A Bilevel Model for Railway Train Set Organizing Optimization

Ya Gao[1] Jie Lu[1] Guangquan Zhang[1] Siwei Gao[2]

[1]Faculty of IT, University of Technology, Sydney, PO Box 123, NSW 2007, Australia
[2]Department of Traffic and Transportation, Southwest Jiaotong University, Emeishan 614202, China

Abstract

Train Set Organization (TSO) is to arrange the train set in railway freight transportation. Bilevel programming techniques were proposed to solve the Stackelberg game in which play is sequential and cooperation is not permitted. In this paper, an optimizing model for TSO is developed by the bilevel techniques. First, we analyzed the multiple level nature of management on TSO and simplified it into two levels. Then, a bilevel model for TSO was develop. Finally, this model was further illustrated by applying it on a railway station.

Keywords: Train set organization, Bilevel decision making, Optimization model

1. Introduction

Railway transportation, as one of the most important vehicles ways, has always been playing an irreplaceable role in social economics. For railway freight transportation, about 80% of the whole transportation time is allotted to the operations of loading/unloading, transferring, and overhauling in railway technical stations [1]. The working state of technical stations, therefore, will influence the whole overpass ability of the railway network. Thus the research on the railway transportation optimization will be bound to focus on the operation of technical stations. The main methodologies used include scheduling theory [1], graph theory [1, 2], mathematical programming [1, 3], and operational theory [1]. Also, some scholars have addressed the problem of traffic controlling from multiple levels' angle [4–7].

Bilevel programming techniques are mainly developed for solving decentralized management problems with decision makers in a hierarchical organization, the upper the leader while the lower the follower. Both the leader and the follower try to optimize their own objective functions and the corresponding decisions do not control but affect those of the other level [8]. Bilevel techniques have been applied with remarkable success in different domains such as decentralized resource planning, electronic power market, logistics, civil engineering, chemical engineering, road network management and risk management [9–15].

Most current research by bilevel techniques on traffic controlling centers on the transforation network design [5, 16] and layout [4, 6]. Little research has been conducted towards TSO problems from multiple levels' angle. In this paper, we used bilevel method to study the problem of TSO in the railway running and management. The rest of this paper is organized as: following the introduction in Section 1, Section 2 analyzed the bilevel nature from TSO. The model was then defined in Section 3. In Section 4, we took the operation in a railway station as an example to further illustrate this bilevel model. Finally, the paper was concluded in Section 5.

2. Problem analysis

The TSO, aiming at arranging the train set in railway freight transportation and with extraordinary professional and technical specialties, is one of the main subjects in railway transportation management. The objectives of TSO include: to make the transportation efficient and even; to use the transporting device reasonably; and to promote the cooperation among different departments involved in the freighting procedure. The term of "organizing" here means arranging, deciding and managing, while "train set organizing" acts to arrange the train set, make decisions on related issues, and manage the procedure in railway transportation.

There exist multiple levels among the running of TSO, i.e. the "national railway network level" the top, the "local bureau railway network level" the second, the "stations" the third, and the "operating group" the bottom. However, as the operating objects of both the national railway network and the local bureau railway network are train set,

while those of the two lower levels are trains, the organization of TSO can be generalized into two levels: the railway network the leader, and the stations the followers. Thus the bilevel programming technique can be used to analyze it.

The main concerns of the railway network are to decide the trains type (pick-up-and-drop-train, district-train, transit-train, or through-train), the trains constitution, the trains number, and the detailed route of the departing train set. The objectives of the railway network include: improving the transportation capacity and service speed, reducing the cost, balancing the working rhythm among divisions, and assigning the break-up and make-up jobs among different stations rationally.

The tasks assigned to the station are to constitute normative train set required by the railway network from all kinds of freight wagons stopped by this station. Involved with these tasks, there also include a series of relevant operations, such as: collecting or delivering, shunting, loading/unloading, and wagon checking. The main concerns of stations include: making the operating efficient, economical and safe; using the transportation devices such as track, shunting locomotive, and hump rationally; deciding the operation steps together with its schedules; and cooperating among steps within the schedule-frame of railway network.

Two levels though TSO can be divided into, the separate levels still share intrinsic consistency. For the upper level, when making a TSO plan, the railway network must consider the influence from the specific operating ability and device conditions of stations, while calculating the influence factors from itself such as the amount and destinations of trains and the track conditions; For stations located at the lower level, when implementing the working goals, they should try their best to harmonize between their own operation abilities and the working arrangement from their top counterpart.

The railway stations can be grouped into two classes: the “through stations" and the “technical stations". Compared with technical stations, through stations are small sized and their daily works, mainly on helping trains go through or two train set from opposite directions meet, are simple and the workload are small. Except for all the functions of through stations, technical stations are to make new train set by breaking up the old ones and adding transship trains and trains originated there. Tasks also include: arrival/departure operating, collection-and-delivering operating, shunting, loading/unloading, and wagon checking. We generalize these operations at technical stations as “shunting and transship operation".

For the reason of facilitating the modeling, we simplify the TSO by assuming that:

1) The railway transportation supply is less than the demand; the aim of TSO is to fully use the transportation ability to provide as much transportation as possible.
2) The topo structure of the railway network is a circle formed by train lines. This is to embody the continuous nature of the net and transportation circulation.
3) The main line is double-track with every track’s direction fixed, which means there allowed two train set running in reverse directions between two stations simultaneously. This is to avoid the meeting problem of two train set with opposite running directions.
4) Within the railway net, there located only technical stations, and runs only one type of trains, the “district trains", which are from a technical station A and to the other technical station B. Between A and B there is no other technical stations.
5) The unit workload of “shunting and transship operation" for all technical stations are the same. In other words, every technical station share the identical amount of operating time for the same train set.

Based on these assumptions above, the decision-maker on the railway network wishes that the density of train set (calculated by the time intervals between any two adjacent running train set) and the length of train set (the number of trains of any train set) as large as possible to obtain the maximal transport capacity. However, for the sake of safety, the density has its upper limit set by railway network. And restricted by the motive power of locomotive and the useful length limit of arrival-departure track, the train set length has its upper limit as well.

Ignoring the constraints by the railway network, the stations, on one hand, wish the length of train set large because the larger the length, the more efficient the operating and the lower the unit operating cost. The operating efficiency is the amount of trains shunted and transshipped per unit time, while the unit operating cost is the cost for every single train. On the other hand, the operating time for shunting and transshipping, which influences the cost, will increase if the length of train set increases. However, the overall effect of the trains set length is that the general unit op-

erating cost will decrease with the increase of the train set length.

>From the analysis above, we can include that, for the variable of the length of train set, the two levels share the same objective: the larger the length the better. For the variable of the density of train set, the decision makers at the upper level pursue its minimum while those at the lower level wish it change with the train set length in the same direction. Generally speaking, the shunting and transshipping time in stations is larger than the safe time intervals of any two adjacent running train set, so the variable of the density of train set is determined by the lower level, the stations, while the variable of the length of train set is controlled by the top level, the railway network.

3. Model building

Based on the analysis above, the bilevel model of TSO is built as:

For $x = (x_1, x_2, ..., x_n) \in X \subset R^n$, $y \in Y \subset R^n$, F, $f : X \times Y \to F(R)$,

Leader: Decision-maker of the railway network

$$\max_{x \in X} F(x, y) = \frac{a_1 \cdot \sum_{i=1}^{n} w_i \cdot x_i}{\sum_{i=1}^{n} w_i \cdot y_i} \tag{1a}$$

$$\text{s.t.} \quad \sum_{i=1}^{n} w_i \cdot x_i < m \tag{1b}$$

$$\sum_{i=1}^{n} w_i \cdot y_i > c_1 \tag{1c}$$

i-th follower: i-th technical station

$$\min_{y_i} f_i(x_i, y_i) = -b_1 \cdot x_i - b_2 \cdot y_i \tag{1d}$$

$$\text{s.t.} \quad c_2 \leq \frac{x_i}{y_i} \leq c_3 \tag{1e}$$

$$y_i > c_4 \tag{1f}$$

Explanation:

1) Variables:
   x_i: the length of train set for the i-th station, which is the number of trains of any train set controlled by the leader, the decision-maker of the railway network.
   y_i: the density of train set for the i-th station, which is the time interval between any two adjacent running train set, controlled by the i-th follower, the i-th technical station.
2) Coefficients and constants:
   n: the number of technical stations in the railway network
   w_i: the relative weight for the i-th station in the railway network.
   a_1: the time interval. If $a_1 = 24$, then $a_1 / \sum_{i=1}^{n} w_i \cdot y_i$ means the number of train set going through the network within 24 hours, $a_1 \sum_{i=1}^{n} w_i x_i / \sum_{i=1}^{n} w_i \cdot y_i$ is the number of trains going through the network per day, and $a_1 > 0$.
   m: the maximum number of trains of any train set regulated by the "Safety Terms". When the trains are empty, the main concern is not to exceed the length limit. When the trains are loaded, the weight limit becomes the decisive factor. However, for the sake of safety, when computing, both the length and weight must meet the requirements. No matter whether it is the weight or length, the ultimate limit is put on the number of trains.
   c_1: the minimum time interval between any trains list regulated by the "Safety Terms".
   b_1 and b_2: the weights set for the influencing power by the length and density to the unit cost.
   c_2 and c_3: the lower and upper number limits of the trains for technical stations to shunt and transship per time unit.
   c_4: the least time for the technical stations to complete the shunting and transshipping.
3) Formula:
   (1a) means the leader aims at obtaining the maximum throughput capacity within certain period of time. $\sum_{i=1}^{n} w_i x_i / \sum_{i=1}^{n} w_i \cdot y_i$ means the number of trains shunted and transshipped per time unit.
   (1b) means the length of train set has its upper limit imposed by the locomotive's motive power and the arrival-departure track's useful length. When the trains are loaded, except for the length limit, there is still weight restriction set upon the train set, which means, the weights of goods loaded together with the weights of trains can not exceed its upper limit.
   (1c) means any two adjacent running train set can not be too close for the sake of safety. c_1 is the minimum time interval between any trains list according to the "Safety Regulation".
   (1d) means the followers wish that the cost is as low as possible. The first part of (1d) means the more the length of trains set, the more efficient of the shunting, and the lower the unit cost. The second part means the longer the train set remains in the station, the higher the

cost.
(1e) means technical stations have their own lower and upper time limit to shunt and transship trains.
(1f) means there exists a least period of time for the technical station to complete the operation.

4. Case study

In this example, we take the railway freight operation in a railway station "Station A" into consideration. Station A is a technical railway station with the duty of managing both the passenger transportation and freight transportation within the precinct of its Railway Bureau. The data collected from Station A cover the duration between November 1, 2006 and December 31, 2006.

Suppose the trains shunted and transshipped are to the direction of Station B, which is another station located near Station A along its downlink. And the weight distribution of trains is listed in Table 1, with the locomotive is SS1(137 ton, 1.9 unit length).

Table 1: Train set distribution

WT	**WS (ton)**	**Load (ton)**	**%**	**EL**
B23	38	40	3	2.1
P64A	26	58	3	1.5
G70	23	58	9	1.1
G60	23	50	50	1.1
G70	23	55	35	1.1

The terms in Table1 are explained as:

WT wagon type, the type of wagon used.

WS wagon suttle, the weight of the empty wagon.

Load the weight of the goods loaded.

EL equivalent length, the equivalent length of a wagon is calculated from the front clasp to the rear clasp, with the unit length as 11 meters. If the equivalent length is "1.1", then its actual length is $11 \times 1.1 = 12.1$ meters.

According to the model defined by (1), the coefficients are calculated and discussed below:

- a_1: as the computation is within the "Basal Daily Working Plan", which is to arrange the wagons assignment and schedule necessary operations based on the "Trains Running Chart", 'Trains Shunting Plan", "Detailed Rules on Technical Station Management", and constraints set by operating spots, the computing of the freighting wagon organization is limited within a working day of 24 hours. So a_1 is set to 24.
- m: limited by the pulling ability of the locomotive and the territorial landform, such as grading, within Station A's precinct, the weight of train set must be not larger than 3500 tons; The departure track used for trains set to the direction to Station B is Track IV, Filed II, whose effective length is 890 meters. Deduced by 30 meters of braking distance, which is left for trains to stop safely, the maximum length for the trains set is 860 meters.
  Taking the constitution of the trains listed in Table 1, we set 1 "unit train" as a virtue train whose equivalent length, denoted by l_1 (meters), and weight, denoted by w_1 (ton), are calculated below:

$$\begin{aligned} l_1 &= 2.1 \times 0.03 + 1.5 \times 0.03 + 1.1 \times 0.09 \\ &\quad + 1.1 \times 0.5 + 1.1 \times 0.35 \\ &= 1.142 \\ w_1 &= (38 + 40) \times 0.03 + (26 + 58) \times 0.03 \\ &\quad + (23 + 58) \times 0.09 + (23 + 50) \times 0.5 \\ &\quad + (23 + 55) \times 0.35 \\ &= 66.95 \end{aligned}$$

  The maximum number of such empty "unit train", denoted by m_e, is $(860 - 1.9 \times 11)/(1.142 \times 11) = 66$, and the maximum number of such loaded "unit train", denoted by m_l, is $(3500 - 137)/66.95 = 50$.
  >From above analyzing and computing, we get $m = \min\{m_e, m_l\} = \min\{66, 50\} = 50$.
- c_1: for the sake of safety, the pursuing distance, the minimum distance interval between any adjacent running trains list, is 10 kilometers, which costs about 0.2 hours in the journey from Station A to Station B. So c_1 is set to 0.2.
- b_1 and b_2: we set the weights of length and density of trains set on the cost of the station as 0.4 and 0.6 respectively.
- c_2 and c_3: the least number of trains Station A can shunt and transship is 30 per hour, while the max number is 150.
- c_4: the least time for Station A to complete the shunting and transshipping for a train set is 0.68 hour.

Thus, the bilevel problem defined by (1) is spe-

cilized as (2) in UStation A:

Leader: Decision-maker of the railway network

$$\textbf{Objective:}\ \max_x F(x,y) = \frac{24x}{y} \tag{2a}$$

$$\text{s.t. } x < 50 \tag{2b}$$

$$y > 0.2 \tag{2c}$$

Follower: Urumchi West Station

$$\textbf{Objective:}\ \min_y f(x,y) = -0.4x - 0.6y \tag{2d}$$

$$\text{s.t. } 30 \le \frac{x}{y} \le 150 \tag{2e}$$

$$y > 0.68 \tag{2f}$$

As illustrated in Fig. 1, the triangle "ABC" depicts the constraint region for this example. The dotted lines with arrows of "F" and "f" represent the optimizing directions for the leader and follower respectively. The optimal solution to this example occurs at the point $(x^*, y^*) = (50, 1.67)$ with $F^* = 718.6$ and $f^* = -21.002$, which means, the railway network will obtain its maximum throughput capacity of 718.6 trains per day, if the decision makers of the railway network set the average number of train to 50, followed by Station A setting the time interval between every two adjacent train sets to 1.67 hour.

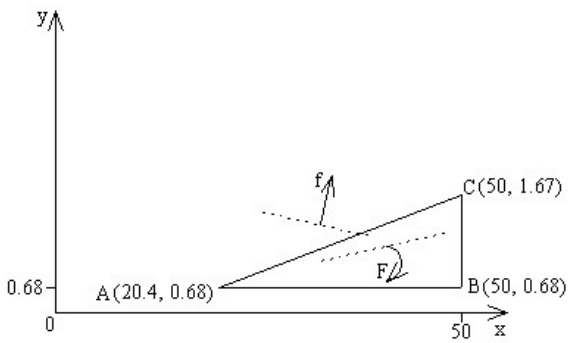

Fig. 1: Geometry of BLP.

5. Conclusions and future study

In this paper, the bilevel nature in train set organization has been first put forward by abstracting and simplifying railway trains management. First, the bilevel model is developed. Then, this model was applied to Station A for a real case study. The testing result obtained from Station A was reasonable and could be helpful to its trains organization. However, as a lot of practical details have to be ignored for articulating the model building, this model has its limitations when applied directly for TSO decision making. Future efforts will be focused on relating more practical and randomly occurred issues from field work.

Acknowledgments

The work presented in this paper was supported by Australian Research Council (ARC) under discovery grants DP0557154 and DP0559213 and China Railway Ministry under Project 2003X033.

References

[1] W.Q. Li and W. Du. *Optimization Models and Algorithms for the Daily Working Plan of Railway Technical Station*. Southwest Jiaotong University Publisher, Chengdu, 2002.

[2] C.G. Wang. Research on the model and algorithm of dynamic wagon-flow allocating in a marshalling station. *Journal of the China Railway Society*, 26(1):1–6, 2004.

[3] M.H. Wang and Q. Zhao. Optimal model and algorithm of stage plan of intelligent dispatching system for marshalling stations. *Journal of the China Railway Society*, 27(6):1–9, 2005.

[4] C. Feng and C. Wen. Bi-level and multi-objective model to control traffic flow into the disaster area post earthquake. *Journal of the Eastern Asia Society for Transportation Studies*, 6:4253 – 4268, 2005.

[5] M. Berezinski, J. Holubiec, and G. Petriczek. Multilevel models for railway traffic control. In *Control'94*, pages 54–56, 1994.

[6] H.Q. Wang Y.P. Le, L.X. Miao and C. Wang. A bi-level programming model and a solution for public logistics terminals planning. In *TRB, 2006*, 2006.

[7] X.H. Li H.H. Mo F. Shi, Q. Fang and Yong Huang. A bi-level p rogramming optimization method for the layout of technical service station. *Journal of the China Railway Society*, (2):1–4, 2003.

[8] W. Candler and R. Townsley. A linear two-level programming problem. *Computers and Operations Research*, 9:59–76, 1982.

[9] E. Aiyoshi and K. Shimizu. Hierarchical decentralized systems and its new solution by a bar-

rier method. *IEEE Transactions on Systems, Man, and Cybernetics*, 11:444–449, 1981.

[10] J. Amat and B. McCarl. A representation and economic interpretation of a two-level programming problem. *Journal of the Operational Research Society*, 32:783–792, 1981.

[11] L. Leblanc and D. Boyce. A bilevel programming algorithm for exact solution of the network design problem with user-optimal flows. *Transportation Research*, 20:259–265, 1986.

[12] P. Marcotte. Network optimization with continuous control parameters. *Transportation Research*, 17:181–197, 1983.

[13] T. Miller, T. Friesz, and R. Tobin. Heuristic algorithms for delivered price spatially competitive network facility location problems. *Annals of Operations Research*, 34:177–202, 1992.

[14] G. Papavassilopoulosm. Algorithms for static stackelberg games with linear costs and polyhedral constraints. In *Proceedings of the 21st IEEE Conference on Decisions and Control*, pages 647–652, 1982.

[15] R. Stulz. Risk management & derivatives. Technical report, Mason: Thomson South-Western, Ohio, 2003.

[16] R. Song Sh.W. He and Q. Zhao. A study on dss of stage plan for railway marshalling station. *System Engineering*, 14(4):62–66, 1996.