Integrated Modelling and Optimization of Train Scheduling and Shunting at Complex Railway Passenger Stations

Xiangchun Qi
Transportation & Economics Research Institute, China Academy of Railway Sciences. No.2 Daliushu Road, Haidian District, Beijing, China. 100081
Phone: 8613801275552; Fax: 86-10-51683124
Email: qixiangchun@sina.com

Yixiang Yue (Corresponding author)
School of Traffic and Transportation, Beijing Jiaotong University, No.3 Shang Yuan Cun, Hai Dian District, Beijing, China. 100044
Phone: 8613511054454; Fax: 86-10-51683124
Email: yixiangyue@yahoo.com

Juntao Han
School of Traffic and Transportation, Beijing Jiaotong University, No.3 Shang Yuan Cun, Hai Dian District, Beijing, China. 100044
Phone: 8613810437903; Fax: 86-10-51683124
Email: 15120813@bjtu.edu.cn

Leishan Zhou
School of Traffic and Transportation, Beijing Jiaotong University, No.3 Shang Yuan Cun, Hai Dian District, Beijing, China. 100044
Phone: 8618601012782; Fax: 86-10-51683124
Email: lszhou@bjtu.edu.cn

Hesham A. Rakha
Charles E. Via, Jr. Department of Civil and Environmental Engineering
3500 Transportation Research Plaza, Blacksburg, VA 24061
Phone: (540) 231-1505 Fax: (540) 231-1555
Email: hrakha@vt.edu

Word count: 5,742 words text + (6 figures + 1 table) x 250 words = 7,492 words

Submission Date: July 30th, 2016

Paper submitted for peer review at the 96th Transportation Research Board Annual Meeting.
ABSTRACT

The operations quality of large train station is critical in determining the efficiency of entire railway network. Large railway stations often have more than two adjacent stations and have to address highly complex train operation patterns. To develop a more efficient operation plan than the existing route-based representations for modeling train conflicts, this paper develops a more systematic track-based resource network (TRN) representation for a railway station that simulates the fixed equipment; based on this high-fidelity TRN, we introduce a heuristic model for integrated scheduling of train operations including shunting movements. The mathematical model simplifies the assignment of station infrastructure devices to train operations. The model is first converted to its dual form, and then a Lagrangian relaxation (LR)—based heuristic is implemented to solve the problem. We have developed a software system to address the operation scheduling problem (OSP) based on this method, and the system has been tested on real-world instances of all stations at the Beijing-Shanghai high-speed railway line. The experimental details of a large-scale high-speed railway station are given. The results demonstrate the quality of the proposed method; the system yields a solution for large-scale terminal stations within a few minutes.

Keywords: Railway station, Train operation, Shunting, Routing, Scheduling, Lagrangian relaxation, Optimization
INTRODUCTION

Large stations or junctions are critical nodes in railway networks, and operations in these nodes are critical for the efficiency of the entire railway line or network. However, large stations usually have more than one adjacent station as well as highly complex train operation patterns. Operations in complex railway stations consist of regular train operations and shunting operations. Regular operations include trains running inbound and outbound and stopping at platforms. In addition to regular train service, some shunting operations are also required: after passengers leave the train, the train units may be shunted to the depot to have the passenger compartments cleaned, catering services performed, and trains turned around and then moved back to the platform for the next round of service. Many possible conflicts can affect the train scheduling process in stations. Dispatchers usually apply simple rules to resolve these conflicts, such as assigning routes to regular train operations first and delaying shunting operations or assigning routes to the trains according to their priority. Such rules are widely used in practice, but achieving the global optimal solutions in large, densely occupied terminal stations is difficult.

In real-world railway traffic management, the development of station operation plans can be divided into two stages. The first stage is conducted while the train timetable is scheduled. In this stage, technicians must draw up a draft train operation plan for each station. Typically, when deciding whether all trains can enter and exit a station, the technicians only consider track capacity. The second stage is conducted during daily operations in which dispatchers must perform all tasks and schedule operations according to the actual delays and the status of the devices. Station dispatchers prevent conflicts between the hundreds of daily trains entering and exiting the station on paths that intersect lines and platforms. The dispatchers also schedule thousands of operations to guarantee that each train has sufficient allowance for headway, dwell time, and travel time. The problem can be described as optimizing the usage of scarce resources for a large number of train operations without conflicts at the railway station. A plan without any “conflicts” is a feasible plan that can be performed by workers in different departments within the station. The detection and prediction of operation conflicts is currently performed by dispatchers, but having a Decision Support System (DSS) perform this job is a main concern of station dispatchers. We have implemented a DSS for China Railway Corporation (CRC) and this DSS has already been tested in more than ten stations and has been demonstrated to provide dispatchers with feasible train operation and shunting plans for stages 1 and 2. We introduce two models in this DSS: a track-based resource network (TRN) and an operation scheduling model (OSM). A TRN depicts a railway station’s infrastructure including tracks, routes, signals, and borders to detect conflicts in the operation plan simulation. Based on a microscopic assignment of infrastructure elements, the OSM formulates the optimal goal and constraints for the problem.

PROBLEM DESCRIPTION AND LITERATURE REVIEW

Before describing the problem, some operational examples are illustrated using a macroscopic schematic of train operations at the Beijing South Railway Station. This station is connected to four adjacent railway stations and a large maintenance depot and has four yards: a conventional train yard, a high-speed train yard, an inter-city train yard, and a temporary storage yard for inter-city trains. The first three train yards belong to the reception-departure yard. The conventional train yard, high-speed train yard, inter-city train yard, and temporary storage yard have 7, 12, 5, and 5 tracks, respectively. Conventional trains are pulled by one or more locomotives. The majority of them pass through the station and some others may proceed into the depot to receive necessary maintenance. Most high-speed trains must enter the depot, except for those trains that
return soon after stopping only at the high-speed yard. Most inter-city trains return to the inter-city train yard, but some of these trains need to go to the temporary storage yard to empty some tracks in the inter-city train yard for upcoming trains. All inter-city trains go to the depot at night for routine maintenance and leave from the depot the next morning. Figure 1 shows the operation process of trains at Beijing South Railway Station.

![Graphical representation of train operations at Beijing South Railway Station](image)

**FIGURE 1 Principal train operation processes of trains in Beijing South Railway Station**

We can categorize trains at stations into four types: terminating trains, originating trains, connecting trains, and passing-through trains. Generally, trains that need to enter the depot or temporary storage yard can be considered terminating trains; trains that come out of the depot or
temporary storage yard can be considered originating trains. Each train has a set of sequential operations. Trains operations can be classified into two types: static or dynamic. In static operations, trains only park on tracks or platforms and do not move during the operation time span; in dynamic operations, trains need to occupy a running route and move from the starting point to the end point of the route. We use tracks as the assigning resources for static operations and running routes for dynamic operations.

For a railway station, the operation time for dynamic operation for each class of trains is a standard value that cannot be changed. The operation time for static operation consists of a time window whose value varies within a range from a minimum to a maximum time. The dynamic train operation time and static operation time window for different classes of trains can be found in the operations manual of the railway station.

The OSP in railway stations can be defined as follows:

Given:
- A railway station and its infrastructure layout;
- A timetable of the arrival and/or departure time of each train at the involved station;
- A predetermined time span for each dynamic operation and a time window for each static operation;
- Train operation rules. The operation rules can be hard rules, such as that each train should go to its target yard and should have its own operation scheme, or soft preferences, such as that higher priority trains should stop at the tracks with the most convenient access.

The OSP in railway stations should determine the operation time and resource used (running route or track) for every operation, without any conflicts, to gain the maximum operation benefits or minimum resource usage costs.

Due to the complexity and difficulty of the OSP in railway stations, earlier studies typically approached the problem from several perspectives and did not take the infrastructure, including depot and access tracks into consideration. The model usually schedules regular in-service train operations but does not include the shunting operations before the depot and terminal.

Corman et al. (2011) presented an innovative optimization framework for real-time rescheduling of trains with different priority classes that can be computed statically or dynamically to include the needs of different stakeholders (1). Cacchiani et al. (2014) presented an overview of recovery models and algorithms for this problem (2). Pellegrini et al. (2014) proposed a mixed-integer linear programming formulation to address the real-time railway traffic management problem, representing the infrastructure at a fine granularity (3). However, the computation time for solving instances is relatively long.

Ariano et al. (2008) and Corman et al. (2009, 2010, 2011) presented a series of good optimization models and algorithms (4) (5) (6) (7). They provided an alternative graph that considers the occupation at block sections as nodes and the operation relations as arcs and used this graph to search for the optimal train dispatch solution. They considered many practical constraints and used a tabu search algorithm to obtain the optimal solution. Their models are more suitable for train timetable scheduling problem and do not include train shunting operations within the station area or train unit moving and maintenance.

For the problem of train dispatching, many scholars have done a lot of research. Dessouky et al. (2006) developed a branch-and-bound procedure to determine the optimal dispatching times for trains traveling in complex rail networks (8). Corman et al. (2012, 2014) performed an extensive review of dispatching and coordination algorithms in terms of feasibility, quality and computation time (9) (10). Meng and Zhou (2014) developed an innovative integer programming
model for this problem on an N-track network by simultaneously rerouting and rescheduling trains (11). However, the model does not have good computational performance for solving complex problems.

As for railway scheduling problem, Carey and Carville (2000, 2003, 2007) performed related work on it. Their earlier papers (Carey and Carville, 2000, 2003) proposed and tested models and algorithms for scheduling trains for a single, busy, complex station (12) (13). Their later work (Carey and Crawford, 2007) extended scheduling from a single station to a rail corridor (14). These methods have the advantage that their solution is based on practical rules, so the scheduling choices are understandable and acceptable to train planners and are easy for the planners to improve if they are not satisfied with the solution. However, the system seems more likely to solve train timetable problems than the OSP in stations because it does not consider other operations in stations, such as entering or exiting a depot. Meng et al. (2010) presented a timetable optimizing model that takes stability as the optimizing goal and a rescheduling model with minimal summary time as the destination (15). An improved particle swarm algorithm was applied in problem solving as a time-adjusting tool.

Many heuristic algorithms (Tomii et al. (1999), Yue et al. (2006)) are proposed for these problems (16) (17). Lee and Chen (2009) proposed an optimization-oriented heuristic to both assign routes and tracks and allocate time slots for a train (18). Liu and Kozan (2009) modeled the train scheduling problem in a railway station as a blocking parallel-machine job-shop scheduling problem and also proposed a feasibility satisfaction heuristic (19). Both of the research problems from Lee and Liu are similar to RDLAP and are also sub-problems of the OSP in stations. Jong et al. (2013) developed a novel two-stage hybrid method for high-speed trains (20). The first stage applied a genetic algorithm to solve train sequencing problems. The resultant sequence was then sent to the second stage to determine the optimal timetable with linear programming techniques.

Yang and Zhou (2014) considered the non-anticipativity constraints associated with the a priori path in a time-dependent and stochastic network and used a Lagrangian relaxation solution approach (21). Their algorithm cannot be applied to a network-wide traffic assignment problem involving multiple OD pairs.

From the above literature review, the following conclusions can be drawn:

(1) Most studies focused on the train routing problem to the platform tracks of railway stations but did not consider the operations within the station, such as train unit shunting, locomotive assignment and routing, or conflicts among shunting operations, the inbound and outbound movements of which are more complex than conflicts between regular inbound and outbound train operations.

(2) Scheduling and routing train operations in stations need to model the use of platform tracks, depot tracks and tracks to and from the depot by shunting operations. Furthermore, integrating all scheduled trains and shunting movements impacts not only the feasibility and quality of the solution but also the efficiency of the solution algorithm.

(3) Most of the algorithms for these problems are heuristics. Combining heuristic and mathematical optimal algorithms such as the Lagrangian relaxation algorithm is a very promising approach for solving real-world large-scale instances of such a problem.

The contribution of this paper is a general TRN that can be used to detect conflicts between scheduled train operations and simulate the whole operation process. Based on the TRN, a resource usage cost model is proposed, and the Lagrangian relaxation algorithm is tested and solved for a large, complex terminal station within a short time. The scheduled regular train
operations as well as the shunting movements between the station tracks and the train depot, including conflict detection and prediction, and the operation plan for a whole day is simulated using a visual DSS.

**TRACK-BASED RESOURCE NETWORK MODEL**

The scheduling operation plan must fit the target station’s infrastructure. The fixed equipment related to train operations mainly includes tracks, switches, crossings, platforms, and signals, and we introduce a graph $G(L,E)$ to represent the railway infrastructure. For each station, using the junction between adjacent track circuits, the center point of the switches, the center point of the crossovers, the mid-point of the platform tracks, the station border, and the yard border are defined as nodes, which we call “Anchor Points,” and the rail line that links two adjacent Anchor Points are defined as arcs. $L$ is the set of Anchor Points, and $E$ is the set of arcs linking the Anchor Points.

Figure 2 shows this simplified $G(L,E)$ graph of the TRN for a railway station topology. A **track** is the railroad linking two nodes, and a **platform track** is a station track that consists of three nodes and two tracks. A **train route** is a set of Anchor Points. If two routes have more than one identical node, other than the starting or ending points of the routes, then the two routes are conflicting routes. In Figure 2, Route$_2$ and Route$_3$ are conflicting routes because they have the same nodes ($l_{22}$ and $l_{29}$).

![Graph topology map of TRN for a railway station](image)

**FIGURE 2** Graph topology map of TRN for a railway station

All train operations can be scheduled on this TRN, and station resources can be represented by nodes in a simplified graph. No set of nodes are simultaneously taken by different trains.

**OPERATION SCHEDULING MODEL**

**Notations for OSP**

In this section, we give a mathematical model of train operation scheduling problem in a railway station. As described in an earlier section, an OSP assigns station resources to each operation without any conflicts. A simple way to accomplish this goal is to consider each route of a TRN as a resource without knowing whether conflicts exist. The complex way is to consider every node in the TRN as a resource. If the number of nodes is huge, we use a compromise method and take the tracks, switches, and crossovers as the resources to be assigned.

**General subscripts**

- $i$ Train index, $N$, the number of trains in this station, $i \in I$, where $I$ is the set of trains
- $j$ Operation index, $J$, the number of operations, $j \in J$, where $J$ is the set of operations
- $k$ Time index, $K$, the number of times, $k \in T$, where $T$ is the set of times
**Parameters**

- $h$: Track index, where $H$ is the number of tracks in this station
- $l$: Switch index, where $L$ is the number of switches in this station
- $m$: Cross index, where $M$ is the number of crossovers in this station
- $r$: Route index, $r \in R$, where $R$ is the set of routes; a route is made up of resources (track, switch or cross)
- $n$: Iteration index

**Decision variables:**

- $\chi$: The operation plan of all trains in this station
- $x_i^r$: The operation plan for train $i$
- $o_{ij}$: Operation $j$ of train $i$
- $\tau_i$: Operation number of train $i$
- $A_i$: The arrival time of train $i$ to this station
- $D_i$: The departure time of train $i$ out of this station
- $T_{ij}^{\min}$: Minimum operation time of $o_{ij}$
- $T_{ij}^{\max}$: Maximum operation time of $o_{ij}$
- $\mu_i$: Train type variable (if 1, train $i$ is a terminating train; if 0, train $i$ is a passing-through train; if -1, train $i$ is an originating train)
- $RT(r)$: The set of next routes of route $r$, $RT(r) \in R$
- $RO(j)$: The set of routes in which operation $j$ can run, $RO(j) \in R$

For dynamic operations, $T_{ij}^{\min} = T_{ij}^{\max}$, meaning that the same class train running a similar route takes a fixed time span. The static operation time can vary from $T_{ij}^{\min}$ to $T_{ij}^{\max}$ because some waiting time is required to avoid existing conflicts. $A_i$ and $D_i$ are given by the timetable. When train $i$ is a terminating train, $D_i = -1$, and when train $i$ is an originating train, $A_i = -1$.

**Decision variables:**

- $S_{ij}$: Starting time of $o_{ij}$
- $E_{ij}$: Finishing time of $o_{ij}$
- $y_{ijkr}$: 1 if Operation $j$ of train $i$ takes up route $r$ at time $k$
- $\delta_{ijkh}$: 1 if Operation $j$ of train $i$ takes up track $h$ at time $k$
- $\theta_{ijkl}$: 1 if Operation $j$ of train $i$ takes up switch $l$ at time $k$
- $\rho_{ijkm}$: 1 if Operation $j$ of train $i$ takes up cross $m$ at time $k$
- $v_{ijk}$: 1 if Operation $j$ of train $i$ is running at time $k$
Mathematical model

As mentioned previously, the goal of OSM is to maximize the total benefit of the operation plan. Let \( f(x) \) denote the total value of operation plan \( x \); the benefit function is the sum of the benefits of the trains scheduled in this station.

\[
f(x) = \sum_{i=1}^{N} \mathcal{f}^R(x_i^r)
\]

\( \mathcal{f}^R(x_i^r) \) is the benefit of train operation plan in this station. In busy railway stations, the goal for dispatchers is to schedule the operation plan to shunt trains to the depot as soon as possible or to minimize the waiting time of all trains at the arrival or departure yard. To minimize the waiting time also means to maximize the handling train number. We introduce the benefit curve for train operation time, as shown in Figure 3. When an operation is a dynamic operation, the operation benefit can be a constant. The benefit curve can take other functional forms, and we could select any functions similar to this curve.

![Figure 3: Benefit curve for each operation time](image)

OSM can be formulated as follows: \textit{Model 1: (P)}

Maximize \( Z = f(x) = \sum_{i=1}^{N} \mathcal{f}^R(x_i^r) \) \hfill (1)

Subject to:

\[
\sum_{i}^{N} \sum_{j=1}^{q_i} \delta_{ijkh} \leq 1 \quad \forall k, h
\] \hfill (2)

\[
\sum_{i}^{N} \sum_{j=1}^{q_i} \theta_{ijkl} \leq 1 \quad \forall k, l
\] \hfill (3)

\[
\sum_{i}^{N} \sum_{j=1}^{q_i} \rho_{ijkm} \leq 1 \quad \forall k, m
\] \hfill (4)

\[
\delta_{ijkh} = \sum_{r \in h} y_{ijk} \quad \forall i, j, k, h
\] \hfill (5)

\[
\theta_{ijkl} = \sum_{r \in l} y_{ijk} \quad \forall i, j, k, l
\] \hfill (6)

\[
\rho_{ijkm} = \sum_{r \in m} y_{ijk} \quad \forall i, j, k, m
\] \hfill (7)

\[
\sum_{j=1}^{r_i} \sum_{r} y_{ijk} \leq \sum_{j=1}^{r_i} v_{ijk} \quad \forall i, k
\] \hfill (8)
}\sum_{r \in R_{O}(j)} y_{ijk} = v_{ijk} \quad \forall i, j, k
\end{equation}

\begin{equation}
y_{ijk} - y_{ijk'} - \sum_{r' \in R_{T}(r)} y_{ij'k'_{r'}} \leq 1 - v_{ijk'} - v_{ij'k'} \quad \forall i, j, k, r, j' = j + 1, k' = k + 1
\end{equation}

\begin{equation}
T_{ij}^{\text{min}} \leq \sum_{k} v_{ijk} \leq T_{ij}^{\text{max}} \quad \forall i, j
\end{equation}

\begin{equation}
p_{ijk} \geq v_{ijk} \quad \forall i, j, k
\end{equation}

\begin{equation}
p_{ijk} \leq p_{ijk'} \quad \forall i, j, k, k' = k + 1
\end{equation}

\begin{equation}
q_{ijk} + v_{ijk} \leq 1 \quad \forall i, j, k
\end{equation}

\begin{equation}
q_{ijk} \leq q_{ijk'} \quad \forall i, j, k, k' = k + 1
\end{equation}

\begin{equation}
q_{ijk} = p_{ij'k} \quad \forall i, j, k, j' = j + 1
\end{equation}

\begin{equation}
S_{ij} = |T| - \sum_{k} p_{ijk} \quad \forall i, j
\end{equation}

\begin{equation}
E_{ij} = |T| - \sum_{k} q_{ijk} \quad \forall i, j
\end{equation}

\begin{equation}
(E_{i1} - A_{i})(1 + \mu_{i}) + (S_{ri} - D_{i})(1 - \mu_{i}) = 0 \quad \forall i
\end{equation}

Equation (1) is the objective function of OSM and maximizes the benefit of all train operation plans in a railway station.

Constraints (2)–(4) define the relationship between trains. These equations are the main constraints of a train’s operation plan model, meaning that at any time, resources (a track, a switch or a cross) can be occupied by no more than one operation and that once an operation takes up the resource, it will last until the end time of the operation.

Constraints (5)–(7) define the relationship between resources and routes for a train. These equations state that when the route is occupied, the resource (a track, a switch or a cross) that belongs to the route is occupied.

Constraints (8)–(10) define the relationship between the route and operation for a train.

Constraint (8) states that a route can be occupied only when train i is conducting an operation. Constraint (9) states that each operation can only occupy certain routes. Constraint (10) is a logic constraint that needs to be satisfied by generating a connected route. When \( y_{ijkr} = 0 \), meaning that operation \( j \) of train \( i \) does not occupy route \( r \) in time \( k \), the constraints can always be satisfied.

When \( v_{ijk'} = 0 \) and \( v_{ij'k'} = 0 \), meaning that train \( i \) does not conduct any operation in the next time \( k' \), the constraints can always be satisfied. When \( v_{ijk'} = 1 \), \( v_{ij'k'} = 0 \) and \( y_{ijkr} = 1 \), due to constraints (8)–(9), \( \sum_{r' \in R_{T}(r)} y_{ij'k'r'} = 1 \).

Constraints (11)–(19) define the relationship between operations for a train.

Constraint (11) is the operation time window, meaning that every operation must be finished within an amount of time longer than the minimum operation time and shorter than the maximum operation time. Constraints (12)–(16) guarantee that the operations of train \( i \) are sequential and that they form an operation chain. The operation set \( \{o_{ij}\} \) that fits constraints (12)–(16) is called the operation chain (OC) of train \( i \). Operation \( j \) of train \( i \) runs during the time span, so the value of \( v_{ijk} \) is one during the time span. The initial values of \( p_{ijk} \) and \( q_{ijk} \) are zero. If the operation has started, the value of \( p_{ijk} \) will be one. Similarly, if the operation has finished, the value of \( q_{ijk} \)
will be one. Constraints (17)–(19) are logic constraints that need to be satisfied by generating a feasible operation plan before solving the model. Equation (19) means that the train operation scheduling must guarantee the timetable of trains arriving at and/or departing from this station.

LAGRANGIAN RELAXATION-BASED HEURISTIC PROCEDURE

Lagrangian relaxation algorithm

In model $P$, constraints (2)–(4) define the main constraint relationship among different trains, referring to major conflicts in the operation plan. A common method of addressing these constraints is Lagrangian relaxation. By relaxing inequalities (2)–(4) using a Lagrangian method, the resulting Lagrangian relaxed problem is formulated: \textbf{Model 2: ($P_{LR}$)}:

$$P_{LR} = \text{Maximize } Z_{LR}$$

$$Z_{LR} = \sum_{i=1}^{N} f^{r}(x^{r}_i) + \sum_{kh} \lambda_{kh} (1 - \sum_{j=1}^{N} \sum_{i=1}^{T_{ij}} \delta_{ijkh}) + \sum_{kl} \phi_{kl} (1 - \sum_{j=1}^{N} \sum_{i=1}^{T_{ij}} \theta_{ijkl})$$

$$+ \sum_{km} \epsilon_{km} (1 - \sum_{j=1}^{N} \sum_{i=1}^{T_{ij}} \rho_{ijklm})$$

$$\lambda_{kh} \geq 0 , \phi_{kl} \geq 0 , \epsilon_{km} \geq 0 , \forall k, h, l$$

s.t. (5)–(19)

Relaxing constraints (2)–(4) means that we can solve the problem without considering route conflicts between different trains. $Z_{LR}$ can be transformed as:

$$Z_{LR} = \sum_{i=1}^{N} [f^{r}(x^{r}_i) - \sum_{kh} \lambda_{kh} \sum_{j=1}^{N} \sum_{i=1}^{T_{ij}} \delta_{ijkh} - \sum_{kl} \phi_{kl} \sum_{j=1}^{N} \sum_{i=1}^{T_{ij}} \theta_{ijkl} - \sum_{km} \epsilon_{km} \sum_{j=1}^{N} \sum_{i=1}^{T_{ij}} \rho_{ijklm}]$$

$$+ (\sum_{kh} \lambda_{kh} + \sum_{kl} \phi_{kl} + \sum_{km} \epsilon_{km})$$

s.t. (5)–(20)

We introduce two variables:

$$z_{i} = [f^{r}(x^{r}_i) - \sum_{kh} \lambda_{kh} \sum_{j=1}^{N} \sum_{i=1}^{T_{ij}} \delta_{ijkh} - \sum_{kl} \phi_{kl} \sum_{j=1}^{N} \sum_{i=1}^{T_{ij}} \theta_{ijkl} - \sum_{km} \epsilon_{km} \sum_{j=1}^{N} \sum_{i=1}^{T_{ij}} \rho_{ijklm}]$$

$$\mathfrak{R} = \sum_{kh} \lambda_{kh} + \sum_{kl} \phi_{kl} + \sum_{km} \epsilon_{km}$$

Model ($P_{LR}$) can also be expressed as follows:

$$P_{LR} = \text{Maximize } Z_{LR} = \sum_{i=1}^{N} z_{i} + \mathfrak{R}$$

s.t. (5)–(20)

The dual problem of $P_{LR}$ is as follows: \textbf{Model 3 - $P_{LD}$}:

$$P_{LD} = \text{min}_{(\lambda_{kh}, \phi_{kl}, \epsilon_{km})} \sum_{i=1}^{N} z_{i} + \mathfrak{R}$$

s.t. (5)–(20)

The upper bound of the model is $P_{LD}^{best} = \text{min}\{P_{LD}^{n}\}$, and the lower bound is $P_{LD}^{best} = \text{max}\{P^{n}\}$. We use sub-gradiente algorithm to update Lagrangian multipliers at each iteration.

Lagrangian relaxation sub-problem

To model ($P_{LR}$), the multipliers are given at every iteration, so $\mathfrak{R}$ is a constant. If we can obtain every maximum value of $z_{i}$, then $Z_{LR}$ will be at its maximum. The value of $z_{i}$ can be interpreted...
as the operation benefit minus the resource usage cost for train $i$. The prime problem can be decomposed into hundreds of sub-problems, and each sub-problem seeks an optimal operation plan for one train.

The generation of a feasible operation plan can be depicted on expanded resource-time networks. Taking the station in Figure 2 as an example, suppose train $i$ is coming from station border $b_1$ and will stop at one station track at time $A_i$ before going to the depot. $T_{l1}^{min} = T_{l1}^{max} = $ 3 minutes, $T_{l2}^{min} = 5$ minutes, $T_{l2}^{max} = 10$ minutes, and $T_{l3}^{min} = T_{l3}^{max} = 4$ minutes. After checking the TRN of the station, one route exists from $b_1$ to $h_1$, $h_2$, and $h_6$. Two routes exist from $b_1$ to $h_4$, and three routes exist from $b_1$ to $h_3$ ($h_5$ - 1 = \{b_1, l_1, l_7, l_{17}, l_{21}, m_5, l_2, l_{29}, l_{31}, l_{43}, h_5\}; h_1 \rightarrow h_5 - 2 = \{b_1, l_1, m_5, l_9, l_{25}, l_{45}, l_{43}, h_5\}; \text{and } h_1 \rightarrow h_5 - 3 = \{b_1, l_1, m_5, l_9, m_3, l_{11}, l_{33}, l_{31}, l_{43}, h_5\}$). Because tracks $h_3$ and $h_4$ are the main lines of the station, their function is to allow passing-through trains to move through the station quickly. Therefore, the dwelling operation cannot be assigned to these tracks. Only one route exists from $h_1$, $h_2$, and $h_6$ to the depot, and two routes exist from $h_3$ to the depot. The restricted path set derived from the feasible operation plan on a resource-time network for this train is illustrated in Figure 4.

![FIGURE 4 Restricted path set derived from a feasible train operation plan](image)

In the restricted path set, the nodes linked by dashed arcs are the same, and we call this a virtual operation. Nodes linked by solid arcs form what we call a real operation. We add two dummy nodes to these operation plan graphs: a starting point and an ending point.

With respect to the operation plans of train $i$, model $P$ needs to calculate the maximum value of the plan. The Lagrangian relaxation sub-problems can find the best path from the starting dummy point to the ending dummy point within the restricted path set.
Lagrangian relaxation–based heuristic algorithm

The scheme of the heuristic algorithm uses Lagrangian relaxation to guarantee the solution quality and uses the heuristic method to decrease the search space and improve the feasibility of the final solution. For each iteration, the algorithm will find the optimal path in the restricted path set for every sub-problem and then aggregate the optimal solution of every sub-problem as the total solution at this iteration, updating the resource penalties according to the resource usage times. After a number of iterations, if the resource penalties reach steady values or the iteration number reaches a pre-set value, then the calculation process terminates.

The process of the heuristic algorithm is described as a flow chart in Figure 5.

---

**FIGURE 5 Flow chart of the heuristic algorithm**
COMPUTATIONAL TEST AND APPLICATION RESULTS

In this section, we provide a large-scale real world application case of the Beijing South Railway Station to test the effectiveness of and validate the OSM and Lagrangian relaxation heuristic algorithm. The algorithm is implemented in C# on a Windows 7.0 platform and evaluated on a personal computer with an Intel(R) Core(TM) i5 CPU and 4 GB of memory.

Beijing South Railway Station includes 56 conventional trains, 172 high-speed trains, and 200 inter-city trains, creating more than one thousand operations every day. It has 29 tracks, 118 switches, and 12 crossovers. Many operation conflicts occur in the left throat of the station because trains running into or out of the station and into or out of the depot all need to use the switches and crossovers located at that throat.

We chose three cases to test this approach. Case A uses the peak hour of the station, 7:00AM–8:00AM, with 27 trains during this one-hour peak. Case B uses a 3-hour period from 7:00AM–10:00AM, with 73 trains during this period. Case C is the 24-hour whole-day operation plan. The maximum iteration numbers are 50 for Case A, 100 for Case B, and 400 for Case C. Figure 6 shows the best values for the three cases.

FIGURE 6 Best values during iteration
The calculation terminates when the maximum iteration number is reached in all cases. The computation results are listed in Table 1. The gap is 1.17% for Case A, 8.92% for Case B, and 17.08% for Case C; the total computation time is 4 seconds for Case A, 56 seconds for Case B, and 840 seconds for Case C, which are acceptable times for such a large-scale problem. For scheduling a one-hour operation plan, it takes less than 10 seconds to obtain a very near optimal solution without any conflicts. For scheduling a 3-hour plan, it takes approximately one minute to obtain a good solution.

<table>
<thead>
<tr>
<th>Cases</th>
<th>Number of trains</th>
<th>Number of iterations</th>
<th>Best value</th>
<th>Gap</th>
<th>Computation time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case A</td>
<td>27</td>
<td>50</td>
<td>3185.9</td>
<td>1.17%</td>
<td>4</td>
</tr>
<tr>
<td>Case B</td>
<td>73</td>
<td>100</td>
<td>11631.6</td>
<td>8.92%</td>
<td>56</td>
</tr>
<tr>
<td>Case C</td>
<td>428</td>
<td>400</td>
<td>129510.5</td>
<td>17.08%</td>
<td>840</td>
</tr>
</tbody>
</table>

CONCLUSIONS

Many railway station dispatchers in China still use pencil and paper to schedule train operations. It takes several years for to dispatchers become familiar with the railway’s infrastructure and standard operations. Before they become eligible dispatchers, a relatively high-quality and fast DSS would be very useful for them. This paper was motivated by the need for a computer-aided railway station dispatching tool.

This paper offers three major contributions. First, a detailed track-based resource network was proposed. The TRN has proven to be very useful to technicians in railway stations as it can help them recognize important track devices and forecast operation conflicts. After the operation plan is established, it can be simulated and provide data to analyze the capacity of the station and identify bottlenecks in the station track devices. Second, a novel approach for train operation scheduling based on microscopic resource (platform tracks, switches and crossovers) assignment was proposed. Moreover, the Lagrangian relaxation heuristic algorithm can generate a good solution within a rather short computation time. Third, we developed several combinatorial methods to properly solve the operation scheduling problem of regular scheduled trains in operation and shunting trains simultaneously for a large number of variables and constraints.

In further research, we will focus on three major aspects: (1) we will improve TRN and OSM, e.g., by considering sectional release of a route, accounting for train acceleration and deceleration times. (2) We will improve the computation speed of the Lagrangian relaxation–based algorithm. To increase the computational performance speed, we need to integrate more search rules into the algorithm to get feasible solution fast. (3) The operation plan presented is limited to train operation and shunting operation. For other operation plans, such as the crew assignment plan and maintenance plan, we intend to simplify their treatment by providing sufficient standard work times to each task.

ACKNOWLEDGEMENTS

This work was financially supported by the Fundamental Research Funds for the Central Universities provided by the Ministry of Education of the People's Republic of China (No. 2014JBZ008) and Project of Science and Technology Research Program of China Railway Corporation (No. 2016X006-C).
REFERENCES


