Proceedings of the International MultiConference of Engineers and Computer Scientists 2010 Vol III,
IMECS 2010, March 17 - 19, 2010, Hong Kong

On the Railway Line Planning Models Considering the Various Halting Patterns*

Bum Hwan Park† Chung-Soo Kim‡ Hag-Lae Rho§

Abstract—**We propose railway line planning optimization models that determines the frequency of each train so that it can meet passenger origin-destination demands while minimizing the related costs. Most line planning models assume that all trains on a supply network stop at all intermediate stations. However to minimize passenger travelling time and provide train service to as many stations as possible, we must consider various halting patterns, that might be given in advance or be sought. Our study considers two line planning problems considering halting patterns, describes the computational complexities for each problem and present the optimization models for both. We also present experimental results obtained for Korean high speed railway network.**

Keywords: railway, line planning, halting pattern, NP-hard

1 Introduction

Since the advent of KTX(Korea Train eXpress) service in 2004, KORAIL(Korea Railroads, Corp.) has been planning to extend the railway network serviced by high-speed trains, and to develop a new high-speed rolling stock, e.g., KTX II which is faster but with a smaller number of seats or HANVIT with tilting function. Presently, KORAIL runs their high-speed trains partly sharing the railway network with conventional trains, *e.g.* Mugunghwa and Saemaul trains, which are much slower.

In these circumstances, an important issue is how many trains are needed to satisfy passenger OD demands for each type of train. This problem is known as *line planning*. In line planning one determines the frequencies of trains on a line to meet the OD demand while minimizing the related costs [3]. However most railway networks are too large to solve line planning model including all train types and such a model would be too complex to reflect the routes passengers would choose among various possibilities, including transfers between train-types. Hence, in most related studies[4, 5, 6], line planning is performed on the *supply network* which is a kind of a logical network comprised of only the links serviced by a single train-type. The network can be derived from the "system split" procedure [2] which is a method for distributing passenger OD demands over an entire network *a priori*. This distribution is based on a value assigned to a route which is primarily determined from the total travelling time and the number of transfers of a route. Thus, using a system split procedure, we can assume that the train-type demands of a specific supply network are given.

We also assumed that the OD demands are given for a specific train-type. By *train-type*, we mean a train service that can be classified by travelling time. For example, train-types in Korea are high-speed (KTX, KTX II), middle-speed(HANVIT) and low-speed(Mugunghwa, Saemaul) trains.

Our contribution in this research is the introduction of various halting patterns into a line planning model within a supply network. This allows trains with the same rolling stock and the same route to have different halting patterns. Presently, the KTXs are operated with various halting patterns between their two end stations. By considering various halting patterns, we expect that a line planning model will provide high-speed train services to as many stations as possible with faster travel times.

In this paper, we describe the optimization models and computational complexities of line planning problems considering various halting patterns. We summarize previous studies in Section 2, and we define our problems, the NP-hardness proofs and the optimization models in Section 3. Our experimental results obtained for the Korean high-speed railway network are described in Section 4.

2 Previous Studies

From a survey of line planning models, previous studies can be summarized as shown in Table 1. The second and third column show the demand type and objective function used in the models, respectively. The fourth column, 'halting patterns' contains an 'O' if the various halting patterns within a supply network were considered, and contains an 'X', otherwise. The last column, 'system split' indicates whether the model uses the assumption of

*This work was supported by "the Railroad Technology Development Program(RTDP)", which is centered at Korea Railroad Research Institute(KRRI).
†Korea National Railroad College, beomi72@hanmail.net
‡Korea National Railroad College, bull6213@hanmail.net
§Korea Railroad Research Institute, hlrho@krri.re.kr

ISBN: 978-988-18210-5-8
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

Table 1: Previous studies

study	input demand type	objective	halting patterns	system split
Bussieck *et al.* [4]	leg load, OD demand	direct travellers	X	O
Claessens *et al.* [6]	leg load	operation cost	X	O
Bussieck *et al.* [5]	leg load	operation cost	X	O
Borndörfer *et al.*[1]	OD demand	operation cost + travelling time	X	X
Goossens *et al.* [9]	leg load	operation cost	X	O
Goossens *et al.* [8]	OD demand	operation cost	X	X
Our study	OD demand	operation cost + travelling time	O	O

a system split.

Most of the studies[4, 5, 6, 9] except [1, 8] are based on *demand-covering model*[10], which find the optimized frequencies so that leg traffic load should be covered. Although not shown in Table 1, complexity analyses for various demand-covering models in path and tree networks are given in [10].

Borndörfer *et al.*[1] and Goosesens *et al.*[8] used OD demand directly to find passenger routes in a heterogeneous network, *i.e.*, they did not assume a system split. However, their models have a shortcoming: the systematically optimized passenger flow differs from passengers' actual choices among the routes. The systematically optimized passenger flow can be unreasonable, especially when various train-types are run on the same route, as in Korea, because the optimized flow might result in all or nothing traffic assignment to each train-type.

Our study is different from the previous ones in that it assumes both a system split and various halting patterns. two cases of halting patterns are considered, depending on whether the halting patterns are given in advance or whether they are to be found in the model. In next section, we define and analyze the models for these two cases. Our models are, to our knowledge, the first reported ones generated with such considerations.

3 Problem Definition and Complexity

We assume that OD demands for a specific train-type are given by a system split, and that there is no passenger transfer between different lines. In this section, we define our two line planning models, based on whether the halting patterns are defined in advance, and prove the NP-hardness of each problem. Most of the complexity analysis in previous studies have been based on demand-covering models. To our knowledge, the following proofs are the first reported for a line planning model considering OD demand directly.

From here on, we use the term *line* as a quadruple(s, e, H, c) where s, e, H, and c represent the start-station, end-station, set of stop stations and seating capacity, respectively.

3.1 Not given halting patterns

First we deal with the problem without prescribed halting patterns(LPWPHP) to find the frequency and halting pattern simultaneously as the problem is solved.

Definition 3.1 *Line planning problem without prescribed halting patterns(LPWPHP)* : to find the frequency and halting pattern of each line simultaneously, minimizing the sum of the total operation cost and total passenger travel time

We prove that LPWPHP is NP-hard by reduction from the PARTITION problem.

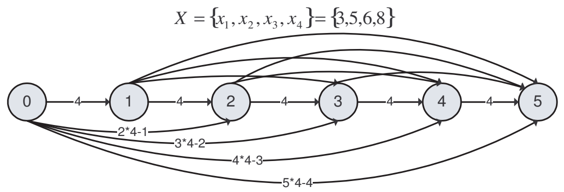

Figure 1: NP-hardness proof of LPWPHP

Theorem 3.2 *LPWPHP is NP-hard*

Proof. It will be shown that PARTITION problem[7] can be polynomially reduced to LPWPHP. Let an instance of PARTITION consist of a set $X = \{x_1, x_2, \cdots, x_n\}$ with $\sum_i x_i = A$. For the construction of the LPWPHP instance, we consider a directed graph where the set of nodes is $V = \{0, 1, 2, \cdots, n, n+1\}$ and the set of arcs is $\{(i,j)|\ i, j \in V\}$ as depicted in Figure 1. In this graph, the arc (i, j) represents non-stop travel from i to j. Every train runs from 0 to $n+1$ with a sufficient large operation cost M and the seating capacity $A/2$. And non-stop travel time from i to j is $n(j-i) - (j-i-1)$ which means that whenever a train

ISBN: 978-988-18210-5-8
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

does not stop at a station, the travelling time decreases by 1. The demand D^{od} from o to d is as follows:

$$D^{od} = \begin{cases} x_o, & o \in \{1, 2, \cdots, n\}, d = n+1 \\ x_d, & o = 0, d \in \{1, 2, \cdots, n\} \\ 0, & o/w \end{cases}$$

We will show that partition (S, S') exists if and only if LPWPHP has an optimal solution with the objective value $2M + n(2n+1)\frac{A}{2}$. First, suppose a partition (S, S') exists. From the partition, we can readily construct line plan with two trains t_1, t_2, each of which has a halting pattern such that each train halts only at stations corresponding S or S', respectively. Each train carries x_i from 0 to i and x_i from i to $n+1$ if the train only stops at i. Note that the total passenger travelling time of demand from 0 to i and from i to $n+1$ on train t_k is $n(n+1)x_i - \delta_k x_i$ where δ_k is the number of non-stop stations for train t_k. Thus, the total passenger travel time for the two trains is $n(n+1)\frac{A}{2} \times 2 - n\frac{A}{2}$ because the total number of non-stop stations for trains t_1 and t_2 is n, *i.e.*, $\sum_{k=1}^{2} \delta_k = n$.

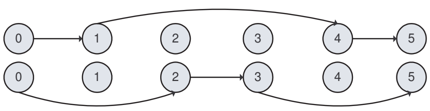

Figure 2: Halting patterns corresponding to partition $(\{x_1, x_4\}, \{x_2, x_3\})$

To prove the converse, suppose that a partition (S, S') does not exist. Note that the necessary number of trains is two so as to meet all OD demands with minimum operation cost. However, because a partition does not exist, at least one OD demand must be split between the two trains. Thus, both trains should stop at a station. This results in fewer non-stop stations than n, *i.e.*, $\sum_{k=1}^{2} \delta_k < n$, so that the total passenger travel time is greater than $(n+1)\frac{nA}{2} \times 2 - \frac{nA}{2}$. Here, M can be set to $n(2n+1)\frac{A}{2} + 1$. This completes the proof. □

3.2 Given halting patterns

Different from LPWPHP, the following problem assumes that the halting patterns are given in advance. This problem appears easier than LPWPHP, but it is still NP-hard.

Definition 3.3 *Line planning problem given halting patterns(LPGHP)* : Given halting patterns, to find the frequency of each line minimizing the sum of the total operation cost and total passenger travel time

We prove that LPGHP is NP-hard by reduction from the exact cover by three sets.

Theorem 3.4 *LPGHP is NP-hard*

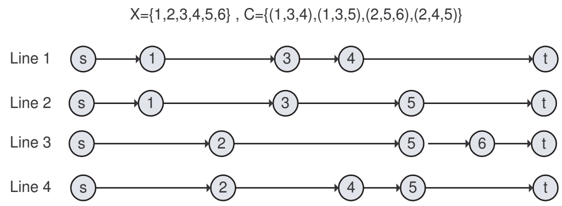

Figure 3: NP-hardness proof of LPGHP

Proof. It will be shown that the exact cover by three sets(X3C)[7] can be polynomially reduced to LPGHP. Let an instance of X3C consist of a set X with $|X| = 3q$ and a collection C of three-element subsets of X. For the reduction, we construct a directed path for all $c = (x_i, x_j, x_k) \in C$ where the set of arcs is $\{(s, x_i), (x_i, x_j), (x_j, x_k), (x_k, t)\}$, as depicted in Figure 3. Each path corresponds to a line for which the operation cost is 1. All arcs have capacity 3 and travel time 0. The demands between two nodes are all 0 except that the demands from s and towards t are all 1. Note that each line can meet the demand of at most 6 OD pairs among totally $6q$ OD pairs with non-zero demand; therefore, the number of lines satisfying OD demands is at least q. This means that if there exists an exact cover C', LPGHP instance has an optimal solution with objective value q and vice versa. □

As described in the above proof, NP-hardness is still valid when the travelling times are all set to zero.

Corollary 3.5 *Even when the objective function of LPGHP is comprised of only operation cost, LPGHP is NP-hard.*

4 Optimization Models

First, we present an optimization model for LPWPHP. The model is based on the multi-commodity flow model in which there are commodities corresponding to the halting pattern of the kth line and commodities used to find the passenger route of the od demand pair.

- A : the set of physical links in the supply network
- $L = \{1, 2, \cdots, K\}$: the set of lines indexed by number k. We assume that K is polynomially bounded in input size
- S : the set of OD pairs indexed by od.
- A_k : the set of arcs representing non-stop travel that line k can traverse
- c_k : operation cost of line k
- t_{ij} : non-stop travelling time between i and j.

ISBN: 978-988-18210-5-8
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

- t_{od}^k : travel time between o and d for line k
- C_k : seating capacity of line k
- $A_k(ij)$: the set of OD pairs traversing the non-stop link ij for line k
- D^{od} : demand between o and d.
- A_{uv} : the set of lines traversing the physical link uv.
- L_{uv} : line capacity of physical link uv.

Decision Variables

- f_k : frequency of line k
- x^k : $x^k = 1$ if the kth line is used; $x^k = 0$ otherwise
- x_{ij}^k : $x_{ij}^k = 1$ if the kth line traverses arc ij, $x_{ij}^k = 0$ otherwise
- λ_{ijk}^{od} : od passenger flow on arc ij of kth line.
- λ_k^{od} : od passenger flow on kth line.

Formulation LPWPHP

$$\min \sum_{k \in L} c_k f_k + \sum_{k \in L} \sum_{od \in S} \sum_{ij \in A_k} t_{ij} \lambda_{ijk}^{od} \quad (1)$$

$$\text{s.t.} \sum_{j: ij \in A_k} x_{ij}^k - \sum_{j: ji \in A_k} x_{ji}^k = \begin{cases} x^k & i = s(k) \\ -x^k & i = t(k) \\ 0 & o/w \end{cases}$$
$$\forall k \in L \quad (2)$$

$$\sum_{j: ij \in A_k} \lambda_{ijk}^{od} - \sum_{j: ji \in A_k} \lambda_{jik}^{od} = \begin{cases} \lambda_k^{od} & i = o \\ -\lambda_k^{od} & i = d \\ 0 & o/w \end{cases}$$
$$\forall od \in S, \forall k \in L \quad (3)$$

$$\sum_{od \in A_k(ij)} \lambda_{ijk}^{od} \le C_k f_k, \ \forall ij \in A_k, \forall k \in L. \quad (4)$$

$$f_k \le M x^k, \ \forall k \in L \quad (5)$$

$$\sum_{od \in A_k(ij)} \lambda_{ijk}^{od} \le M x_{ij}^k, \ \forall ij \in A_k \forall k \in L \quad (6)$$

$$\sum_k \lambda_k^{od} = D^{od}, \ \forall od \in S \quad (7)$$

$$\sum_{k \in A_{uv}} f_k \le L_{uv}, \ \forall uv \in A \quad (8)$$

$$x_{ij}^k, x^k \in \{0,1\}, others \ge 0, integer. \quad (9)$$

By (1), the model minimizes a combination of total operating cost and total passenger travel time. This objective function enables us to minimize the total passenger travel time among line plans with the minimum total operation cost by taking the product of a given weight and the first term in the objective function. In our experiment, a sufficiently large weight was used. Here, (2) and (3) indicate flow conservation of the train and passenger, respectively and (4) indicates that the sum of OD demand traversing ij link within the kth line must be no more than the total seating capacity of the kth line which is the frequency times seating capacity. Also, (5) and (6) enable the variables to be consistent. The sum over all k of od demands is equal to D^{od} by (7).

The mathematical formulation for LPGHP can be readily derived from the above formulation because LPGHP is the case when x_{ij}^k, x^k are given in the LPWPHP problem. Hence the formulation is much simpler.

Formulation LPGHP

$$\min \sum_{k \in L} c_k f_k + \sum_{k \in L} \sum_{od \in S} t_{od}^k \lambda_k^{od} \quad (10)$$

$$\text{s.t.} \sum_{od \in A_k(ij)} \lambda_k^{od} \le C_k f_k, \ \forall ij \in A_k, \forall k \in L \quad (11)$$

$$\sum_{k \in L} \lambda_k^{od} = D^{od}, \forall od \in S \quad (12)$$

$$\sum_{k \in A_{uv}} f_k \le L_{uv}, \ \forall uv \in A \quad (13)$$

$$\lambda_k^{od}, f_k \ge 0, integer \quad (14)$$

5 Experimental Results

In the section, we give the results obtained for the Korean high-speed railway network assuming various scenarios that exist in the actual network.

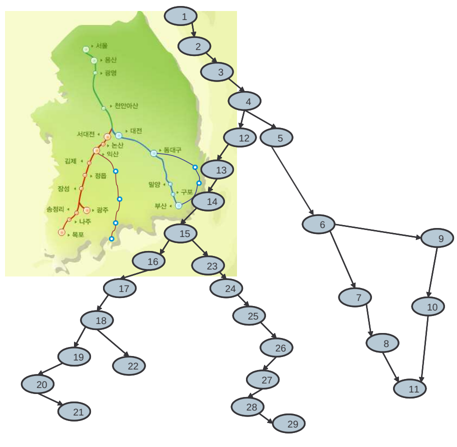

Figure 4: Korean railway network expected to be serviced by high-speed trains in 2011.

Figure 4 shows the railway network expected to be serviced by high-speed trains in 2011. Currently, route 6-9-10-11 is under construction ; tains will be able to run faster on this route than on 6-7-8-11 because the tracks are exclusive for high-speed trains. Route 15-23- $\cdots$ -29 is used only by conventional trains, currently not used by any high-speed train. From 2011, new high-speed rolling

ISBN: 978-988-18210-5-8
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

Table 2: Experimental results - Scenario 1

rolling stock	line route	LPGHP # of patterns (# of trains)	LPWPHP # of patterns (# of trains)	Current(2007) # of patterns (# of trains)
KTX	1-11 (via 7)	8(36)	6(36)	15(47)
KTX	1-6	2(22)	3(22)	2(9)
KTX	1-5	1(2)	1(2)	1(1)
KTX	2-21	2(4)	1(4)	6(8)
KTX	2-22	4(9)	3(8)	9(10)
computation time(s)		5.16*	792.75**	n/a
obj. value(10^6)		2628.6	2593.5	2859.5
travelling time(min)		6,455,776	6,128,284	6,526,909
Total number of patterns		17	14	33
Total number of trains		73	72	75

*,** : the number of branching nodes is restricted to 10,000. * : MIP gap = 1.76%, ** : MIP gap = 47.39%,

stock KTX II will be introduced, which has a lower operation cost than KTX because a KTX II train consists of 10 cars (seating capacity 544), while a KTX train consists of 20 cars (seating capacity 935). In our experiment, we assume that the operation cost is proportional to the driving distance, and the operation cost of KTX II is 0.6 times as large as that of KTX. The number of candidate patterns used in the LPGHP model is the same as that of the actual pattern currently serviced by KTX, which is shown in the last column of Table 2. The maximum number of generated patterns is restricted to six for the routes 1 to 11 and to four in the others. To obtain quickly a feasible solution to LPWPHP, some prominent halting patterns used to reduce the travelling time were fixed in advance.

We generate a new optimized line plan for the high-speed railway network under three scenarios. Table 2∼4 list the results from these scenarios, which are as follows: only KTX is used with current high-speed railway network(Scenario 1), only KTX is used with the Figure 4 network(Scenario 2) and both KTX and KTX II are used with the Figure 4 network(Scenario 3). For Scenario 1, we compared the solution with the current line plan. The third and fourth column in each table indicate the number of generated patterns and the number of trains for LPGHP and LPWPHP, respectively. The formulations LPGHP and LPWPHP were implemented on a PC with an Intel Core Duo 2.20-GHz CPU using CPLEX 10.2.

Table 2 shows how much our line planning model could improve the current line plan. The total number of trains was decreased to 73 in the LPGHP and to 72 in the LPWPHP case, from 75. Trains with non-stop pattern are rare in the current line plan. However in the LPWPHP cases, there were many trains with non-stop pattern which result in reduced passenger travel time. Although there were many trains with non-stop patterns, the number of patterns was much lower than that of the current schedule. Lower number of patterns enables us more easily to develop train operation plan including timetable. Obviously a decrease in the total number of trains indicates lower operation cost.

The objective values of LPWPHP were not much better than those of the LPGHP case because both models minimized the total operation cost first by weighting a large number to the operation cost. However, in all scenarios, the total passenger travel time was decreased in LPWPHP model because it generated optimized patterns to do exactly that. We expect that the LPWPHP model can play an important role in developing an optimized line plan with various halting patterns when there are no halting patterns to refer to, such as when new train service is launched on a new railway network. The total passenger travel time in the LPWPHP model was significantly decreased by 4 ∼ 5 % in all scenarios, compared with the LPGHP model.

6 Concluding Remarks

We present complexity analyses and optimization models for a railway line planning model considering the halting patterns of trains. We ran the models on a comparatively small network, serviced by high-speed trains. Although the model was implemented for a small network, the computation time and the MIP gap of both LPGHP and LPWPHP are tremendous as shown in Table 2∼4. To apply these models to larger networks, it will be necessary to develop a more efficient algorithm that can guarantee a tighter MIP gap and a reduced computation time, such as a branch-and-cut or column-generation algorithm.

References

[1] R. Borndörfer, M. Grötschel, and M. E. Pfetsch. A column generation approach to line planning in public transport. *Transportation Science*, 41(1):123–132, 2007.

Table 3: Experimental results- Scenario 2

rolling stock	line route	LPGHP # of patterns (# of trains)	LPWPHP # of patterns (# of trains)
KTX	1-11 (via 7)	5(9)	3(9)
KTX	1-11 (via 9)	4(29)	6(29)
KTX	1-6	2(21)	3(21)
KTX	1-5	1(2)	1(2)
KTX	2-21	2(4)	1(4)
KTX	2-22	3(6)	2(6)
KTX	2-29	2(7)	3(7)
computation time(s)		8.81*	2991.39**
obj. value(10^6)		2874.1	2873.8
travelling time(min)		6,410,163	6,058,557
Total number of patterns		19	19
Total number of trains		78	78

*,** : the number of branching nodes is restricted to 10,000. * : MIP gap = 1.16%, ** : MIP gap = 78.83 %

Table 4: Experimental results- Scenario 3

rolling stock	line route	LPGHP # of patterns (# of trains)	LPWPHP # of patterns (# of trains)
KTX / KTX II	1-11 (via 7)	6(11) / 0(0)	4(12) / 0(0)
KTX / KTX II	1-11 (via 9)	3(27) / 0(0)	5(26) / 0(0)
KTX / KTX II	1-6	1(20) / 1(1)	4(20) / 1(1)
KTX / KTX II	1-5	1(2) / 1(2)	1(2) / 0(0)
KTX / KTX II	2-21	2(4) / 0(0)	3(4) / (0)
KTX / KTX II	2-22	4(6) / 0(0)	3(6) / 0(0)
KTX / KTX II	2-29	2(4) / 1(4)	2(4) / 2(4)
computation time(s)		9.48*	6331.89**
obj. value(10^6)		2837.0	2836.7
travelling time(min)		6,469,524	6,209,656
Total number of patterns		21	25
Total number of trains		79	79

*,** : the number of branching nodes is restricted to 10,000. * : MIP gap = 0.48%, **: MIP gap = 81.19%

[2] A. Bouma and C. Oltrogge. Linienplanung und simulation für öffentliche verkehrswege in praxis und theorie. *ETR Eisenbahntechnische Rundschau*, 43(6):369–378, 1994.

[3] M. R. Bussieck. *Optimal Lines in Public Rail Transport.* PhD thesis, TU Braunschweig, 1997.

[4] M. R. Bussieck, P. Kreuzer, and U. T. Zimmermann. Optimal lines for railway systems. *European Journal of Operations Research*, 96:54–63, 1996.

[5] M. R. Bussieck, T. Lindner, and M. E. Lübbecke. A fast algorithm for near optimal line plans. *Math. Methods Opers. Res.*, 59:205–220, 2004.

[6] M. T. Classens, N. M. van Dijk, and P. J. Zwaneveld. Cost optimal allocation of rail passenger lines. *European Journal of Operations Research*, 110:474–489, 1998.

[7] M. R. Garey and D. S. Johnson. *Computers and Interactability : A Guide to the Theory of NP-completeness.* W. H. Freeman and Company, San Francisco, 1976.

[8] J.-W. Goosens, S. V. Hoesel, and L. Kroon. On solving multi-type line plannning problems. *METEOR Research Memorandum RM/02/009, University of Maastricht*, 2002.

[9] J.-W. Goosens, S. V. Hoesel, and L. Kroon. A branch-and-cut approach for solving railway line-planning problems. *Transportation Science*, 38(3), 2004.

[10] L. M. Torres, R. Torres, R. Borndörfer, and M. E. Pfetsch. Line planning on paths and tree networks with applications to the quito trolebús system. *ZIB-Report 08-53*, 2008.