}_q^k y_q^k + \Phi \sqrt{\sum_{k \in K} \sum_{q \in Q^k} (\hat{\tau}_q^k x_q^k)^2} \leq 0, \quad (16)$$

$$-y_q^k \leq v_q^k - x_q^k \leq y_q^k \quad \forall k \in K; \quad \forall q \in Q^k, \quad (17)$$

$$-\sum_{q \in Q^k} v_q^k + d^k p^k + \hat{d}^k s^k + \Gamma^k \sqrt{(\hat{d}^k r^k)^2} \leq 0 \quad \forall k \in K, \quad (18)$$

$$-s^k \leq p^k - r^k \leq s^k \quad \forall k \in K, \quad (19)$$

$$y_q^k \geq 0, \quad s^k \geq 0 \quad \forall k \in K; \quad \forall q \in Q^k,$$

Eqs. (2), (4)-(6), and (14).

It has been shown that the optimal solution of (RP3) will satisfy Constraints (12) and (13) with probability  $1 - \exp(-\Phi^2/2)$  and  $1 - \exp(-(\Gamma^k)^2/2)$ , respectively, where  $\Phi \geq 0$  and  $\Gamma^k \geq 0$  [26].

To simplify model (RP3), the substitution of  $p^k = 1$  for all  $k \in K$  into Constraints (18) and (19) yields the following constraints:

$$-\sum_{q \in Q^k} v_q^k + d^k + \hat{d}^k s^k + \Gamma^k \hat{d}^k |r^k| \leq 0, \quad \forall k \in K, \quad (20)$$

$$-s^k \leq 1 - r^k \leq s^k, \quad \forall k \in K. \quad (21)$$

The term  $|r^k|$  introduces non-linearity to Constraint (20). Two continuous variables ( $\lambda_1^k$  and  $\lambda_2^k$ ) and one binary variable ( $\lambda_3^k$ ) are introduced for each  $k \in K$ ; then,  $r^k$  and  $|r^k|$  are replaced by  $\lambda_1^k - \lambda_2^k$  and  $\lambda_1^k + \lambda_2^k$  in model (RP3) [39]. To show the relationship among  $\lambda_1^k$ ,  $\lambda_2^k$ , and  $\lambda_3^k$ , the following logical constraints must be considered, where  $M$  is an arbitrary, large value:

$$\lambda_1^k - M\lambda_3^k \leq 0, \quad (22)$$

$$\lambda_2^k - M(1 - \lambda_3^k) \leq 0. \quad (23)$$

Thus, finally, the robust counterpart of model (P), denoted by (RP4), is formulated as follows:

$$(RP4) \quad \min_{v_q^k, y_q^k, x_q^k, p^k, s^k, \lambda_1^k, \lambda_2^k, \lambda_3^k, \Psi, z^a} \Psi, \quad (24)$$

s.t.

$$-\sum_{q \in Q^k} v_q^k + d^k + \hat{d}^k s^k + \Gamma^k \hat{d}^k (\lambda_1^k + \lambda_2^k) \leq 0$$

$$\forall k \in K, \quad (25)$$

$$-s^k \leq 1 - (\lambda_1^k - \lambda_2^k) \leq s^k, \quad \forall k \in K, \quad (26)$$

$$\lambda_1^k, \lambda_2^k \geq 0, \quad \lambda_3^k \in \{0, 1\}, \quad \forall k \in K, \quad (27)$$

Eqs. (2), (4)-(6), (16), (17), (22), and (23).

After discarding the nonlinearity of Constraint (18), model (RP4) is an integer program with only one non-linear constraint, i.e., Constraint (16). However, this problem remains difficult to solve, even for small-sized problems; thus, an effective solution algorithm is required. The next section is devoted to development of the algorithm.

#### 4. Solution algorithm

This section develops a specialized branch-and-cut algorithm, that is, a branch-and-bound algorithm is integrated into cut generation to solve the relaxation

of model (*RP4*) at various branches of the branch-and-bound tree. To solve the model (*RP4*), Constraint (16) is removed from model (*RP4*), leaving only one integer linear program that possibly leads to a solution that violates Constraint (16). This violation can be managed by deriving a cut to remove the violated solution from the current feasible region.

The main component of the proposed algorithm approximates the nonlinear terms involved in Constraint (16). If  $f(\Psi, v_q^k, y_q^k, x_q^k)$  is equal to:

$$-\Psi + \sum_{k \in K} \sum_{q \in Q^k} \tau_q^k v_q^k + \sum_{k \in K} \sum_{q \in Q^k} \hat{\tau}_q^k y_q^k + \Phi \sqrt{\sum_{k \in K} \sum_{q \in Q^k} (\hat{\tau}_q^k x_q^k)^2},$$

then the linear approximation of this function for variables  $(\Psi, v_q^k, y_q^k, x_q^k)$  about  $(\bar{\Psi}, \bar{v}_q^k, \bar{y}_q^k, \bar{x}_q^k)$ , using the linear part of the Taylor series, is given by:

$$\begin{aligned} f(\Psi, v_q^k, y_q^k, x_q^k) \Big|_{\text{at } (\bar{\Psi}, \bar{v}_q^k, \bar{y}_q^k, \bar{x}_q^k)} &\approx f(\bar{\Psi}, \bar{v}_q^k, \bar{y}_q^k, \bar{x}_q^k) \\ &+ \left( \frac{\partial f}{\partial \Psi} \right) (\Psi - \bar{\Psi}) + \sum_{k \in K} \sum_{q \in Q^k} \left( \frac{\partial f}{\partial v_q^k} \right) (v_q^k - \bar{v}_q^k) \\ &+ \sum_{k \in K} \sum_{q \in Q^k} \left( \frac{\partial f}{\partial y_q^k} \right) (y_q^k - \bar{y}_q^k) \\ &+ \sum_{k \in K} \sum_{q \in Q^k} \left( \frac{\partial f}{\partial x_q^k} \right) (x_q^k - \bar{x}_q^k). \end{aligned} \quad (28)$$

The derivatives of  $f(\cdot)$  with respect to  $\Psi$ ,  $v_q^k$ ,  $y_q^k$ , and  $x_q^k$  are  $(-1)$ ,  $(\tau_q^k)$ ,  $(\hat{\tau}_q^k)$ , and  $\Phi \hat{\tau}_q^{k2} \bar{x}_q^k / \sqrt{\sum_{k \in K} \sum_{q \in Q^k} (\hat{\tau}_q^k \bar{x}_q^k)^2}$ , respectively; therefore, Expression (28) can be rewritten as follows:

$$\begin{aligned} f(\Psi, v_q^k, y_q^k, x_q^k) \Big|_{\text{at } (\bar{\Psi}, \bar{v}_q^k, \bar{y}_q^k, \bar{x}_q^k)} &\approx -\Psi + \sum_{k \in K} \sum_{q \in Q^k} \tau_q^k v_q^k + \sum_{k \in K} \sum_{q \in Q^k} \hat{\tau}_q^k y_q^k \\ &+ \Phi \sum_{k \in K} \sum_{q \in Q^k} \frac{\hat{\tau}_q^{k2} \bar{x}_q^k}{\sqrt{\sum_{k \in K} \sum_{q \in Q^k} (\hat{\tau}_q^k \bar{x}_q^k)^2}} x_q^k \\ &+ \Phi \left[ \sqrt{\sum_{k \in K} \sum_{q \in Q^k} (\hat{\tau}_q^k \bar{x}_q^k)^2} \right. \\ &\left. - \sum_{k \in K} \sum_{q \in Q^k} \frac{\hat{\tau}_q^{k2} \bar{x}_q^{k2}}{\sqrt{\sum_{k \in K} \sum_{q \in Q^k} (\hat{\tau}_q^k \bar{x}_q^k)^2}} \right]. \end{aligned} \quad (29)$$

Suppose that  $(\bar{\Psi}, \bar{v}_q^k, \bar{y}_q^k, \bar{x}_q^k, \bar{z}^a, \bar{s}^k, \bar{\lambda}_1^k, \bar{\lambda}_2^k, \bar{\lambda}_3^k)$  is the optimal value of variables  $(\Psi, v_q^k, y_q^k, x_q^k, z^a, s^k, \lambda_1^k, \lambda_2^k, \lambda_3^k)$  in the relaxation of (*RP4*). If this solution meets the integrality of  $z^a$  and  $\lambda_3^k$  and yet violates Constraint (16), the following cut, (Relation (30)), is added to the list of constraints to cut off this solution in the remainder of the process:

$$f(\Psi, v_q^k, y_q^k, x_q^k) \Big|_{\text{at } (\bar{\Psi}, \bar{v}_q^k, \bar{y}_q^k, \bar{x}_q^k)} \leq 0. \quad (30)$$

In the following, the detailed process of the proposed branch-and-cut algorithm is displayed to solve model (*RP4*):

**Step 0 (Initialization):** Set  $ub = +\infty$ ,  $lb = -\infty$  and initialize the input parameters;

**Step 1 (Termination test):** If the termination condition is met, then the current solution is optimal; hence, go to Step 6;

**Step 2 (Branch selection):** If  $\Upsilon$  (i.e., a list of active branches) is empty, then it indicates that all of the possible branches have been investigated, so go to Step 6; otherwise, choose branch  $i \in \Upsilon$  using the best-first-search strategy;

**Step 3 (Relaxation problem):** Solve the relaxation of model (*RP4*) (which is model (*RP4*) without Constraint (16) and without integrality constraints for  $z^a$  and  $\lambda_3^k$ ) for branch  $i$ . Let  $(\bar{\Psi}, \bar{v}_q^k, \bar{y}_q^k, \bar{x}_q^k, \bar{z}^a, \bar{s}^k, \bar{\lambda}_1^k, \bar{\lambda}_2^k, \bar{\lambda}_3^k)$  be the optimal solution and set  $lb = \max(lb, \bar{\Psi})$ ;

**Step 4 (Cut generation):** If  $(\bar{z}^a, \bar{\lambda}_3^k)$  is not the integer solution, then go to Step 3. If the value of:

$$\begin{aligned} -\bar{\Psi} + \sum_{k \in K} \sum_{q \in Q^k} \tau_q^k \bar{v}_q^k + \sum_{k \in K} \sum_{q \in Q^k} \hat{\tau}_q^k \bar{y}_q^k \\ + \Phi \sqrt{\sum_{k \in K} \sum_{q \in Q^k} (\hat{\tau}_q^k \bar{x}_q^k)^2} \end{aligned}$$

is strictly greater than zero, then add cut (30) to branch  $i$ ,  $\Upsilon \leftarrow \Upsilon + \text{branch } i$ , and go to Step 1. Otherwise (i.e.,  $-\bar{\Psi} + \sum_{k \in K} \sum_{q \in Q^k} \tau_q^k \bar{v}_q^k + \sum_{k \in K} \sum_{q \in Q^k} \hat{\tau}_q^k \bar{y}_q^k + \Phi \sqrt{\sum_{k \in K} \sum_{q \in Q^k} (\hat{\tau}_q^k \bar{x}_q^k)^2} \leq 0$ ), set:

$$\begin{aligned} ub = \min(ub, \sum_{k \in K} \sum_{q \in Q^k} \tau_q^k \bar{v}_q^k + \sum_{k \in K} \sum_{q \in Q^k} \hat{\tau}_q^k \bar{y}_q^k \\ + \Phi \sqrt{\sum_{k \in K} \sum_{q \in Q^k} (\hat{\tau}_q^k \bar{x}_q^k)^2}), \end{aligned}$$

and go to Step 1;

**Step 5 (Branching):** Branch on an integer variable with a fractional value in the optimal solution in Step 3, add two branches to set  $\Upsilon$ , and then go to Step 1;

**Step 6 (Exit):** End and report the optimal solution.

The algorithm begins at Step 0 by determining the initial value for  $lb$  and  $ub$ , and inputting the data into the blocking network. The sets of blocks ( $A$ ) and blocking paths ( $Q^k$ ) for each shipment  $k \in K$  are among the input sets. The set of blocks is chosen from all of the possible blocks being made from each origin to each yard, from each origin to each destination, from each yard to each destination, and from each yard to another yard. There exist many blocking paths for each shipment in the blocking network. The candidate set of blocking paths can be chosen in different fashions. One way is to use expert judgment to choose reasonable blocking paths. The other way is to use the  $k$ -shortest path algorithm and, then, choose the blocking paths with travel times that are less than the shortest travel time multiplied by a factor for each shipment. Barnhart et al. [5] suggested that this factor is equal to 1.5 to remove very long paths from the candidate set of blocking paths, assuming that no blocking path potentially in the optimal solution is ignored.

In Step 1, two termination conditions are defined, and the algorithm terminates as long as one condition is met. These conditions are: the relative gap (i.e.,  $(ub - lb)/ub$ ) is less than  $\varepsilon_r$ ; the absolute gap (i.e.,  $ub - lb$ ) is less than  $\varepsilon_a$ .

Step 2 is dedicated to choosing a branch from set  $\Upsilon$ . Two classical branch selection strategies are the depth-first-search and the best-first-search. The former selects the deepest branch of the tree, and the latter follows the opposite strategy by choosing the unexplored branches with the smallest lower bound [40]. In this paper, the best-first-search is used, which is the default setting of CPLEX [41]. Step 3 computes the solution of model (RP4) without Constraint (16) and without integrality constraints for  $z^a$  and  $\lambda_3^k$  (called relaxed model), and it obtains the optimal solution  $(\bar{\Psi}, \bar{v}_q^k, \bar{g}_q^k, \bar{x}_q^k, \bar{z}^a, \bar{s}^k, \bar{\lambda}_1^k, \bar{\lambda}_2^k, \bar{\lambda}_3^k)$ . In Step 4, a feasible solution to (RP4) is found provided that integer-feasible solution for  $z^a$  and  $\lambda_3^k$  is identified and it does not violate Constraint (16). Otherwise, the linear approximation of this constraint (or cut (30)) is added to the tree to cut off the current solution from the potential solutions. Step 5 performs the process of splitting the integer variable, which is not integer feasible, into two branches. For example, consider the solution for a branch with  $\bar{z}^a = 0.4$ . Two branches are created from this branch: The first branch is associated with  $z^a = 0$ , and the second branch is associated with  $z^a = 1$ . Finally, Step 6 reports the optimal solution obtained from the procedure.

## 5. Computational results

This section reports the computational results of the robust formulation of the railroad blocking problem under uncertainty in demand value and travel time for a number of hypothetical instances as well as the Iranian rail network. The proposed algorithm was coded in Visual Basic and compiled with Microsoft Visual Studio, version 2010 using CPLEX software, version 12.2, with default settings and callback features. The experiments were performed on a Dell Latitude computer featuring a 2.66 Intel processor with 4 GB of RAM running Windows 7, X32 edition. To test the efficiency of the proposed algorithm, the result of the algorithm was compared with the optimal solution found by GAMS 23.5 for model (RP4) with the COUENNE solver. All tests were run within a one-hour time limit, and  $\varepsilon_r$  and  $\varepsilon_a$  were assumed to be  $1e-4$  and  $1e-5$ , respectively.

### 5.1. Test instances

Our tests were performed on 14 instances to study the efficiency of the proposed algorithm. The instances are labeled with  $(A, B, C)$ , where  $A$ ,  $B$ , and  $C$  are the number of origins, classification yards, and destinations, respectively. Each instance is generated inside an  $(A - 1) \times 100^{\text{km}} \times (A - 1) \times 100^{\text{km}}$  region. Origin  $a (= 1, \dots, A)$ , classification yard  $b (= 1, \dots, B)$ , and destination  $c (= 1, \dots, C)$  are located at coordinates  $(0, 100 \times (a - 1))$ ,  $(50 \times (A - 1), 100 \frac{A-1}{B-1} \times (b - 1))$ , and  $(50 \times (A - 1), 100 \frac{A-1}{C-1} \times (c - 1))$ , respectively. The generic region is illustrated in Figure 2.

It is necessary to limit the number of blocking paths for each shipment to ensure that GAMS does not terminate due to out-of-memory error. Herein, we assume that all blocking paths, which contribute to achievement of the optimal solution, are in the candidate set of blocking paths; therefore, in this paper, the blocking paths are considered for each shipment whose travel time is less than the shortest travel time multiplied by 3 (instead of 1.5 reported by Barnhart et al. [5]).

### 5.2. Evaluation of the proposed algorithm

Fourteen different instances from (2,2,2) to (20,20,20) were considered;  $\Phi = 0.1$ ,  $\Gamma^k = 0.1$  for all  $k \in K$  were assumed. The demand ( $\tilde{d}^k$ ) ranges were  $[d^k - \hat{d}^k, d^k + \hat{d}^k]$ , where  $d^k = 1000$  and  $\hat{d}^k = 200$  for all  $k \in K$ ; the travel time ( $\tilde{\tau}_q^k$ ) ranges were  $[\tau_q^k - \hat{\tau}_q^k, \tau_q^k + \hat{\tau}_q^k]$ , where  $\tau_q^k$  is calculated based on the average train speed being equal to 60 km/h, and the value of  $\hat{\tau}_q^k$  was set as a random value between  $0.1 \times \tau_q^k$  to  $0.2 \times \tau_q^k$ . Table 2 shows the test results for 14 instances and compares the effectiveness of the proposed algorithm and GAMS.

This table shows that the proposed algorithm outperforms the GAMS results, and the algorithm is able to solve the instances optimally within one hour.



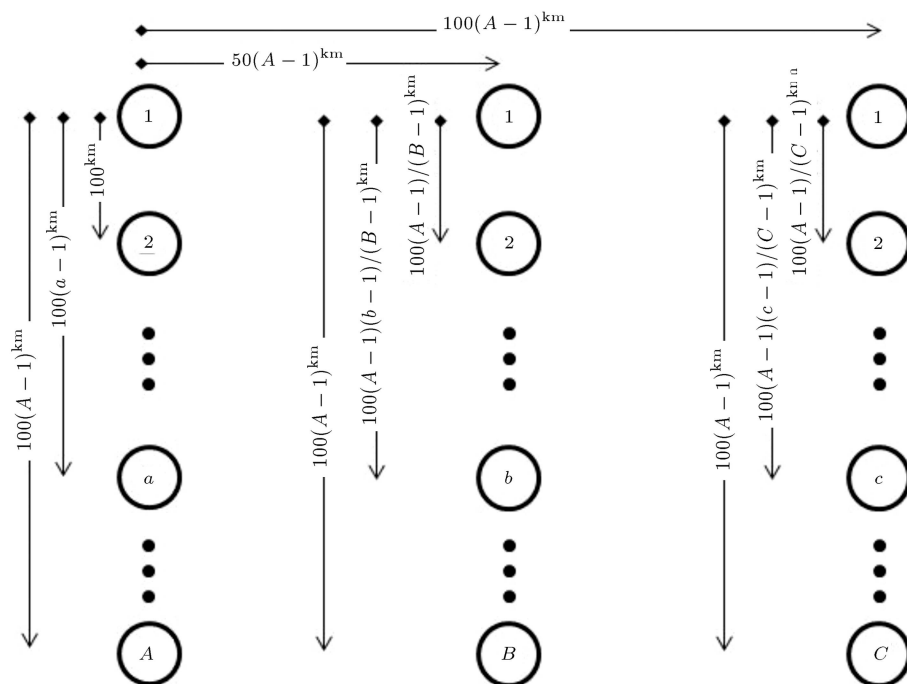
Figure 2. A scheme of the instances ( $A$ ,  $B$ ,  $C$ ).

Table 2. Computational results of 14 test instances.

Instance ( $A^1$ , $B^2$ , $C^3$ )	GAMS					Our algorithm					
	CPU time (s)	IP solution <sup>4</sup>	Lower bound	$A_{\text{gap}}^5$	$R_{\text{gap}}^6$	CPU time (s)	#Cuts	IP solution	Lower bound	$A_{\text{gap}}$	$R_{\text{gap}}$
2.2.2	8.31	141.93	141.92	0.01	6e-5	0.05	16	141.94	141.92	0.01	9e-5
3.3.3	37.90	347.76	347.75	0.01	5e-5	0.06	15	347.77	347.74	0.03	7e-5
4.4.4	> 1hr	909.03	907.38	1.65	1.8e-3	0.22	24	909.02	908.93	0.08	9e-5
5.5.5	> 1hr	1874.26	1870.70	3.56	1.9e-3	0.40	35	1873.98	1873.82	0.17	8e-5
6.6.6	> 1hr	3352.26	3346.66	5.6	1.7e-3	1.08	48	3351.65	3351.33	0.32	9e-5
7.7.7	> 1hr	5451.35	5443.61	7.74	1.4e-3	2.72	65	5450.53	5450.05	0.48	8e-5
8.8.8	—	—	—	—	—	5.93	72	8279.55	8278.82	0.73	8e-5
9.9.9	—	—	—	—	—	18.24	89	11947.40	11946.25	1.15	9e-5
10.10.10	—	—	—	—	—	42.48	101	16600.13	16598.53	1.60	9e-5
12.12.12	—	—	—	—	—	164.65	126	29071.42	29068.53	2.88	9e-5
14.14.14	—	—	—	—	—	439.02	142	46675.08	46670.47	4.62	9e-5
16.16.16	—	—	—	—	—	1371.9	155	70244.03	70237.05	6.98	9e-5
18.18.18	—	—	—	—	—	2389.9	156	100648.22	100638.18	10.03	9e-5
20.20.20	—	—	—	—	—	3498.3	180	138757.45	138743.65	13.80	9e-5

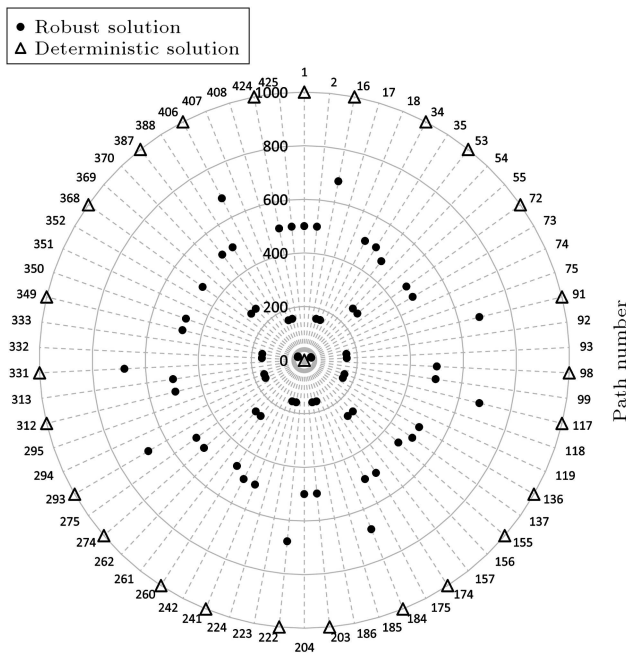
1. Number of origins; 2. Number of classification yards; 3. Number of destinations;

4. IP optimal value of the objective function in the unit of car-hr; 5. Absolute gap; 6. Relative gap.

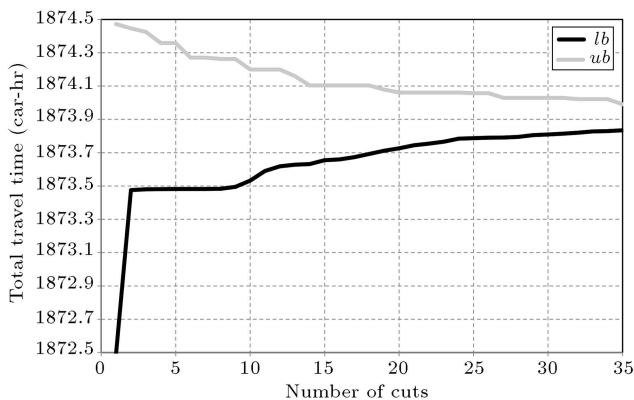
However, GAMS could only solve 2 of the 14 instances optimally within one hour. Moreover, it is notable that this software was not even able to find a feasible solution for 8 of the 14 instances.

Figure 3 shows the optimal values of flow variables for the deterministic and robust models in instance (5,5,5). In this figure, when moving away from the cen-

ter of the circle, the value of the flow variable increases on the logarithmic scale. It can be observed that the deterministic model addresses fewer paths with more flow value, whereas the robust model addresses more paths with less flow value. Specifically, there are 25 paths (a single path for each of 25 shipments) with positive flow and the remainder of the paths with zero



**Figure 3.** Illustration of number of paths for deterministic and robust optimal solutions for instance (5,5,5).



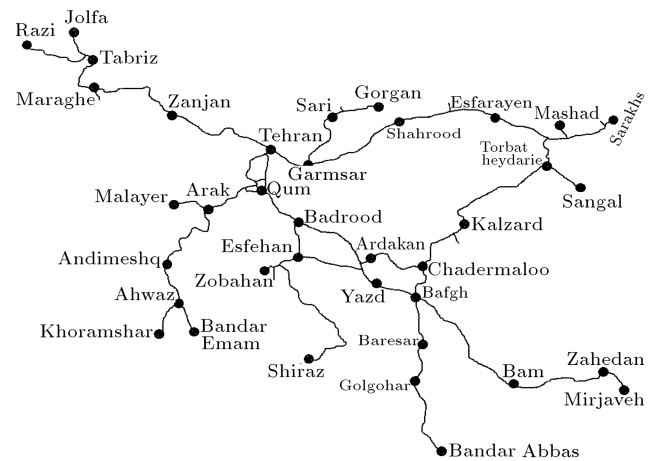
**Figure 4.** The value of the lower and upper bounds versus the number of cuts for instance (5,5,5).

values, overlapping in the center of the circle. However, the robust solution contains 66 paths with positive values, and each shipment is shipped with more than one path.

Figure 4 plots the value of the lower bound and upper bound versus the number of cuts, for instance, (5,5,5) in Step 4. As expected, the gap decreases as cuts are added.

### 5.3. A case-study: Iran rail network

The real data were provided by studying railways in Iran to illustrate the application of the proposed algorithm to the blocking problem under demand and travel time uncertainty. The Iran rail network, which is approximately 12000 km length, has 334 stations, 44 classification yards, 1883 shipments, and 5058 blocks.



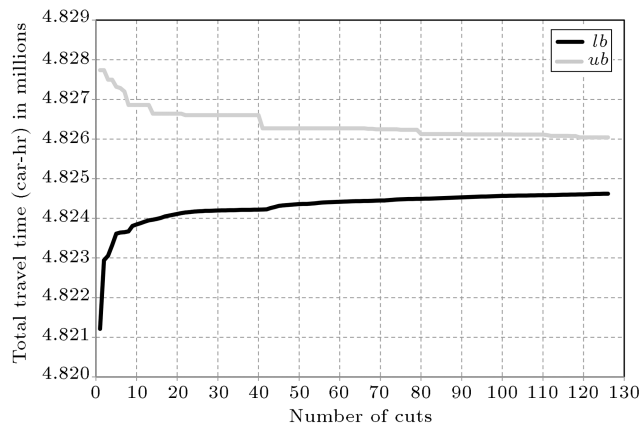
**Figure 5.** The abstract form of Iran rail network.

An abstract illustration of the Iranian rail network is provided in Figure 5.

Similar to Section 5.1, the set of blocking paths for any shipment is limited by the paths with average travel times which are three times that of the path with the shortest average travel time for this shipment. The Iranian network includes 1883 O-D pairs with a demand value approximately varying from 5 to 200000 tons for a three-month horizon in the railways of Iran. To set the value of demand fluctuations, the time history data provided by the railways of Iran were analyzed.  $\hat{d}^k$  was calculated as  $\max_{o \in \Pi^k} (d_o^k - d^k)$ , where  $d_o^k$  is an  $o$ th observation of demand value for shipment  $k$ ;  $d^k$  is the average demand value;  $\Pi^k$  is the set of all observations for shipment  $k$ . Similarly,  $\hat{\tau}_q^k$  was determined by  $\max_{p \in \Delta_q^k} (\tau_{q,p}^k - \tau_q^k)$ , where  $\tau_{q,p}^k$  is the  $p$ th

observation of travel time for  $q \in Q^k$   $k \in K$ ;  $\tau_q^k$  is the average travel time;  $\Delta_q^k$  is the set of all observations of travel time. Further, it was supposed that  $\Phi = 2.8$  and  $\Gamma^k = 2.8$  for all of the shipments. According to the values of  $\Phi$  and  $\Gamma^k$ , the optimal solution of model (RP4) would be valid with Constraints (12) and (13) with a probability more than 0.98, and the rail manager could guarantee (with a probability of 0.98) the practicality of the decisions to be made.

Figure 6 illustrates the values of the lower and upper bounds for the objective function, i.e., total travel time, obtained by the proposed algorithm for the Iranian rail network. The algorithm hits the one-hour time limit after adding 128 cuts and reaches the relative gap as equal to 0.034%, which is acceptable for the real-sized problem. The optimal value of model (P) with the deterministic data was approximately 4.3 million car-hr, and the optimal value of the robust model was approximately 4.8 million car-hr, which is roughly a 10% increase; the robust solution also guarantees the validity of Constraints (12) and (13) with a probability of more than 0.98. Using the values of time, the



**Figure 6.** The value of lower and upper bounds for Iran rail network versus number of cuts.

economic benefits of a deterministic solution can be quantified to compare it to the robust solution. It has been reported that the value of time for freight rail transportation is approximately 34 dollars (30 euros) per ton-hr in Europe [42]. Therefore, the price of robustness against uncertainty is approximately 425 million dollars in the three-month time horizon (where net car load is 25 tons).

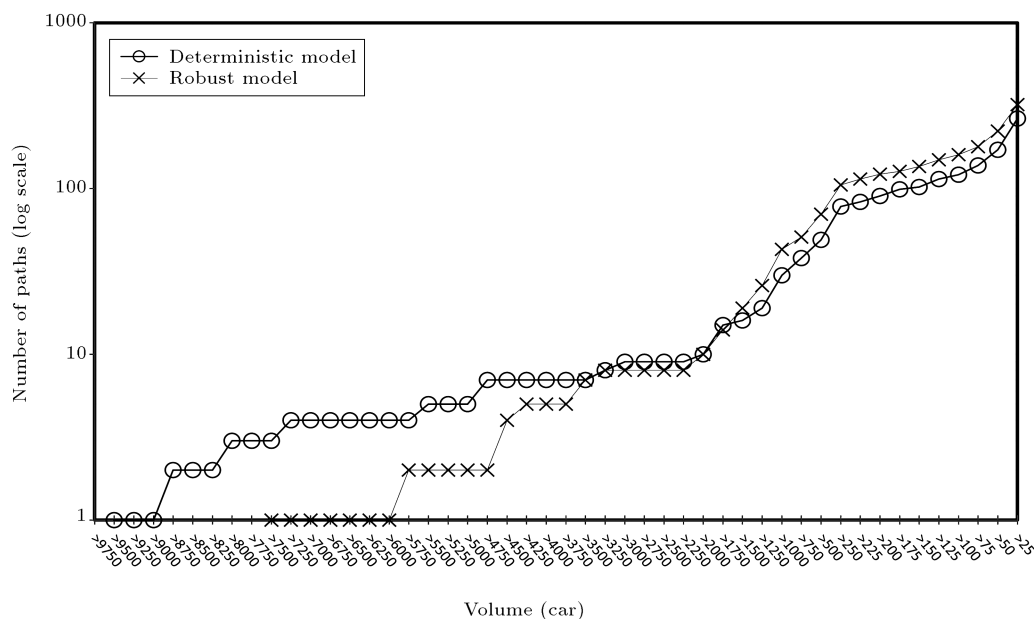
The optimal solution to the robust model involves 1855 selected blocks and 1937 selected blocking paths, while the optimal solution to the deterministic model involves 1822 selected blocks and 1903 selected blocking paths. To emphasize the difference between the optimal solutions of the deterministic and robust models in detail, the optimal solutions for block variables and flow variables from the two models are given in Table 3. For the sake of convenience in showing the results, the optimal solutions for the deterministic

**Table 3.** The comparison of the optimal solutions for deterministic and robust models.

Condition	Number of cases happened
$z^{a,D} = 1$ and $z^{a,R} = 1$	1270
$z^{a,D} = 0$ and $z^{a,R} = 1$	585
$z^{a,D} = 1$ and $z^{a,R} = 0$	552
$v_q^{k,D} > 0$ and $v_q^{k,R} > 0$	943
$v_q^{k,D} = 0$ and $v_q^{k,R} > 0$	994
$v_q^{k,D} > 0$ and $v_q^{k,R} = 0$	960

model are denoted by  $z^{a,D}$  and  $v_q^{k,D}$ , and the optimal solutions for the robust model are denoted by  $z^{a,R}$  and  $v_q^{k,R}$ . For block variables, 1270 blocks were common in the optimal solution of both models; however, 1137 blocks were not involved in the optimal solution of either model. For flow variables, 934 paths for the robust model were also involved in the deterministic model; however, 960 paths in the robust solution were equal to zero that were positive in the deterministic solution. This observation shows that the robust solution includes more paths than the deterministic solution does, which is also illustrated in [43,44] for the shortest robust path problem.

Figure 7 illustrates the number of blocking paths with volumes greater than the amount of flow shown on the horizontal axis. For example, consider the value of 5000 cars from the horizontal axis; then, the corresponding values on the vertical axis indicate the numbers of blocking paths with volumes greater than 5000 cars. This figure illustrates that, when moving from left to right on the horizontal axis, the deterministic line meets the robust line at a point in



**Figure 7.** The volume of blocking paths for deterministic and robust models.

the middle. It can be derived from the figure that the optimal solution to the deterministic model consists of more blocking paths with a larger volume. In particular, the deterministic model yields four paths with volumes greater than 6000 cars; however, the optimal solution to the robust model has only one path with a volume greater than 6000 cars. However, this figure also suggests that the robust solution includes more blocking paths with less volume than the deterministic model. The authors also generated similar figures for different networks, and found that these figures follow the same trends.

## 6. Conclusion

The railroad blocking problem is an important issue at the tactical planning level. This problem includes identifying the list of visited classification yards for each shipment to improve the performance of car shipping while satisfying railway restrictions. Because the blocking parameters change monthly, they greatly affect the quality of the deterministic solution and can possibly render this solution meaningless in practice. Thus, the blocking plan must be re-optimized for new realizations of uncertain parameters; then, the next stage problems (such as train scheduling) should also be re-optimized. A methodology addressing this need is robust optimization [26], which produces a solution immunized against uncertain data. The robust model is an optimization program with continuous flow variables, integer block variables, and nonlinear functions in the constraints. To solve the program, we developed a specialized branch-and-cut algorithm. This method is viewed as a generalization of the branch-and-bound method, with additional cuts generated on eligible branches of the branch-and-bound tree. To test the efficiency of the proposed algorithm, the results of the proposed algorithm were compared with the optimal solution found by GAMS on 14 different instances. We observed that the proposed algorithm solved all instances optimally within one hour. However, GAMS could only solve 2 of the 14 instances optimally within one hour, and it was not able to find a feasible solution for 8 of the 14 over one hour. Finally, the proposed algorithm was applied to a real-sized blocking problem. It was shown that, by ignoring 10% of the objective function of the deterministic model, which is approximately 425 million dollars for a three-month horizon time, we were able to develop an optimal solution that was immune to data uncertainty with a probability of more than 0.98. In addition, the optimal solution of the robust model was compared with that of the deterministic model, and our investigation showed that there is a fundamental difference between the optimal values of the block and path variables for the two models.

## References

1. Crainic, T.G. and Laporte, G. "Planning models for freight transportation", *Eur. J. Oper. Res.*, **97**, pp. 409-438 (1997).
2. Crainic, T.G. "Service network design in freight transportation", *Eur. J. Oper. Res.*, **122**, pp. 272-288 (2000).
3. Assad, A.A. "Modelling of rail networks: Toward a routing/makeup model", *Transport Res. B-Meth*, **14**, pp. 101-114 (1980).
4. Fügenschuh, A., Homfeld, H., and Schülldorf, H. "Single-car routing in rail freight transport", *Transport Sci.*, **49**, pp. 130-148 (2013).
5. Barnhart, C., Jin, H., and Vance, P.H. "RailRoad blocking: A network design application", *Oper. Res.*, **48**, pp. 603-614 (2000).
6. Bontekoning, Y. and Priemus, H. "Breakthrough innovations in intermodal freight transport", *Transport Plan Techn.*, **27**, pp. 335-345 (2004).
7. Bodin, L.D., Golden, B.L., Schuster, A.D., and William, R. "A model for the blocking of trains", *Transport Res B-Meth*, **14**, pp. 115-150 (1980).
8. Marín, A. and Salmerón, J. "Tactical design of rail freight networks. Part I: Exact and heuristic methods", *Eur. J. Oper. Res.*, pp. 26-44 (1996).
9. Lin, B.-L., Wang, Z.-M., Ji, L.-J., Tian, Y.-M., and Zhou, G.-Q. "Optimizing the freight train connection service network of a large-scale rail system", *Transportation Research Part B: Methodological*, **46**, pp. 649-667 (2012).
10. Newton, H.N., Barnhart, C., and Vance, P.H. "Constructing railroad blocking plans to minimize handling costs", *Transport Sci.*, **32**, pp. 330-345 (1998).
11. Keaton, M.H. "Designing optimal railroad operating plans: Lagrangian relaxation and heuristic approaches", *Transportation Research: Part B*, **23**, pp. 415-431 (1989).
12. Hasany, R.M., Shafahi, Y., and Kazemi, S.F. "A comprehensive formulation for railroad blocking problem", *ECMS*, pp. 758-763 (2013).
13. Ahuja, R.K., Jha, K.C., and Liu, J. "Solving real-life railroad blocking problems", *Interfaces*, **37**, pp. 404-419 (2007).
14. Martinelli, D.R. and Teng, H. "Optimization of railway operations using neural networks", *Transport Res C-Emer*, **4**, pp. 33-49 (1996).
15. Yaghini, M., Momeni, M., Sarmadi, M., Seyedabadi, M., and Khoshraftar, M.M. "A fuzzy railroad blocking model with genetic algorithm solution approach for Iranian railways", *Appl. Math. Model*, **20**, pp. (2015).
16. Yue, Y., Zhou, L., Yue, Q., and Fan, Z. "Multi-route railroad blocking problem by improved model and ant colony algorithm in real world", *Comput. Ind. Eng.*, **60**, pp. 34-42 (2011).

17. Yaghini, M., Ahadi, H.R., Barati, E., and Saghian, Z. "Tabu search algorithm for the railroad blocking problem", *J. Transp. Eng.-ASCE*, **139**, pp. 216-222 (2012).
18. Ahuja, R.K., Cunha, C.B., and Sahin, G. "Network models in railroad planning and scheduling", *Tut. Oper. Res.*, **1**, pp. 54-101 (2005).
19. Voll, R. and Clausen, U. "Branch-and-price for a European variant of the railroad blocking problem", *Electronic Notes in Discrete Mathematics*, **41**, pp. 45-52 (2013).
20. Yaghini, M., Rahbar, M., Karimi, M., and Khoshkroudi, M. "A branch-and-price algorithm for solving the railroad blocking problem", *International Journal of Engineering Science (2008-4870)*, **25**, pp. 99-108 (2014).
21. Yang, L., Gao, Z., and Li, K. "Railway freight transportation planning with mixed uncertainty of randomness and fuzziness", *Appl. Soft Comp.*, **11**, pp. 778-792 (2011).
22. Gao, Y., Yang, L., and Li, S. "Uncertain models on railway transportation planning problem", *Appl. Math. Model*, **40**, pp. 4921-4934 (2016).
23. Meng, Q., Hei, X., Wang, S., and Mao, H. "Carrying capacity procurement of rail and shipping services for automobile delivery with uncertain demand", *Transport Res. E-Log*, **82**, pp. 38-54 (2015).
24. Milenković, M.S., Bojović, N.J., Švadlenka, L., and Melichar, V. "A stochastic model predictive control to heterogeneous rail freight car fleet sizing problem", *Transport Res. E-Log*, **82**, pp. 162-198 (2015).
25. Jin, H., *Designing Robust Railroad Blocking Plans*, Massachusetts Institute of Technology (1998).
26. Ben-Tal, A. and Nemirovski, A. "Robust solutions of Linear Programming problems contaminated with uncertain data", *Math. Prog.*, **88**, pp. 411-424 (2000).
27. Prékopa, A., *Stochastic Programming*, Springer Science & Business Media (2013).
28. Wong, W.G., Niu, H., and Ferreira, L. "A fuzzy method for predicting the demand for rail freight transportation", *J. Adv. Transport*, **37**, pp. 159-171 (2003).
29. Carvalho, G., *Methodology for Railway Demand Forecasting Using Data Mining*, SAS Global Forum (2007).
30. Gorman, M.F. "Statistical estimation of railroad congestion delay", *Transport Res. E-Log*, **45**, pp. 446-456 (2009).
31. Matas, A., Raymond, J.L., Gonzalez-Savignat, M., and Ruiz, A. "Predicting the demand: Uncertainty analysis and prediction models in Spain", *Working Paper in Economic Evaluation of Transportation Projects*, pp. 1-31 (2009).
32. Dong, Y. "Modeling rail freight operations under different operating strategies", PhD Thesis, MIT, Cambridge, MA (1997).
33. Birge, J.R. and Louveaux, F., *Introduction to Stochastic Programming*, Springer (2011).
34. Soyster, A.L. "Convex programming with set-inclusive constraints and applications to inexact linear programming", *Oper. Res.*, **21**, pp. 1154-1157 (1973).
35. Ben-Tal, A. and Nemirovski, A. "Robust convex optimization", *Mathematics of Operations Research*, **23**, pp. 769-805 (1998).
36. Seref, O., Ahuja, R.K., and Orlin, J.B. "Incremental network optimization: Theory and algorithms", *Oper. Res.*, **57**, pp. 586-594 (2009).
37. Han, J., Lee, C., and Park, S. "A robust scenario approach for the vehicle routing problem with uncertain travel times", *Transport Sci.*, **48**, pp. 373-390 (2013).
38. Smith, J.C., Ahmed, S., Cochran, J.J., Cox, L.A., Keshinocak, P., Kharoufeh, J.P., and Smith, J.C. "Introduction to robust optimization", *Wiley Encyclopedia of Operations Research and Management Science*, John Wiley & Sons, Inc. (2010).
39. Nemhauser, G.L. and Wolsey, L.A. "Integer programming and combinatorial optimization", W., Chichester, G.L. Nemhauser, M.W.P. Savelsbergh, and G.S. Sigismondi, *Constraint Classification for Mixed Integer Programming Formulations*, *COAL Bulletin*, **20**, pp. 8-12 (1988).
40. Lee, J. and Leyffer, S., *Mixed Integer Nonlinear Programming*, Springer Science & Business Media (2011).
41. ILOG, CPLEX OPTIMIZATION, INC. 2009. Using the CPLEX Linear Optimizer.
42. Van Essen, H., Boon, B., den Boer, L., Faber, J., van den Bossche, M., Vervoort, K., and Rochez, C., *Marginal Costs of Infrastructure Use - Towards a Simplified Approach*, Delft (2003).
43. Bertsimas, D. and Sim, M. "Robust discrete optimization and network flows", *Math. Prog.*, **98**, pp. 49-71 (2003).
44. Ben-Tal, A., El Ghaoui, L., and Nemirovski, A., *Robust Optimization*, Princeton University Press, New Jersey (2009).

## Appendix

Ben-Tal and Nemirovski [26] introduced the following methodology for generating the robust formulation from linear model while the technological coefficient are uncertain. Consider the following optimization problem [26]:

$$(AP1) \min \sum_{i=1}^n c_i x_i, \quad (A.1)$$

s.t.

$$\sum_{i=1}^n \tilde{a}_{ij} x_i \leq b_j, \quad j = 1, \dots, m, \quad (A.2)$$

$$l_i \leq x_i \leq u_i, \quad i = 1, \dots, n. \quad (A.3)$$

Suppose that, in constraint  $j$ ,  $\tilde{a}_{ij} = (1 + \varepsilon \xi_{ij})|a_{ij}|$  where  $\xi_{ij}$  is the independent random variables symmetrically distributed in interval  $[-1, 1]$ ,  $\varepsilon$  is the tolerance affecting  $\tilde{a}_{ij}$ , and  $|a_{ij}|$  is the average value of  $\tilde{a}_{ij}$ . Ben-Tal and Nemirovski [26] proved the following theorem to generate the robust formulation of (AP1).

**Theorem.** Let  $\hat{\mathbf{x}}$  be the optimal solution to the following optimization problem:

$$(AP2) \min \sum_{i=1}^n c_i x_i, \quad (A.4)$$

s.t.

$$\sum_{i=1}^n a_{ij} x_i + \sum_{i=1}^n \varepsilon |a_{ij}| y_{ij} + \varepsilon \Omega_j \sqrt{\sum_{i=1}^n a_{ij}^2 z_{ij}^2} \leq b_j, \quad j = 1, \dots, m, \quad (A.5)$$

$$-y_{ij} \leq x_i - z_{ij} \leq y_{ij}, \quad i = 1, \dots, n; \quad j = 1, \dots, m, \quad (A.6)$$

$$l_i \leq x_i \leq u_i, \quad i = 1, \dots, n, \quad (A.7)$$

$$y_{ij} \geq 0, \quad i = 1, \dots, n; \quad j = 1, \dots, m. \quad (A.8)$$

Then,  $\hat{\mathbf{x}}$  is feasible for (AP1) with probability  $(1 - \exp(-\Omega_j^2/2))$  for constraint  $j$  where  $\Omega_j$  is a positive parameter [26].

**Proof:** Let  $\hat{\mathbf{x}}$  be feasible for (AP2). Then:

$$\begin{aligned} & \Pr \left[ \sum_{i=1}^n \tilde{a}_{ij} x_i > b_j \right] \\ & \stackrel{(a)}{=} \Pr \left[ \sum_{i=1}^n (a_{ij} + \varepsilon |a_{ij}| \xi_{ij}) x_i > b_j \right] \\ & = \Pr \left[ \sum_{i=1}^n a_{ij} x_i + \sum_{i=1}^n \varepsilon |a_{ij}| x_i \xi_{ij} > b_j \right] \\ & \stackrel{(b)}{=} \Pr \left[ \sum_{i=1}^n a_{ij} x_i + \sum_{i=1}^n \varepsilon |a_{ij}| x_i \xi_{ij} \right. \\ & \quad \left. + \sum_{i=1}^n \varepsilon |a_{ij}| y_{ij} \xi_{ij} - \sum_{i=1}^n \varepsilon |a_{ij}| y_{ij} \xi_{ij} > b_j \right] \\ & = \Pr \left[ \sum_{i=1}^n a_{ij} x_i + \sum_{i=1}^n \varepsilon |a_{ij}| (x_i - y_{ij}) \xi_{ij} \right. \\ & \quad \left. + \sum_{i=1}^n \varepsilon |a_{ij}| y_{ij} \xi_{ij} > b_j \right] \end{aligned}$$

$$\stackrel{(c)}{\leq} \Pr \left[ \sum_{i=1}^n a_{ij} x_i + \sum_{i=1}^n \varepsilon |a_{ij}| z_{ij} \xi_{ij} + \sum_{i=1}^n \varepsilon |a_{ij}| y_{ij} \xi_{ij} > b_j \right]$$

$$\stackrel{(d)}{\leq} \Pr \left[ \sum_{i=1}^n \varepsilon |a_{ij}| z_{ij} \xi_{ij} > \Omega_j \varepsilon \sqrt{\sum_{i=1}^n a_{ij}^2 z_{ij}^2} \right]$$

$$\stackrel{(e)}{=} \Pr \left[ \sum_{i=1}^n \eta_{ij} \xi_{ij} > \Omega_j \sqrt{\sum_{i=1}^n \eta_{ij}^2} \right]$$

$$\stackrel{(f)}{=} \Pr \left[ \sum_{i=1}^n \eta_{ij} \xi_{ij} > \Omega_j \right]$$

$$\stackrel{(g)}{\leq} \exp(-\Omega_j^2) \mathbf{E} \left\{ \exp \left\{ \Omega_j \sum_{i=1}^n \eta_{ij} \xi_{ij} \right\} \right\}$$

$$\stackrel{(h)}{=} \exp(-\Omega_j^2) \prod_{i=1}^n \mathbf{E} \{ \exp \{ \Omega_j \eta_{ij} \xi_{ij} \} \}$$

$$\stackrel{(i)}{=} \exp(-\Omega_j^2) \prod_{i=1}^n \left[ \sum_{l=0}^{\infty} \frac{(\Omega_j \eta_{ij})^{2l}}{(2l)!} \right]$$

$$\stackrel{(j)}{=} \exp(-\Omega_j^2) \exp \left( \Omega_j^2 / 2 \times \sum_{i=1}^n \eta_{ij}^2 \right)$$

$$\stackrel{(k)}{=} \exp(-\Omega_j^2) \exp(\Omega_j^2 / 2) = \exp(-\Omega_j^2 / 2).$$

In (a), we substitute  $\tilde{a}_{ij}$  into  $(1 + \varepsilon \xi_{ij})|a_{ij}|$ , in (b) we add and subtract  $\sum_{i=1}^n \varepsilon |a_{ij}| y_{ij} \xi_{ij}$ , in (c) we consider  $x_i - y_{ij} \leq z_{ij}$ , in (d) we consider  $\sum_{i=1}^n a_{ij} x_i + \sum_{i=1}^n \varepsilon |a_{ij}| y_{ij} + \varepsilon \Omega_j \sqrt{\sum_{i=1}^n a_{ij}^2 z_{ij}^2} \leq b_j$ , in (e) we substitute  $\eta_{ij}$  into  $|a_{ij}| z_{ij}$ , in (f) we assume the case of  $\sum_{i=1}^n \eta_{ij}^2 = 1$ , in (g) we use the Tschebyshev inequality, in (h) we use independence of  $\eta_{ij}$ , in (i) we use the definition of expected value of EXP and symmetry of the distribution of  $\eta_{ij} = [-1, 1]$ , in (j) we use Taylor series, and in (k) we assume the case of  $\sum_{i=1}^n \eta_{ij}^2 = 1$ .  $\square$

In the above discussion, Only the uncertainty which is inherent in the technological coefficients is considered for model (AP1) (used by Ben-Tal and Nemirovski [26]). However, the blocking problem involves uncertainty in both the objective function and right-hand side of the constraint. In the paper, we applied some modifications to the uncertain model to

adopt the idea of Ben-Tal and Nemirovski [26] (stated above) to present the robust formulation for blocking problem.

### Biographies

**Reza Mohammad Hasany** received his Bachelor of Science degree in Civil Engineering in 2007 and his Master of Science degree in Transportation Planning in 2010 from Sharif University of Technology, Tehran, Iran. Currently, He is a Doctoral Student at Department of Civil Engineering in Sharif University of Technology. In his PhD's thesis, he worked on modeling and solution algorithm for the railroad blocking problem. His research interests include mathematical program-

ming, large-scale optimization, and optimization under uncertainty.

**Yousef Shafahi** obtained his BS from Shiraz University, Shiraz, Iran; MS from Isfahan University of Technology, Isfahan, Iran; and PhD from University of Maryland, College Park, MD, USA. He joined the Department of Civil Engineering in Sharif University of Technology, Tehran, Iran as an Assistant Professor in 1998 and is now a Professor and the Head of Transportation Division group at this Department. His research interests are urban transportation planning; road, rail, and air transportation; transit networks design and scheduling; applications of operation research, soft computing and simulation in transportation.