

Research Article

Turnback Capacity Assessment and Delay Management at a Rail Transit Terminal with Two-Tail Tracks

Zhibin Jiang,¹ Yuyan Tan,² Feng Wang,³ and Lei Bu³

¹*School of Transportation Engineering, Key Laboratory of Road and Traffic Engineering of the Ministry of Education, Tongji University, Shanghai 201804, China*

²*Institute of Railway Systems Engineering and Traffic Safety, Technical University of Braunschweig, Pockelsstraße 3, 38106 Braunschweig, Germany*

³*Department of Civil and Environmental Engineering, Jackson State University, P.O. Box 17068, Jackson, MS, USA*

Correspondence should be addressed to Yuyan Tan; y.tan@tu-bs.de

Received 25 December 2014; Revised 4 February 2015; Accepted 6 February 2015

Academic Editor: Anders Eriksson

Copyright © 2015 Zhibin Jiang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Terminal capacity and performance have become a major concern for rail transit agencies in China due to the ever increasing passenger demand. This paper develops a mixed integer programming (MIP) optimization model to estimate the turnback capacity and performance of a rail transit terminal with two-tail tracks. The capacity evaluation and delay propagation are described and assessed as an N -track integrated model with minimal time span and train delay. Operations and design parameters such as tail track allocation strategies, maximum layover time, headway pattern, buffer time distribution scheme, and primary delay are also considered in this model. The effectiveness of the model is tested by a case study with computation results drawn from one terminal station in Shanghai, China. The case study results show that unfixed platform time and flexible tail track allocation strategies can improve the capacity of turnback operation, and the strategy of allowing swapping of the tail tracks has a significantly positive impact on delay absorption.

1. Introduction

A rail transit terminal is a key point for train movements, where trains reverse direction (turn back) to provide continuing service in the opposite direction. Careful planning and operations of a terminal are essential for transit agencies to provide quality service to passengers effectively and efficiently [1]. The line capacity of a rail transit is often governed by the frequency with which trains can be reversed at terminal stations and intermediate turnback sites [2]. Terminal capacity and performance have become a major concern for rail transit agencies in China due to the ever increasing passenger demand, which has constantly driven the need for more frequent service up to the capacity that the system was originally designed for. Take Shanghai as an example; as of March 2014, there are 14 rail transit lines operating in Shanghai, with a total operating route length of 538 kilometers and 329 stations. On each weekday, over 8 million passengers ride on the lines of Shanghai rail transit network,

and five of the 14 rail lines in Shanghai have become nearly saturated due to the limiting capacities of the terminals. Similar issues are happening in other cities in the world as well. From a survey of the Transit Cooperative Research Program (TCRP), nine out of 58 responding transit systems cited turnbacks as a constraint [3].

However, there is a lack of well-established concepts and tools in the existing rail transit literature to optimize the tail track allocation strategies, maximum layover time, headway pattern, recovery time, and buffer time distributions at a heavily utilized terminal. This paper deals with the problem of estimating the capacity of a rail transit terminal with two-tail tracks, which is a popular configuration for rail transit terminals with high demands in China. The main objectives of this paper are to construct a new MIP optimization model for the turnback capacity and disruption management problems of a rail transit terminal and to develop a robust numerical tool that can help transit agencies to better assess the capacity and performance of their terminals to serve the

ever increasing demands, while considering the major factors that affect the capacity of a transit terminal. The application of this model may also be applied to the planning, design, and safety assessment of a rail transit terminal.

2. Background

Both in literature and in practice, quite a lot of effort has been devoted to the assessments of the capacity and delays of high-intensity terminals in a rail transit network. A simple model was presented in TCRP Report 13 [3], and the capacity of a 2-track stub-end terminal based on infrastructure layout and train movement characteristics was estimated. More recently, in TCRP Report 165, Kittelson et al. [4] presented a formula to extend the approach to estimating terminal layover time and found that the layout and configuration design of a terminal played a key role in providing delay recovery and service performance.

Lee [1] stated that terminal capacity and performance have become a major concern for public transport authorities, and the research results confirmed the effects of terminal configuration on service quality, but the results were limited to only one type of terminals (with no tail track). Based on simulation studies, van Oort and van Nes [5] presented their research of the effects of various terminal configurations on reliability of service and calculated the average delay per vehicle for three main types of terminals.

Obviously, terminal capacity is closely related to timetable scheduling. Yet, the train scheduling problem in a rail transit system is quite complex [6]. Numerous proposed scheduling models and algorithms have addressed different but related optimization objectives, including capacity, route service, minimal delay, passenger demand, operating cost, and transit unit (TU) circulation [7–12]. However, although the parameters of minimum headway and layover time at terminals were considered in these studies, other important factors, such as the tail track allocation strategy and delay recovery time distribution, were not included in these studies.

Actually, the major factors that affect the capacity and reliability of a high-intensity terminal should include track allocation strategies, train operation process, and recovery time and buffer time distributions. Since terminal congestion happens most often when high frequencies are scheduled [13, 14], it would be necessary to understand the distributions of the recovery time and buffer time to study the terminal capacity. Furthermore, maximum capacity has a significant effect on terminal reliability and delay management, which has often been identified as the main performance criterion for customer satisfaction on rail transit system [15].

Due to the high passenger volumes and high utilization of rail transit lines, delay management has also received much attention from the perspectives of passengers or rail transit companies. Schmöcker et al. [15] stated that the terminal recovery times can often be used to reduce both the cumulative delay and the number of affected trains. Carrel et al. [13] also proposed that the capacity of terminals and reversing points has a direct influence on the propagation of delays throughout the line. Puong and Wilson [16] stated that delay recovery is preferably performed at terminal stations

to minimize the negative impacts of the disruption on the reverse direction. The result of Wang et al. [17] showed that sections with switches and turnbacks at terminal stations are both capacity bottlenecks for a moving block railway line. Furthermore, Dollevoet et al. [18] developed a model for the delay management problem that considered the capacities of the stations and allowed the rescheduling of the platform track assignment under a real-time operating scenario.

There are quite some published works on the real-time delay management in rail transit terminals. Cacchiani et al. [19] presented an overview of recovery models and algorithms for real-time railway disturbance and disruption management. Flamini and Pacciarelli [20] addressed a scheduling problem arising in the real-time management of a metro rail terminus in a rail transit system. The scheduling problem is modeled as a bicriteria job shop scheduling problem with additional constraints. Mannino and Mascis [21] developed a real-time automated traffic control system to operate trains in metro stations. For each candidate routing an instance of the blocking, no-wait job-shop scheduling problem with convex costs is solved to optimality by branch and bound.

However, in these studies, a lot of attention has been paid to passenger demand, train unit circulation, or line capacity, but few of the studies looked at service regularity, tail track allocation strategies, maximum layover time, buffer time distribution scheme, and the delay management at terminals. In practice, in China terminal capacity is calculated by transit agencies who always allocate trains to one fixed tail in the process of timetable scheduling, and, in actual operation, the trains would maintain the turnback track as planned in the original timetable. Obviously, this operating strategy is not an optimal method of using the resources of tail tracks and platforms at a terminal.

Therefore, this study introduces a number of new features to capture the influence of major elements that have not been comprehensively considered in past studies on the problem of estimating terminal capacity and delay management. Firstly, we will focus on developing and refining capacity concepts relevant to rail transit terminals with tail tracks. Secondly, an N -track based MIP model, which considers important operating and design parameters, such as tail track allocation strategies, maximum layover time, and buffer time distribution scheme, will be developed. Finally, a case study in Shanghai will be used to illustrate the practical value of the developed model.

3. Problem Description

3.1. Rail Transit Terminal Types and Characteristics. Rail transit terminals may be designed in many forms: hundreds of types exist over the world. In China, a stub-end terminal with one of the four main types is commonly chosen in rail transit lines, as shown in Figure 1.

Figure 1(a) is a typical stub-end terminal with crossovers located in advance of the station. Trains arrive and depart from the same platform and all the passenger and crew activities happen during the time the train stays on the platform. Figure 1(b) is a typical terminal with far-side crossovers and tail tracks beyond the station; it can also be called a relay

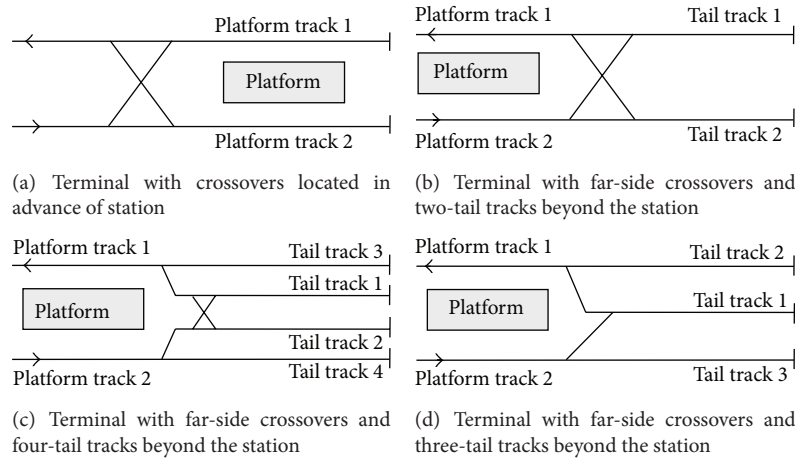


FIGURE 1: Four main types of stub-end terminal.

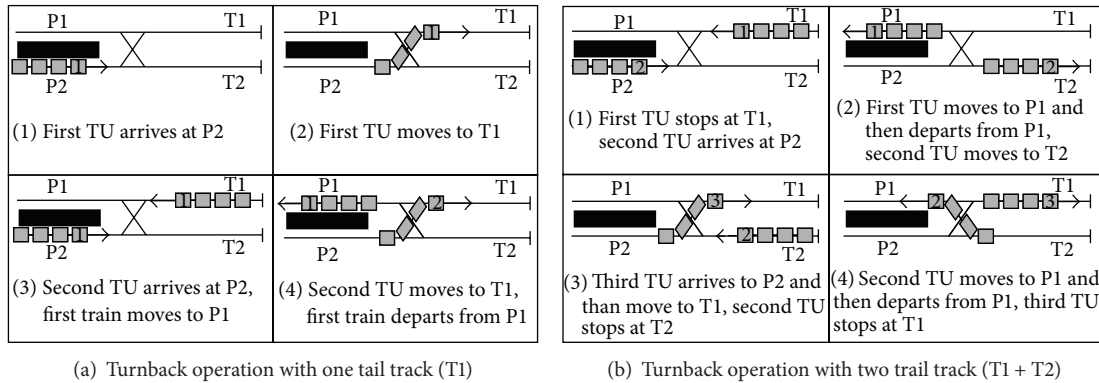


FIGURE 2: Two turnback operation procedures of a terminal with two-tail tracks.

terminal [1]. Trains use separate tracks/platforms for arrival and departure. The arriving train pulls into one platform, then pulls into one of the tail tracks, changes direction, and then returns to pick up passengers from the other platform.

Figures 1(c) and 1(d) are terminals with four and three-tail tracks, respectively. These terminals are often designed as a function of operation by short turning at the middle position of a line and allow a storage track or tracks for spare or disabled trains—a useful, if not essential, failure management facility. Normally, if all the trains need to turn back at the terminal, the terminal types of Figures 1(b) and 1(c) have the same procedure for train turnbacking, because only two-tail tracks (tail tracks 1 and 2) of Figure 1(c) can be used for turnback operation. Because of its popularity in use for high demands, we will take the terminal with two-tail tracks (Figure 1(b)) as the terminal type for our study.

3.2. Turnback Operation at Terminal with Two-tail Tracks. Under low frequency operations, only 1 tail track (the preferred) may be used for train turnbacks. In high frequency operations, both tail tracks would be needed for train processing with three operation modes: only at tail track 1 (T1), only at tail track 2 (T2), and both at tail tracks 1 and 2 (T1 + T2). Figure 2 shows the two turnback operation procedures of a terminal with two-tail tracks.

It is evident that tail track allocation is an integral component at both the strategic planning level and the operational level. In determining the capacity of a rail transit line at the strategic planning level, one must allocate the associated turnback track over a certain time horizon in such a way that the maximum number of trains can be scheduled. Furthermore, on an operational level, when a disruption to the planned timetable occurs, one must reallocate the tracks in such a way that the negative impact of the disruption is minimized.

3.3. Layover Time. Layover time is the time between the scheduled arrival and departure of a vehicle at a transit terminal station. The layover time at a terminal with tail tracks consists of several time elements, which are illustrated in Figure 3, and the detail explanation of these elements was introduced by van Oort and van Nes [5]. Adequate buffer time at a terminal is always needed to absorb delays due to variability of train arrivals and terminal processes. This helps provide reliability of train departure times. However, too much buffer time reduces productivity, constrains terminal capacity, and may result in terminal congestion.

3.4. Theoretical Capacity and Delay Management of Terminal. Theoretical capacity of a terminal is defined as [1]: maximum

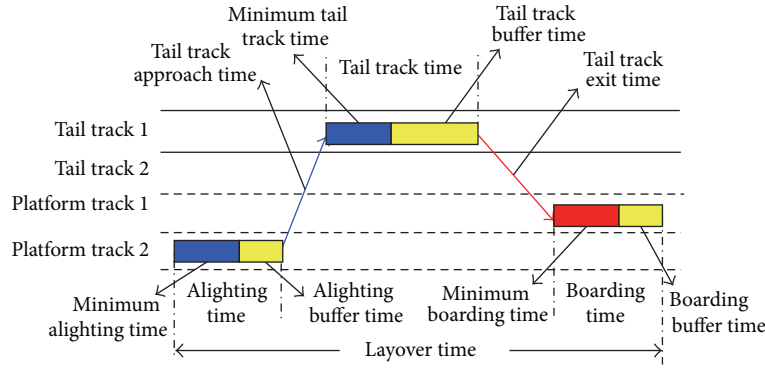


FIGURE 3: Time compositions of layover time at terminal with tail track.

number of incoming trains a terminal can process per unit time based on the existing infrastructure and type of operations. When a system operates at close to its capacity, any significant delay may use up a train's terminal layover time, resulting in potential delay to the trip in the reversed direction [4]. The capacity of terminals and reversing points has a direct influence on the propagation of delays throughout the line, because the layover time scheduled at a terminal helps put trains back into sequence or on schedule. A "delay recovery strategy" can be defined as the series of decisions made by the operations management to maximize the service performance following the delays that occur after an incident. Since terminal capacity is closely related to delay management, this study followed the result of Törnquist and Persson [22], which considered the rescheduling of trains under disturbance on a so called N -tracked network, where N refers to the number of tracks in a segment. We extended their model by adding crossovers or block sections to form multiple parallel tracks in a segment of high service demand.

4. Optimization Model

4.1. Problem Definition. In this paper we denote a segment as a collection of one or multiple tracks between two points where a point merely is the connection between two segments. A track can only be occupied by one train at a time and the track can be either unidirectional (i.e., only permitting one-way traffic) or bidirectional. In the model of Törnquist and Persson [22], the switches between the tracks were not defined explicitly. That means, if two trains use the same track within a segment and thereby are separated by time, their paths are not considered to be in conflict, but in practice they could be. The proposed model in this paper has handled this issue.

Figure 4 shows the illustration of the terminal with two-tail tracks, and the terminal is divided into three segments. Segment 1 represents the platform tracks (A and B), segment 2 represents double crossovers (D) and two main tracks (C and E), and segment 3 represents tail tracks (G and F). Segments 1 and 3 are double-tracked and segment 2 is triple-tracked. It is also important to know in which direction(s) the track can be operated. A bidirectional track can be used in any direction (but still only by one train at a time) while a unidirectional

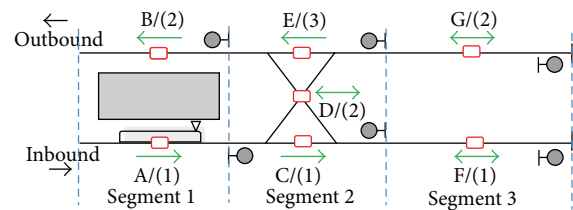


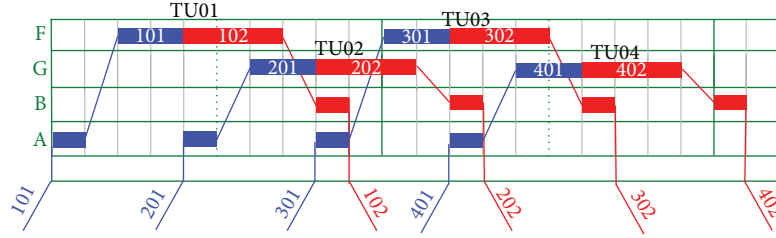
FIGURE 4: The N -tracked model of terminal with two-tail tracks.

only in one specific direction. Therefore, according to the operations practice of turning back, the directional rule of each track at all segments is shown in Figure 4.

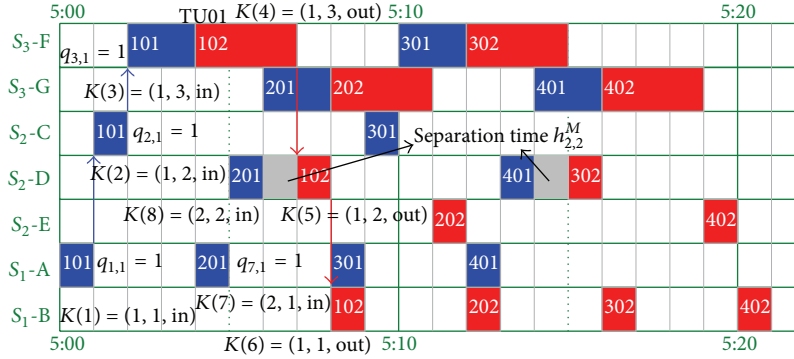
The temporary resource request by a train to occupy a segment is hereafter referred to as an event. An event has initial start and end times for a track within the segment but then needs to be updated with new start and end times (and possibly track name) during operations. Figures 5(a) and 5(b) illustrate the platform and tail track occupation diagram and event allocation diagram. The event allocation diagram shows the resources for segments 1–3 and how the segments are allocated to the events. Each box represents a scheduled event and is numbered with respect to the train the event is associated with.

Figure 6 shows the possible sequence of events of each train trip. For inbound trains, the events must start at track A in segment 1, and then select one track (track C or track D) in segment 2. If one train selects track C in segment 2, it must select track F in segment 3; otherwise it must select track D in segment 2 and track G in segment 3. Then the first event of the connecting trip, which is linked to the same TU, must select the same track in segment 3. If an outbound train selects track F, it must select track D and track B subsequently, or if it selects track G, then it must select track E and track B. The time taken to traverse segments 1 and 2 must equal the minimum required traversing time, which means that none of the trains can add extra stop time in segment 1 and segment 2.

Every transit unit (TU) has six events and must follow a fixed sequence of segments. In modeling, we define two main train trips: inbound train trips and outbound train trips and all trips are grouped by pairs, with each pair of train



(a) Platform and tail track occupation diagram



(b) Events allocation diagram

FIGURE 5: Track occupation diagram and events allocation diagram of N -tracked model.

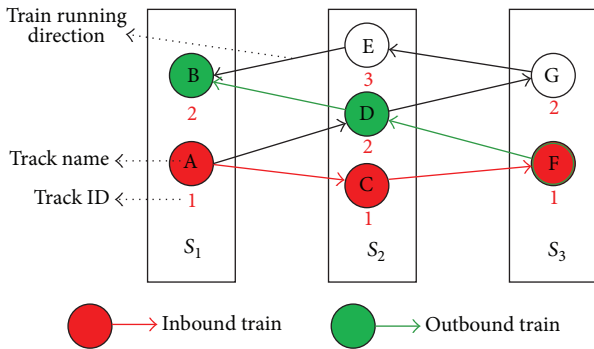


FIGURE 6: Possible sequence of events of each train.

trips denoting a same TU. In Figure 5 there are 8 train trips and therefore 4 TUs, for example, trip 101 and trip 102 are connected by TU01, and trip 102 must share the same track with trip 101 at segment 3.

Different events of train movements may cause conflicts due to the share use of the same equipment (track and/or switch). The conflicts are detected whenever different events try to occupy the same track at the same time, and these events require a gap time to separate them (i.e., a gap time between a train leaving the track and another train entering the track). The length required for the gap time is different depending on the moving direction of the conflicting events (meeting at opposite directions or following at a same direction). Figure 7(a) shows an example of events conflicting at track F, where train 2 arriving at track F must ensure that train 1 has left track F. Figure 7(b) shows an example of events conflicting at track D (crossover), where train 4 arriving at

track F (through track D) must ensure that train 3 has left track D.

4.2. Model Formulation. In this section, a Mixed integer programming (MIP) optimization model is developed for the problem of terminal capacity assessment with delay management. It is noted that the time units of following parameters and decision variables are all in seconds.

4.2.1. Sets. The sets below are used for the mathematical model:

U : the set of TUs,

S : the set of segments, in the model of Figure 4, $S = (1, 2, 3)$,

S^T : the set of segments with tail tracks, in the model of Figure 4, $S^T = 3$,

S^P : the set of segments with platform tracks, in the model of Figure 4, $S^P = 1$,

S^C : the set of segments with crossover tracks, in the model of Figure 4, $S^C = 2$,

S^W : the set of segments where trains can wait extra time at tracks,

K : the ordered set of events where an event is a resource request by a certain train for a specific segment, $K = [(i, j, in/out), i \in U, j \in S]$, which means the event of TU i at segment j in inbound or outbound direction,

$K_{i,j}^{in}$: the ordered set of events for TU i at segments j in inbound direction, $i \in U, K_{i,j}^{in} \in K, j \in S$,

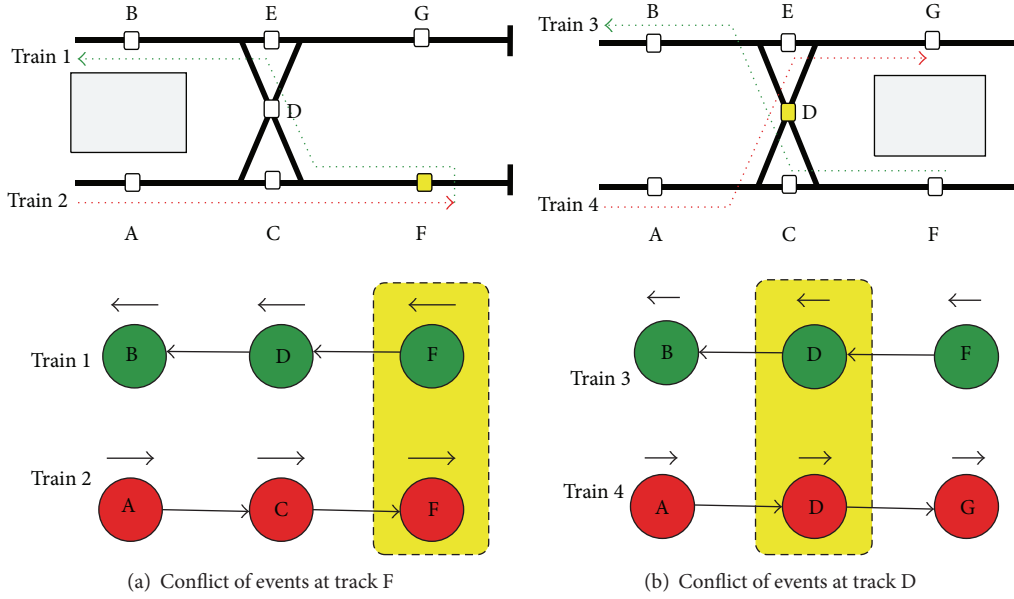


FIGURE 7: Events conflict detection at different track.

$K_{i,j}^{\text{out}}$: the ordered set of events for TU i at segments j in outbound direction, $i \in U$, $K_{i,j}^{\text{out}} \in K$, $j \in S$,

K_j^S : the ordered set of events of segment j , $j \in S$,

P_j : the set of parallel tracks of segment j , $j \in S$,

f_i : the ordered set of sequence numbers for first events in K for TU i , $i \in U$,

l_i : the ordered set of sequence numbers for last events in K for TU i , $i \in U$.

Accordingly, the ordered set of events K for TU01 in Figure 5 can be expressed as $K = [(1, 1, \text{in}), (1, 2, \text{in}), (1, 3, \text{in}), (1, 3, \text{out}), (1, 2, \text{out}), (1, 1, \text{out})]$.

4.2.2. *Parameters.* The model uses the following parameters, which are assumed to be in integer values:

n^{TU} : the number of initial TUs,

n : the number of events of each TU, in the model of Figure 5, $n = 6$,

f_e : the sequence number of the first event in K , in the model of Figure 4, $f_e = 1$,

l_e : the sequence number of the last event in K , in the model of Figure 4, $l_e = n \times n^{\text{TU}}$,

$d_{j,t}^{\text{in}}$: the minimal occupation time for track t at segment j in the inbound direction, it specifies the minimum nonstop running time for the segment, or it represents the minimal dwell time (including the running time of the train entering or leaving the track) at platform or tail track, $j \in S$, $t \in P_j$, $j \in S$,

$d_{j,t}^{\text{out}}$: the minimal occupation time for track t at segment j in the outbound direction, $j \in S$, $t \in P_j$, $j \in S$,

$h_{j,t}^M$: the required gap time for track t at segment j if the trains meet, $j \in S$, $t \in P_j$, $j \in S$,

$h_{j,t}^F$: the required gap time for track t at segment j if one train follows the other, $j \in S$, $t \in P_j$, $j \in S$,

t^b : the start time of the first event f_e ,

t^{delay} : the primary delay time,

π_k^{begin} : the start time of event k in the initial timetable, $k \in K$,

π_k^{end} : the end time of event k in the initial timetable, $k \in K$,

o_k : the run direction of event k , $k \in K$,

M : a sufficiently large positive constant, here given the value $3600 * 24 = 86,400$,

$\theta_{k,t} = \{1, \text{ if } k \text{ uses track } t \text{ (as in the initial timetable), } k \in K_j^S, t \in P_j, j \in S^T; 0, \text{ otherwise}\}$,

$b_{k,j} = \{1, \text{ if event } k \text{ runs inbound direction at segment } j, k \in K_j^S, j \in S; 0, \text{ otherwise}\}$,

N^{fix} : the sequence number of fixed tail track.

4.2.3. *Decision Variables.* The following decision variables are defined in the model:

x_k^{begin} : the start time of event k , $k \in K$,

x_k^{end} : the end time of event k , $k \in K$,

$q_{k,t} = \{1, \text{ if event } k \text{ uses track } t, k \in K_j^S, t \in P_j, j \in S; 0, \text{ otherwise}\}$,

$\gamma_{k\hat{k}} = \{1, \text{ if event } k \text{ occurs before event } \hat{k}, k, \hat{k} \in K_j^S, j \in S, k < \hat{k}; 0, \text{ otherwise}\}$.

4.2.4. *Objective Function.* We consider objectives in the view of the following two aspects.

- (a) The maximal number of TUs can be operated in a given time period, which can be represented by minimizing the time span of all trains:

$$\text{Min} (x_{le}^{\text{end}} - x_{fe}^{\text{begin}}). \quad (1)$$

- (b) Least delay of all train operations, which can be presented by minimizing the total arrival and departure delays of trains on platform tracks:

$$\text{Min} \sum_{i \in U} (x_{fi}^{\text{begin}} - \pi_{fi}^{\text{begin}}). \quad (2)$$

4.2.5. *Constraints.* In this section, we will focus on the constraints of the model, which are listed as follows.

(1) *Event Time Constraints.* Consider

$$x_{fe}^{\text{begin}} = t^b, \quad (3)$$

$$x_{k+1}^{\text{begin}} = x_k^{\text{end}}, \quad k \notin K_{i,j}^{\text{out}}, \quad i \in U, \quad j \in S, \quad (4)$$

$$x_{fi+1}^{\text{begin}} \geq x_{fi}^{\text{begin}}, \quad i \in U, \quad i < n^{\text{TU}}, \quad (5)$$

$$x_{li+1}^{\text{end}} \geq x_{li}^{\text{end}}, \quad i \in U, \quad i < n^{\text{TU}}, \quad (6)$$

$$x_k^{\text{end}} \geq x_k^{\text{begin}} + b_{k,j} \cdot d_{j,t}^{\text{in}} + (1 - b_{k,j}) \cdot d_{j,t}^{\text{out}}, \quad (7)$$

$$k \in K, \quad j \in S^W, \quad t \in P_j,$$

$$x_k^{\text{end}} = x_k^{\text{begin}} + b_{k,j} \cdot d_{j,t}^{\text{in}} + (1 - b_{k,j}) \cdot d_{j,t}^{\text{out}}, \quad (8)$$

$$k \in K, \quad j \notin S^W, \quad t \in P_j,$$

$$x_{fi+2}^{\text{begin}} - x_{fi+1}^{\text{begin}} = x_{fi+1}^{\text{begin}} - x_{fi}^{\text{begin}}, \quad (9)$$

$$i \in U, \quad i < n^{\text{TU}} - 1,$$

$$x_{li+2}^{\text{end}} - x_{li+1}^{\text{end}} = x_{li+1}^{\text{end}} - x_{li}^{\text{end}}, \quad (10)$$

$$i \in U, \quad i < n^{\text{TU}} - 1.$$

Constraint (3) defines the start time of first event to be t^b . Constraint (4) specifies that each train event is directly succeeded by the next one in the ordered set of events for the train trip. Constraint (5) defines the entering sequence of all arrival trains. Constraint (6) ensures that the first entering train must depart first; namely, all the trains must obey the first in first out rule. At the segments that train can wait a long time; actual occupation time should be no less than the minimum occupation time; otherwise, actual occupation time must be equal to the minimum occupation time. These facts are depicted in constraints (7) and (8). Constraints (9)

and (10) specify identical gaps between all arrival or departure trains at platform tracks.

(2) *Constraints at Segments and Tracks.* Consider

$$\sum_{t \in P_j} q_{k,t} = 1, \quad k \in K_j^S, \quad j \in S, \quad (11)$$

$$q_{k,1} = 1, \quad k \in K_{i,j}^{\text{in}}, \quad i \in U, \quad j \in S^P, \quad (12)$$

$$q_{k,2} = 1, \quad k \in K_{i,j}^{\text{out}}, \quad i \in U, \quad j \in S^P, \quad (13)$$

$$q_{k,t} = q_{k-1,t}, \quad k \in (K_{i,j}^{\text{in}} \cup K_{i,j}^{\text{out}}), \quad (14)$$

$$t \in P_j, \quad i \in U, \quad j \in S^T,$$

$$q_{k,3} = q_{k-1,2}, \quad k \in K_{i,j}^{\text{out}}, \quad i \in U, \quad j \in S^C, \quad (15)$$

$$q_{k,2} = q_{k-1,1}, \quad k \in K_{i,j}^{\text{out}}, \quad i \in U, \quad j \in S^C, \quad (16)$$

$$x_{\hat{k}}^{\text{begin}} - x_k^{\text{end}} \geq h_{j,t}^M \cdot \gamma_{k\hat{k}} - M \cdot (1 - \gamma_{k\hat{k}}), \quad (17)$$

$$k, \hat{k} \in K_j^S, \quad j \in S, \quad k < \hat{k} - 1, \quad o_{\hat{k}} \neq o_k,$$

$$q_{k,t} + q_{\hat{k},t} = 2, \quad t \in P_j,$$

$$x_{\hat{k}}^{\text{begin}} - x_k^{\text{end}} \geq h_{j,t}^F \cdot \gamma_{k\hat{k}} - M \cdot (1 - \gamma_{k\hat{k}}), \quad k, \hat{k} \in K_j^S, \quad j \in S, \quad (18)$$

$$k < \hat{k} - 1, \quad o_{\hat{k}} = o_k, \quad q_{k,t} + q_{\hat{k},t} = 2, \quad t \in P_j,$$

$$x_k^{\text{begin}} - x_{\hat{k}}^{\text{end}} \geq h_{j,t}^M \cdot (1 - \gamma_{k\hat{k}}) - M \cdot \gamma_{k\hat{k}}, \quad k, \hat{k} \in K_j^S, \quad j \in S, \quad (19)$$

$$k < \hat{k} - 1, \quad o_{\hat{k}} \neq o_k, \quad q_{k,t} + q_{\hat{k},t} = 2, \quad t \in P_j,$$

$$x_k^{\text{begin}} - x_{\hat{k}}^{\text{end}} \geq h_{j,t}^F \cdot (1 - \gamma_{k\hat{k}}) - M \cdot \gamma_{k\hat{k}}, \quad k, \hat{k} \in K_j^S, \quad j \in S, \quad (20)$$

$$k < \hat{k} - 1, \quad o_{\hat{k}} = o_k, \quad q_{k,t} + q_{\hat{k},t} = 2, \quad t \in P_j.$$

Constraint (11) enforces that every event must use exactly one track per relevant segment. Constraints (12) and (13) ensure that the entering and departing trains at segment 1 must select track 1 and track 2, respectively. Constraint (14) ensures that if a train enters segment 3, it must select the same track ID of the foregoing event. Constraints (15) and (16) ensure that if a train departs from track G, it must select track E, and if a train departs from track F, it must select track D. Constraints (17) through (20) specify that one event at a segment must end with the elapse of a required gap time before a next event may start on the same segment, if the events are using the same track of the segment. The length of the gap time depends on whether the conflicting trains meet ($h_{j,t}^M$) or follow ($h_{j,t}^F$) on each other.

4.3. Evaluation Strategy. Four strategies, strategies 1 through 4, are considered in this study regarding the rule of turnbacking and delay management. Strategies 1 and 2 are applied to quantify the maximal turnback capacity of the terminals with different operating methods, and strategies 3 and 4 are mainly applied to assess the delay propagation by different ways of tail track allocation.

4.3.1. Strategy 1: Select Fixed Tail Track. This strategy allows trains to select one fixed tail track. The strategy adopts a formulation including objective function (1) and constraints (3) though (20) in addition to the following constraint:

$$q_{k,t} = 1, \quad t = N^{\text{fix}}, \quad k \in K_j^S, \quad j \in S^T. \quad (21)$$

4.3.2. Strategy 2: Select from the Two-tail Tracks and Every Two TUs Are Grouped and Operated with a Cyclic Pattern. This strategy allows trains to select any one of the tail tracks. In order to assure that all trains selecting the same turnback track keep the same occupation process, every two TUs are grouped and run in the cyclic pattern; that is, the third TU has the same occupation time and track as the first TU and the same for the fifth. The strategy adopts a formulation including objective function (1) and constraints (3) through (20) in addition to the following constraints:

$$\begin{aligned} x_{k+3n}^{\text{begin}} - x_{k+n}^{\text{begin}} &= x_{k+2n}^{\text{begin}} - x_k^{\text{begin}}, \quad k \in K, \quad k \leq n \times (n^{\text{TU}} - 3), \\ x_{k+3n}^{\text{end}} - x_{k+n}^{\text{end}} &= x_{k+2n}^{\text{end}} - x_k^{\text{end}}, \quad k \in K, \quad k \leq n \times (n^{\text{TU}} - 3), \\ q_{k+2n,t} &= q_{k,t}, \quad k \in K, \quad t \in P_j, \quad j \in S, \quad k \leq n \times (n^{\text{TU}} - 2), \\ x_{k+2n}^{\text{end}} - x_{k+2n}^{\text{begin}} &= x_k^{\text{end}} - x_k^{\text{begin}}, \quad k \in K, \quad k \leq n \times (n^{\text{TU}} - 2). \end{aligned} \quad (22)$$

Constraints (22) are used to group and operate TUs with a cyclic pattern.

4.3.3. Strategy 3: Delay Recovery, Allowing Swaps of Tail Tracks. This strategy allows trains to swap the turnback tracks. The strategy adopts a formulation including objective function (2) and constraints (3) through (20) in addition to the following constraints:

$$t^b = \pi_{f_e}^{\text{begin}} + t^{\text{delay}}, \quad (23)$$

$$x_{f_i}^{\text{begin}} \geq \pi_{f_i}^{\text{begin}}, \quad i \in U, \quad (24)$$

$$x_{l_i}^{\text{end}} \geq \pi_{l_i}^{\text{end}}, \quad i \in U. \quad (25)$$

Constraint (23) sets the first arrival train with the delay time of t^{delay} and Constraints (24) and (25) ensure that the actual arrival or start time at the platform must be no earlier than as planned.

4.3.4. Strategy 4: Delay Recovery, Maintaining the Tail Track in Original Timetable. This strategy requires that the trains

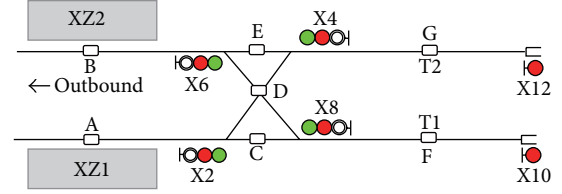


FIGURE 8: Track map of Xinzhuang (XZ) Station of line 1 in Shanghai, China.

to maintain the turnback track as in the original timetable, which means that the trains maintain the same turnback track at segment S^T . The strategy adopts a formulation including objective function (2), constraints (3) through (20), and (23) through (25) in addition to the following constraint:

$$q_{k,t} = \theta_{k,t}, \quad k \in K_j^S, \quad t \in P_j, \quad j \in S^T. \quad (26)$$

5. Application on Shanghai Rail Transit Line 1

The applicability of the proposed model for estimating the terminal capacity and delay was tested by a case study of the Xinzhuang terminal of line 1 in Shanghai, which is one of the most congested terminals of the Shanghai rail transit network. The length of this line is approximately 36.39 km, with 28 stations (including 8 transfer stations). This line uses TUs each composed of 8 high capacity vehicles (310 persons/vehicle). The minimal headway of this line is 2.75 min in peak hours and the line is operated extremely busy with a daily ridership exceeding 1 million passengers. A schematic track layout of the terminal area of this station is shown in Figure 8.

The proposed models were implemented in Visual Studio 2013 using IBM ILOG CPLEX 12.5 as a black-box MIP solver running on a personal computer with an Intel (R) Core (TM) i7-3520M CPU at 2.90 GHz and 4 GB of RAM under Windows 8 64-bit environment. The original timetable was obtained from the TPM (Train Plan Maker) software [9], which has been applied to designing operational timetables for all rail transit lines in Shanghai. The general input parameters of this terminal, as shown in Table 1, were adopted from the rail transit operation company of Shanghai.

5.1. Theoretical Capacity Analysis. As summarized in Table 2, four scenarios were tested to evaluate the proposed model under different operational conditions: these scenarios differed in three ways: (1) allocation strategy of tail tracks, (2) fixed or unfixed platform time, and (3) layover time. Other operating parameters are $n^{\text{TU}} = 8$, $n = 6$, $f_e = 1$, $l_e = 48$, and $t^b = 0$.

For every combination of scenario and layover time, 112 computation runs were carried out. The result from the experiments is shown in Figure 9.

These diagrams in Figure 9 show that both the operational schemes of tail tracks and the platform time have very significant impacts on minimum headway of arrival or departure trains. Figure 9 shows that scenario 2 and scenario

TABLE 1: Terminal train operations parameters (unit: second).

Segment track	Inbound occupation time ($d_{j,t}^{in}$)	Outbound occupation time ($d_{j,t}^{out}$)	Separation time of meeting events ($h_{j,t}^M$)	Separation time of following events ($h_{j,t}^F$)
A	30	—	—	60
B	—	30	—	60
C	45	—	—	60
D	50	50	20	60
E	—	45	—	60
F	20	20	115	60
G	20	20	115	60

TABLE 2: Computational scenarios.

Scenario ID	Operational scheme of tail tracks	Strategy	Platform time	Layover time (sec)
1	T2	1	Unfixed ($S^W = 1,3$)	200, 210, 220, ..., 480
2	T2	1	Fixed ($S^W = 3$)	200, 210, 220, ..., 480
3	T1 + T2	2	Unfixed ($S^W = 1,3$)	200, 210, 220, ..., 480
4	T1 + T2	2	Fixed ($S^W = 3$)	200, 210, 220, ..., 480

TABLE 3: Results of delay estimation.

Primary delay (s)	Strategy 3				Strategy 4			
	CPU Time	Total delay (s)	Delay trains	Delay recovery time (s)	CPU Time	Total delay	Delay trains	Delay recovery time (s)
60	24	60	1	60	5	60	1	60
120	23	120	1	120	5	120	1	120
180	24	219	3	395	5	291	8	870
240	24	435	7	715	6	895	12	1,240
300	294	876	9	1,035	5	1,935	26	2,385
360	95	1,468	11	1,095	5	3,589	31	3,065
420	277	2,191	13	1,420	5	5,863	44	4,055
480	335	2,981	15	1,570	5	8,608	46	4,115
540	795	3,986	17	1,770	5	11,368	46	4,175
600	410	5,152	21	2,090	5	14,128	46	4,235
660	869	6,450	24	2,395	5	16,888	46	4,295
720	277	7,960	26	2,440	5	19,648	46	4,355
780	1,175	9,579	28	2,760	5	22,408	46	4,415
840	848	11,359	32	3,065	5	25,168	46	4,475
900	754	13,316	33	3,113	5	27,928	46	4,535

4 result in a high minimum headway when layover time is less than 310 s. However, the minimum headway of scenario 4 is much lower than scenario 2 when layover time is more than 310 s. Figure 9 also indicates that the minimum headway reaches the lowest 116 s when layover time is 300 s at scenario 3. It is also interesting to notice that the minimum headway is not a linear relationship with layover time when layover time is less than 300 s at scenario 3, and the results do not change when layover time is less than 320 s at scenario 1.

The comparison of strategy 1 versus strategy 2 shows that unfixed platform time and flexible tail track allocation strategies can improve the capacity of turnback operation. The figure also suggests that the minimum headway could vary quite significantly even at a same layover time. Figures

9(a)–9(d) show the track occupation diagrams for the four different scenarios when the layover time is 320 s. Apparently, the fixed platform time produces longer wait times at the tail track, which makes the tail track constraining to the terminal capacity. Figure 9 also indicates that the reason for the headway to drop between 290 s and 300 s in scenario 3 is due to the identical gaps between all arrival or departure trains at platform tracks.

5.2. *Delay Analysis.* The initial timetable of line 1 was an actual Friday operation timetable in April 2014. This timetable was named 148-2. The average headway of arriving or departing trains at the Xinzhuang terminal was planned at 164 s during the morning peak. In our study, we selected

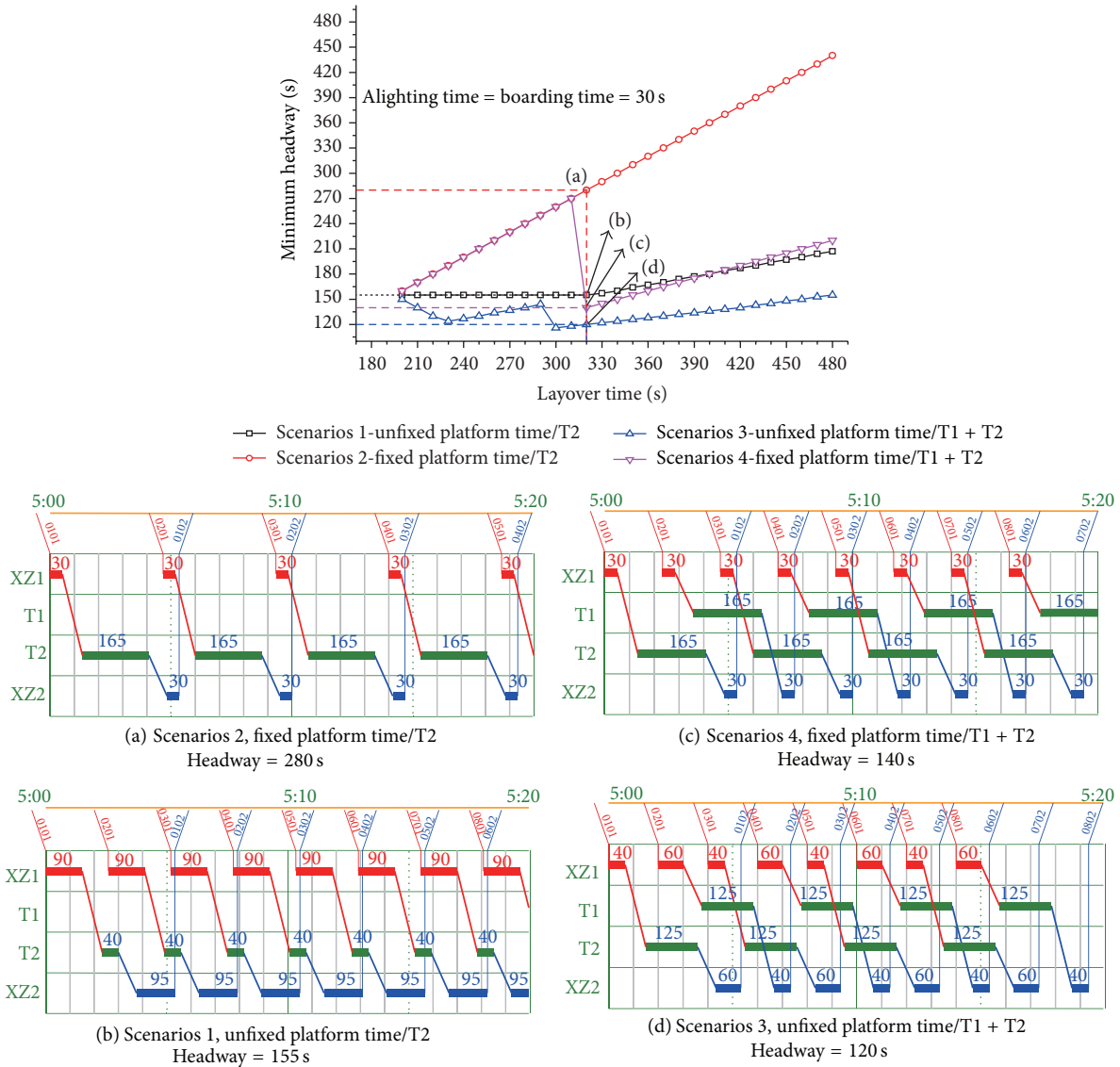


FIGURE 9: Minimum headway versus layover time and track occupation diagrams.

6:58 am to 8:06 am as the computational time horizon, which included 46 trains and 23 TUs. The minimum alighting and boarding times were 45 s and the minimal tail track time was 40 s, assuming that the platform time was not fixed and the primary delay train was the first arrival train.

Table 3 presents the results yielded by strategy 3 and strategy 4; in this table, the values for CPU time, total delay, delay trains, and delay recovery time are reported with respect to different primary delays. Figure 10 shows more details for the arrival and departure delays. The figure and table show that strategy 3 (which allows swapping of the tail tracks) has a significantly positive impact on absorbing of the delay. When the primary delay is less than 180 seconds, the trains do not suffer from knock-on delay at all, whereas the total delay increased rapidly when the primary delay increases to 240 seconds. Furthermore, the total delay of the trains is affected directly by the allocation strategy of the

tail tracks. Figures 10(a) and 10(b) show the new tail track occupation schemes obtained by the optimization results with a primary delay of 300 seconds and constraint modeling strategy 4 and strategy 3, respectively.

The choice of strategy had a noticeable effect on computation time. Strategy 3 performed longer CPU time than strategy 4 at a same primary delay. The CPU time was related to the primary delay under strategy 3, which needed very long computation times when the primary delay was more than 240 seconds. However, the CPU time was quite short and did not vary when strategy 4 was applied.

Scenario 2 in Table 2 is the most popular and practical operation in China, because it is quite simple and easy to be applied in daily operations. Actually, Scenario 3 is the best operational strategy in terminal with maximal capacity and more flexibility in operation according to the computational results.

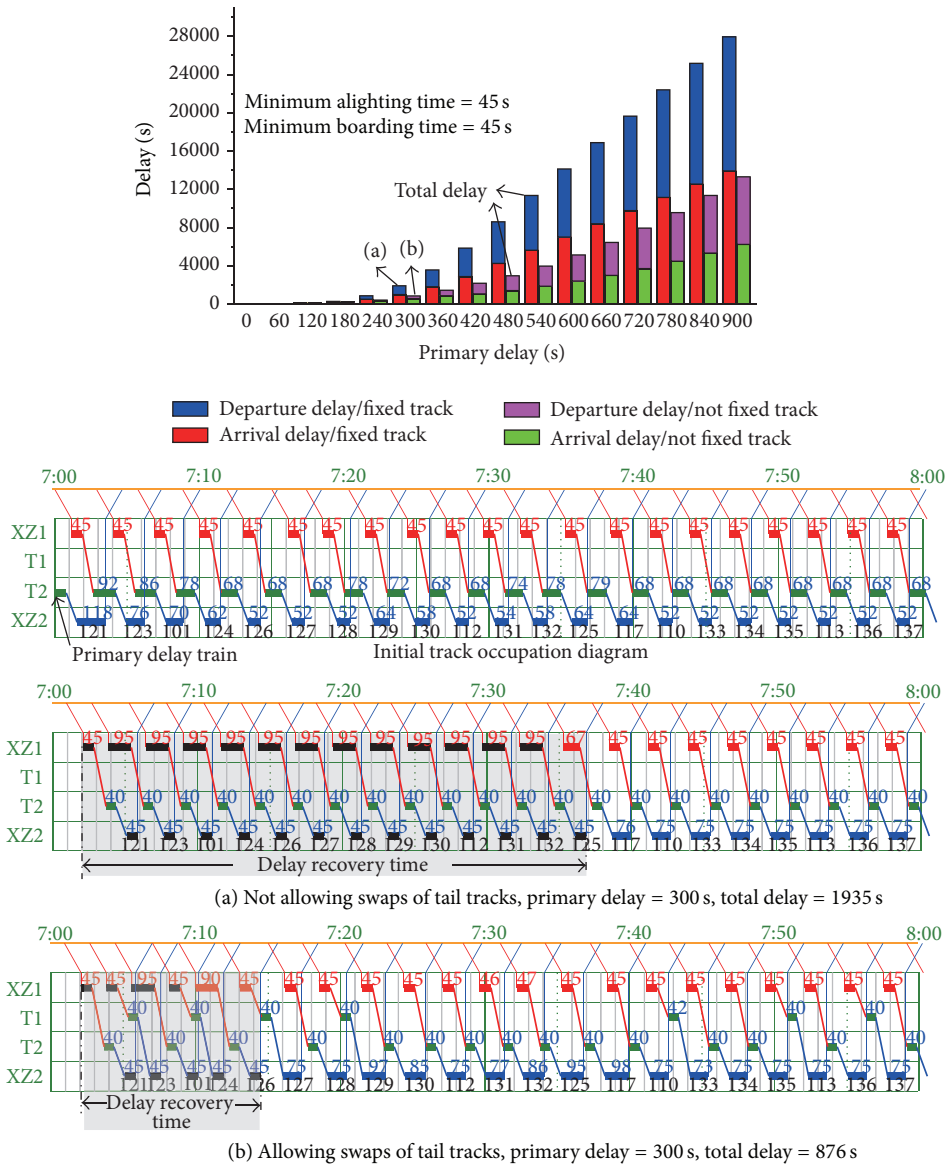


FIGURE 10: Delay index and track occupation diagram by the different primary delay.

6. Conclusions

In this paper, we have proposed a model and problem formulation for estimating turnback capacity and delay at a rail transit terminal with two-tail tracks and considered the capacity evaluation and delay propagation as an integrated optimization model with two objectives of minimizing time span and minimizing train delay. Operations parameters such as tail track allocation strategies, maximum layover time, headway pattern, buffer time distribution scheme, and primary delay were also considered in the model. Since the model built in this paper is a generic model, the method can be easily changed to adapt to the changes in track occupation time or initial timetable and can also be fit for other types of terminals (crossovers in advance of terminal with two or more platform tracks) with small changes in constraints. A case study in Shanghai rail transit line 1 illustrated that unfixed platform time and flexible tail track

allocation strategies can improve the capacity of turnback operation, and the strategy of allowing swapping of the tail tracks has a significantly positive impact on absorbing of the delay. The example also illustrated that rail transit agencies could use this proposed optimization model to compute for a reasonable capacity value with different managerial goals and operational strategies and for a delay recovery plan with high reliability.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grants no. 61473210) and the funds

from Shanghai Shentong Metro group and iRAGS of Siemens AG in Braunschweig. The acquisition of the transit operations data in the paper was supported by the Shanghai Shentong Metro Operations Management Center. The authors are grateful to these supports.

References

- [1] D. S. Lee, *Understanding capacity and performance of urban rail transit terminals [M.S. thesis]*, Massachusetts Institute of Technology, 2002.
- [2] D. C. Gill, "Assessment of mass transit turn-back capacity using dynamic simulation models," in *Proceedings of the International Conference on Computers in Railways: Computers in Railways VII*, pp. 1077–1086, Bologna, Italy, 2000.
- [3] T. Parkinson and I. Fisher, "Rail transit capacity," TCRP Report 13, Transportation Research Board, Washington, DC, USA, 1996.
- [4] Kittelson & Associates, Parsons Brinckerhoff, KFH Group, Texas A&M Transportation Institute, and Arup, *Transit Capacity and Quality of Service Manual*, TCRP Report 165, Transportation Research Board, Washington, DC, USA, 3rd edition, 2013.
- [5] N. van Oort and R. van Nes, "Impact of rail terminal design on transit service reliability," *Transportation Research Record*, vol. 2146, pp. 109–118, 2010.
- [6] M. T. Isaai, A. Kanani, M. Tootoonchi, and H. R. Afzali, "Intelligent timetable evaluation using fuzzy AHP," *Expert Systems with Applications*, vol. 38, no. 4, pp. 3718–3723, 2011.
- [7] M. Carey and I. Crawford, "Scheduling trains on a network of busy complex stations," *Transportation Research Part B: Methodological*, vol. 41, no. 2, pp. 159–178, 2007.
- [8] V. Cacchiani, A. Caprara, and P. Toth, "A Lagrangian heuristic for a train-unit assignment problem," *Discrete Applied Mathematics*, vol. 161, no. 12, pp. 1707–1718, 2013.
- [9] J. Zhibin, G. Jia, and X. Ruihua, "Circle rail transit line timetable scheduling using Rail TPM," in *Proceedings of the 12th International Conference on Computer System Design and Operation in the Railways and other Transit Systems (COMPRAIL '10)*, pp. 945–952, Beijing, China, September 2010.
- [10] Y. Lee and C.-Y. Chen, "A heuristic for the train pathing and timetabling problem," *Transportation Research Part B: Methodological*, vol. 43, no. 8-9, pp. 837–851, 2009.
- [11] H. Niu and X. Zhou, "Optimizing urban rail timetable under time-dependent demand and oversaturated conditions," *Transportation Research Part C: Emerging Technologies*, vol. 36, pp. 212–230, 2013.
- [12] Ö. Yalçinkaya and G. Mirac Bayhan, "A feasible timetable generator simulation modelling framework for train scheduling problem," *Simulation Modelling Practice and Theory*, vol. 20, no. 1, pp. 124–141, 2012.
- [13] A. Carrel, R. G. Mishalani, N. H. M. Wilson, J. P. Attanucci, and A. B. Rahbee, "Decision factors in service control on high-frequency metro line: importance in service delivery," *Transportation Research Record*, vol. 2146, pp. 52–59, 2010.
- [14] Z.-B. Jiang, F. Li, R.-H. Xu, and P. Gao, "A simulation model for estimating train and passenger delays in large-scale rail transit networks," *Journal of Central South University*, vol. 19, no. 12, pp. 3603–3613, 2012.
- [15] J. D. Schmöcker, S. Cooper, and W. Adeney, "Metro service delay recovery: comparison of strategies and constraints across systems," *Transportation Research Record*, no. 1930, pp. 30–37, 2005.
- [16] A. Puong and N. H. M. Wilson, "A train holding model for urban rail transit systems," in *Computer-Aided Systems in Public Transport*, vol. 600 of *Lecture Notes in Economics and Mathematical Systems*, pp. 319–337, Springer, Berlin, Germany, 2008.
- [17] H. Wang, N. Zhao, and L. Chen, "An integrated capacity evaluation method for CBTC-system-equipped urban rail lines," *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, pp. 1–12, 2013.
- [18] T. Dollevoet, D. Huisman, L. Kroon, M. Schmidt, and A. Schöbel, "Delay management including capacities of stations," *Transportation Science*, pp. 1–19, 2014.
- [19] V. Cacchiani, D. Huisman, M. Kidd et al., "An overview of recovery models and algorithms for real-time railway rescheduling," *Transportation Research, Part B: Methodological*, vol. 63, pp. 15–37, 2014.
- [20] M. Flamini and D. Pacciarelli, "Real time management of a metro rail terminus," *European Journal of Operational Research*, vol. 189, no. 3, pp. 746–761, 2008.
- [21] C. Mannino and A. Mascis, "Optimal real-time traffic control in metro stations," *Operations Research*, vol. 57, no. 4, pp. 1026–1039, 2009.
- [22] J. Törnquist and J. A. Persson, "N-tracked railway traffic rescheduling during disturbances," *Transportation Research Part B: Methodological*, vol. 41, no. 3, pp. 342–362, 2007.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

