A hybrid method to solve railroad passenger scheduling problem

Masoud Yaghini[a*], Alireza Alimohammadian[a], and **Samaneh Sharifi[b]**

[a]Department of Railroad Transportation, Iran University of Science & Technology, Tehran, Iran
[b]Department of Industrial Engineering, KNTU University of Technology, Tehran, Iran

A R T I C L E I N F O	A B S T R A C T
	Railroad transportation planning is strategically a long term and an important decision making problem especially in the area of travelling passengers. There have been literally various methods to use in order to provide optimum traveling schedule such as direct or indirect methods. Direct solutions involve the implementation of mixed integer programming, which is often hard to solve for real-world applications. The proposed model of this paper uses a column generation method to decompose a large-scale railroad passenger-scheduling problem into some smaller scale problems, which are easier to solve. The primary concern with the resulted problem is that final solutions of the method need to be integer and this is in contrast with convexity assumption of column generation techniques. We propose heuristic method to handle this problem and apply the proposed model for some examples. The preliminary results indicate that the proposed model of this paper could provide optimal solutions for small-scale problems and it can reach some reasonable solutions for larger problems when direct implementation fails to do in reasonable amount of time.

1. Introduction

Train scheduling is one of the most important issues in railroad industry. During the past few decades, there have been tremendous changes in world's railroad industry and it is believed to be one of the most and secure facilities for transporting passengers between cities. During the past few decades, there have been significant attempts to provide efficient methods for cost management of train scheduling.

D'Ariano et al. (2007) studied a train scheduling problem with railway infrastructure managers when they are in real-time traffic control. Once train operations are faced with uncertainty, a new conflict-free timetable of feasible arrival and departure times needs to be re-computed so that we can minimize the deviation from the original one. The problem can be stated as a big job shop problem with no-store constraints. They estimated time separation among trains, and modeled the scheduling problem with an alternative graph formulation. They also developed a branch and bound method, which enhances implication rules enabling to increase the computation. They applied their method on

* Corresponding author. Tel: +982177240540
E-mail addresses: yaghini@iust.ac.ir (M. Yagini)

a bottleneck area of the Dutch rail network and explained that a truncated version of the method could give optimal or near optimal solutions in less computational time. Peeters and Kroon (2008) proposed a model and a branch-and-price technique to determine an efficient railway rolling stock circulation on a set of train lines. Lindner (2000) and Scholl (2005) in their Phd work presented a comprehensive study on railroad scheduling problems using different techniques such as branch-and-price method. Ghoseiri and Ghannadpour (2010) presented a hybrid genetic algorithm to solve a multi-depot homogenous locomotive assignment problem with time windows. They first introduced a mathematical model using vehicle routing problem with time windows (VRPTW) and then they used a cluster-first, route-second approach to change the multi-depot locomotive assignment into a set of single depot problems and provided the solutions for each problem, independently.

They performed some experimental results with a set of 15 completely random generated instance problems and reported some encouraging results. Chung et al. (2009) presented a hybrid genetic algorithm for train sequencing in the Korean railway similar to what Ghoseiri and Ghannadpour did. There are also several works to use heuristic methods for solving railways problems (Lee & Chen, 2009). There are also several works, which use Neural network to solve railway problems. Tsai et al. (2009), for instance used Neural network based temporal feature models for short-term railway passenger demand forecasting. Another alternative solution to determine the capacity of train scheduling is to use system dynamic methods (Suryani et al., 2010). Bussieck et al. (1996) presented fuzzy theory for train scheduling problems. Claessens et al. (1998) used location techniques for cost allocation of railroad problems. Burdett and Kozan (2009) presented various techniques for inserting additional trains into existing timetables.

Liu and Kuzan (2009) considered the train scheduling problem as a blocking parallel-machine job shop scheduling (BPMJSS) problem. In their modeling formulation, all trains, single-track sections and multiple-track segments, respectively, are synonymous with jobs, single and parallel machines, and an operation was considered as the movement/traversal of a particular train across a section. They solved a parallel-machine job-shop-scheduling (PMJSS) problem using an improved shifting bottleneck procedure (SBP) method without considering blocking conditions for a real-world application from Queensland Rail and analyzed the results. The results indicated that their proposed methodology would be a very useful tool for the real-life train scheduling problems.

Mesquita and Paias (2008) used set partitioning/covering-based methods for the integrated vehicle and crew scheduling problem. They presented an algorithm with two stages, which starts with a pre-processing phase to define the set of tasks and to reach an initial set of duties. In a second phase, they used a linear programming relaxation of the models using a column generation scheme and used branch-and-bound whenever the resulting solution does not provide integer solution.

2. Problem statement

The following notations are used for the proposed model of this paper.

Counters
e	counter of the arcs in network $e \in E$
r	Counter of lines $r \in R$
a,b	Counters associated with the origin and destination of station $a,b \in V$

Sets
V	Sets of all stations
E	Sets of all arcs
R	Sets of all suggested lines

Parameters
$d_{a,b}$	Demand for trip from origin a to destination b

$u(e)$	Upper bound for the number of permitted trip on arc e
c	The capacity of train in terms on passengers

Decision variables

$y_{r,a,b}$	An integer variable determines number passengers directly traveled from station a to b by line r.
x_r	An integer variable indicating the number of trips on line r

The proposed model of this paper is formulated as follows.

$$\max \quad \sum_{a,b \in V} \sum_{e \in r | r \in R_{a,b}} y_{r,a,b} \tag{1}$$

$$\sum_{e \in r | r \in R_{a,b}} y_{r,a,b} \leq d_{a,b} \qquad \forall a,b \in V \tag{2}$$

$$\sum_{e \in r | r \in R_{a,b}} y_{r,a,b} \leq c \, x_r \qquad \forall r \in R, e \in r \tag{3}$$

$$\sum_{e \in r | r \in R_{a,b}} x_r \leq u(e) \qquad \forall e \in E \tag{4}$$

$$x_r \geq 0 \ \& \ \text{Integer}, \ y_{r,a,b} \geq 0 \ \& \ \text{Integer} \qquad \forall r \in R, \forall a,b \in V \tag{5}$$

The objective function in Eq. (1) maximizes the number of passengers directly traveled from an origin to a destination. Eq. (2) specifies that total number of passengers travel between each origin and destination must be less than the total demand. Eq. (3) specifies the limit of the capacity of train. Eq. (4) assigns an upper limit on the number of trips and finally Eq. (5) specifies the type of each variable.

3. The proposed model

The proposed model of this paper has some integer variables and there are some direct and indirect solution strategies to solve the resulted problem (Goossens et al., 2004; Barnhart & Laporte, 2007). One alternative solution is to use column generation technique to solve such problem. This method was originally used to solve large scale linear programming problems and recently there are varieties of this kind of methods for different railroad applications (Goossens et al., 2006; Brønmo et al., 2010; Dezső, 2010; Rönnberg & Larsson, 2010; Oppen et al., 2010). In column generation method, a problem is divided into two main and sub-problem and we do not handle a big size problem in each iteration. At each iteration of the column generation method, we use dual variables of the original problem with new variables of the sub-problem. Since this type of solution procedure is designed to solve large-scale linear programming problems, there is no guarantee that the final solution would be integer. The proposed model of this paper provides a heuristic method, which could help obtain integer solutions from column generations technique. Fig. 1 shows details of relationship between the sub-problem and the main problem.
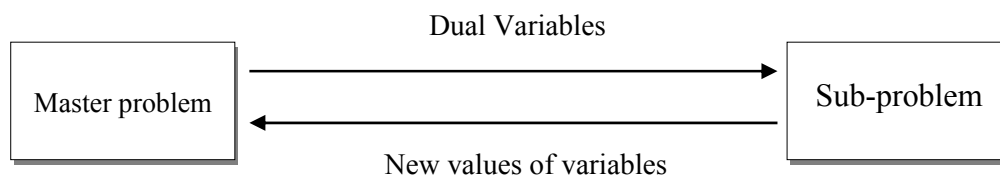

Fig. 1. Procedure of column generation method

As we can observe from Fig 1, in each iteration of the proposed model dual variables of the main problem are used by the sub-problem. In column generation method, we can use Eq. (3) in the main problem and by defining two sub-problems and try to satisfy other constraints through these sub-

problems. Therefore, we solve three problems, which increase the speed of the algorithm for large scale problems. Eq. (6) to Eq. (10) summarize the master problem.

Master Problem:

$$\min \quad \sum_{j}(-\sum_{a,b\in V}\sum_{e\in r|r\in R_{a,b}} y_{r,a,b})\lambda_j \tag{6}$$

$$\sum_{j}((\sum_{e\in r|r\in R_{a,b}} y_{r,a,b})\lambda_j + (-c\ x_r)\gamma_j) + s = 0 \qquad \forall r\in R, e\in r \tag{7}$$

$$\sum_{j}\lambda_j = 1 \tag{8}$$

$$\sum_{j}\gamma_j = 1 \tag{9}$$

$$\lambda_j \geq 0, \quad \gamma_j \geq 0, \quad s \geq 0 \tag{10}$$

In this model, the objective function is the same as the original one, λ and γ are decision variables and Eq. (7) is the same as capacity constraint of train with the difference that we added a non-negative slack variable to change it into equality constraint. Eq. (8) is the convexity constraint for the first sub-problem and Eq. (9) is the convexity constraint for the second sub-problem, and finally constraint set (10) is associated with the non-negativity of variables. Vectors x and y are inputs of the master problem from the sub-problem and dual variables of Eq. (7), Eq. (8) and Eq. (9) are w, α and β, respectively. The first sub-problem is formulated as follows,

Sub problem 1:

$$\max \quad ((w\times A)+1)y + \alpha \tag{11}$$

$$\sum_{e\in r|r\in R_{a,b}} y_{r,a,b} \leq d_{a,b} \qquad \forall a,b\in V \tag{12}$$

$$y \geq 0 \tag{13}$$

The objective function of the first sub-problem determines optimal values of variable y for the main problem. As we explained earlier, w is the dual variables for Eq. (7) and A are coefficients of variables y in Eq. (7). Eq. (13) determines decision variables for the first sub-problems. Since the nature of the sub-problem needs to maintain convexity form of the optimization, we cannot specify integer for variables. On the other hand, the variables represent the number of passengers travelled and normally it is a big number so rounding the number does not change the nature of the problem. The second sub-problem determines the optimal values of variables x, which is as follows,

Sub Problem 2:

$$\max \quad (W\times A)x + \beta \tag{14}$$

$$\sum_{e\in r|r\in R_{a,b}} x_r \leq u(e) \qquad \forall e\in E \tag{15}$$

$$x \geq 0 \tag{16}$$

Again, the optimal solution needs to be integer but since the variables determine the number of dispatches we can use imply the integer value using the heuristic approach method we present. The termination criteria is also written as follows,

$$\max \quad (W\times A)x + \beta \leq 0 \quad \& \quad \max \quad ((w\times A)+1)y + \alpha \leq 0 \tag{17}$$

According to Eq. (17), the column generation method terminates one the optimal values of both sub-problems are less than equal zero. The optimal results are then entered into the heuristic approach to

achieve integer solutions. The heuristic approach rounds the sum of dispatched passengers on each arc and difference is calculated and based on the resulted difference we add or subtract some arcs.
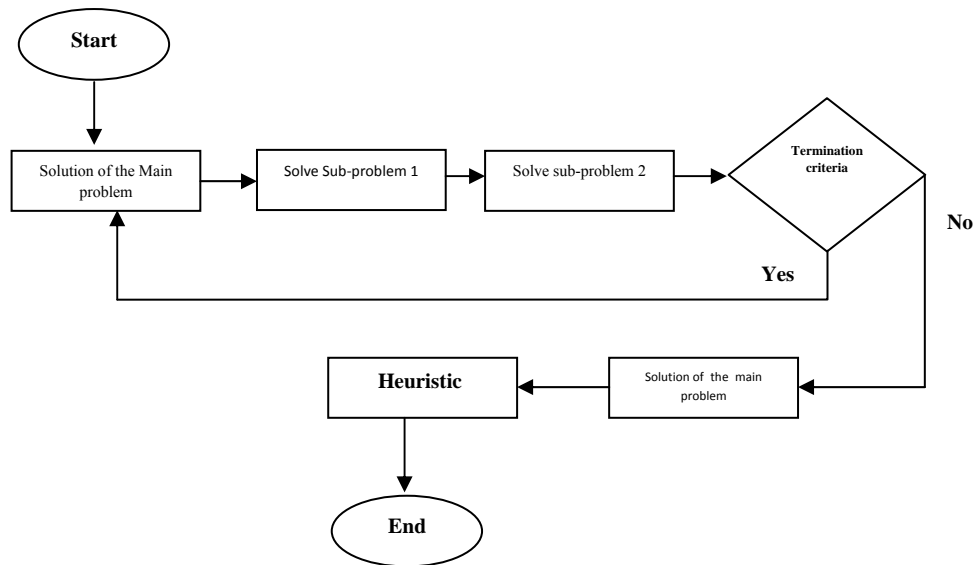

Fig. 2. Details of the proposed method

4. Results

In this section, we present details of the implementation of the proposed model. In our implementation, all optimization codes were performed using GAMS software package and we have setup 16 sample problems in different sizes given in Table 1. All problems are solved using direct implementation of mixed integer programming and the proposed model of this paper. The numbers of stations are set to 5, 10, 15 and 20. The suggested lines are also recommended in two different conditions. In the first case, F, we recommend new lines within the existing stations. For the second case, 50% of the suggested lines in the first case are suggested as the new lines. Of course, these lines are designed so that we could directly transfer all passengers. The suggested lines are denoted by H and we have considered two values of low and high for $u(e)$ denoted by T and L, respectively.

Table 1
The performance of the proposed method versus direct mixed integer problem

Example	Number of stations	Suggested lines	Capacity of each arc	GAMS CPU solution in sec.	Objective function	Proposed method CPU in sec.	Objective function
1	5	F	T	2.1	330	4.1	330
2			L	0.1	500	2.1	500
3		H	T	.60	330	8.0	330
4			L	5.0	500	7.0	500
5	10	F	T	2.420	1540	7.64	1540
6			L	8.378	2250	5.58	2250
7		H	T	8.45	1540	3.10	1540
8			L	1.42	2250	6.9	2250
9	15	F	T	.67098	3170	1.735	3170
10			L	.16543	5250	2.673	5250
11		H	T	.1832	3170	8.112	3170
12			L	8.749	5250	.5992	5250
13	20	F	T	-	-	9.2348	5990
14			L	-	-	.62153	9500
15		H	T	7.6834	5990	.6574	5990
16			L	4.6151	9500	4.502	9500

As we can observe from the implementation of the proposed model of this paper, both methods could reach optimal solutions for small scale problems. However, as the size of the problem increases, the proposed can reach final solutions in all cases and in reasonable amount of times.

5. Conclusion

In this paper, we have proposed a column generation method to decompose a large-scale railroad passenger-scheduling problem into some smaller scale problems, which are easier to solve. The primary concern with the resulted problem was that final solutions of the method needed to be integer and this was in contrast with convexity assumption of column generation techniques. We proposed a heuristic method to handle this problem and applied the proposed model for some examples. The preliminary results showed that the proposed model of this paper could provide optimal solutions for small-scale problems and it could reach some reasonable solutions for larger problems when direct implementation fails to do in reasonable amount of time.

References

Barnhart,C. & Laporte, G. (2007). *Handbook in OR & MS*. Vol. 14, Elsevier B.V.

Bussieck, M. R., Kreuzer, P., & Zimmermann, U. T. (1996). Optimal lines for railway systems. *European Journal of Operational Research*, 96, 54–63.

Brønmo, G., Nygreen, B., & Lysgaard, J. (2010). Column generation approaches to ship scheduling with flexible cargo sizes. *European Journal of Operational Research*, 200(1), 139-150.

Burdett, R. L., & Kozan, E. (2009). Techniques for inserting additional trains into existing timetables. *Transportation Research, Methodological*, 43(8-9), 821-836.

Claessens, M. T., van Dijk, N. M., & Zwaneveld, P. J. (1998). Cost optimal allocation of passenger lines. *European Journal of Operational Research*, 110, 474–489.

Chung, J., Moo, S., & Choi, I. (2009). A hybrid genetic algorithm for train sequencing in the Korean railway. *Omega*, 37(3), 55-65.

D'Ariano, A., Pacciarelli, D., & Pranzo, M. (2007). A branch and bound algorithm for scheduling trains in a railway network. *European Journal of Operational Research*, 183(2), 643-657.

Dezső, B., Jüttner, A., & Kovács, J. (2010). Column generation method for an Agent Scheduling Problem. Electronic Notes in Discrete Mathematics, 36, 829-839.

Ghoseiri, K., & Ghannadpour, F. (2010). A hybrid genetic algorithm for multi-depot homogenous locomotive assignment with time windows. *Applied Soft Computing*, 10(1), 53-65.

Goossens, J. W., Hoesel, S. V., & Kroon, L. (2006). On solving multi-type railway line planning problem. *European Journal of Operational Research*, 168(2), 403-424.

Goossens, J. H. M., van Hoesel, C. P. M., & Kroon, L. G. (2004). A branch-and-cut approach for solving railway line-planning problems. *Transportation Science*, 38, 379–393.

Lindner, T. (2000). Train schedule optimization in public rail transport. PhD thesis, TU Braunschweig.

Liu, S.Q., & Kozan, E. (2009). Scheduling trains as a blocking parallel-machine job shop scheduling problem. Computers & Operations Research, 36(10), 2840-2852.

Lee, Y., & Chen, C. (2009). A heuristic for the train pathing and timetabling problem. *Transportation Research, Methodological*, 43(8-9), 837-851.

Mesquita, M., & Paias, A. (2008). Set partitioning/covering-based approaches for the integrated vehicle and crew scheduling problem. *Computers & Operations Research*, 35(5) 1562-1575.

Oppen, J., Løkketangen, A., & Desrosiers, J. (2010). Solving a rich vehicle routing and inventory problem using column generation. *Computers & Operations Research*, 37(7), 1308-1317.

Peeters, M., & Kroon, L. (2008). Circulation of railway rolling stock: a branch-and-price approach. *Computers & Operations Research*, 35(2), 538-556.

Rönnberg, E., & Larsson, T. (2010). Column generation in the integral simplex method. *European Journal of Operational Research*, 129(1), 333-342.

Scholl, S. (2005). Customer-oriented line planning. PhD thesis, University of Kaiserslautern.

Suryani, E., Chou, S., & Chen, C. H. (2010). Air passenger demand forecasting and passenger terminal capacity expansion: A system dynamics framework. *Expert Systems with Applications*, 37(3), 2324-2339.

Tsai, T. H., Lee, C. K., & Wei, C. H. (2009). Neural network based temporal feature models for short-term railway passenger demand forecasting. *Expert Systems with Applications*, 36(2), 3728-3736.