Hindawi
Journal of Advanced Transportation
Volume 2017, Article ID 3097681, 17 pages
https://doi.org/10.1155/2017/3097681

*Research Article*

# Optimized Skip-Stop Metro Line Operation Using Smart Card Data

**Peitong Zhang, Zhanbo Sun, and Xiaobo Liu**

*School of Transportation and Logistics, Southwest Jiaotong University, Chengdu, China*

Correspondence should be addressed to Xiaobo Liu; xiaobo.liu@swjtu.cn

Received 4 June 2017; Accepted 14 August 2017; Published 13 November 2017

Academic Editor: Giulio E. Cantarella



Skip-stop operation is a low cost approach to improving the efficiency of metro operation and passenger travel experience. This paper proposes a novel method to optimize the skip-stop scheme for bidirectional metro lines so that the average passenger travel time can be minimized. Different from the conventional "A/B" scheme, the proposed Flexible Skip-Stop Scheme (FSSS) can better accommodate spatially and temporally varied passenger demand. A genetic algorithm (GA) based approach is then developed to efficiently search for the optimal solution. A case study is conducted based on a real world bidirectional metro line in Shenzhen, China, using the time-dependent passenger demand extracted from smart card data. It is found that the optimized skip-stop operation is able to reduce the average passenger travel time and transit agencies may benefit from this scheme due to energy and operational cost savings. Analyses are made to evaluate the effects of that fact that certain number of passengers fail to board the right train (due to skip operation). Results show that FSSS always outperforms the all-stop scheme even when most passengers of the skipped OD pairs are confused and cannot get on the right train.

## 1. Introduction

The time required for stops at stations is usually a major travel time component for urban transit systems. Cortés et al. [1] discussed different approaches and operation strategies that can reduce total passenger travel time. With the goal of reducing travel time and increasing overall transit operation efficiency, skip-stop operation is a scheme that can have fleets not making all designated stops along a route. Skip-stop operation has been widely applied in bus transit based on the A/B mode. Under the A/B mode, operating fleets are labeled as either type A or type B; stations are also predetermined as type A, type B, or type AB stations. Type A fleets can only stop at type A and type AB stations and type B fleets can only stop at type B and type AB stations [2, 3]. Eberlein et al. [4] proposed a real-time deadheading strategy to determine the dispatching time. Deadheading vehicles and skip operations were optimized to minimize the total passenger cost. Sun and Hickman [5] formulated the real-time skip-stop problem for bus operation and presented an enumeration method with good computation speed. Yu et al. [6] studied the service reliability of a bus route in the city of Dalian and proposed a heuristic algorithm-based deadheading optimization strategy. Fu et al. [7] presented a new strategy to have all-stop and skip-stop trains operate in pairs. The interests of operators and passengers were balanced in the model by using an exhaustive search procedure. Skip-stop operation is applied in practice by the Transmilenio system in Bogota and Metro Rapid system in Los Angeles for bus operation and has been proved to be highly effective according to the US National Research Council [8] and Leiva et al. [9].

For urban metro systems, the A/B mode-based skip-stop strategy was first developed for the metro systems in Chicago city since 1948 [10, 11]. It was later implemented in Philadelphia and the City of New York [12]. However, due to budget constraints and large train headways caused by operating different types of trains, it switched back to the all-stop mode after several years. A/B mode has also been implemented in the metro system of Santiago, Chile, since 2007, which has demonstrated a significant benefit to passengers and transit operators [10]. Lee et al. [13] and Jamili and Aghaee [14] reported that, under skip-stop operation, passengers experienced shorter travel time and operators benefited from the scheme by having reduced fleet size and operational costs.

In addition, skip-stop operation has been commonly used as a recovery strategy in case of special events such as demand surges and major disruptions according to Cao et al. [15].

With respect to the use of skip-stop metro operation, there are some challenges that have not been fully addressed in the existing literature. The first is that since overtaking is prohibited on single-track metro lines, a larger departure headway is needed for the case that the target train travels faster (due to skip operation) than the previous train. This may also reduce the line capacity and cause some difficulty in timetable design. The second challenge is brought by the infrastructure constraints at the terminal stations, where trains need to wait and/or be stored at the turnaround terminal before serving the backward direction. Therefore the safe and efficient operation at the turnaround terminals needs to be properly considered. The third challenge is the precise estimation of dynamic passenger demand pattern, which significantly affects the design of skip-stop scheme. Last but not least, skip-stop operation usually leads to passenger confusion. As a result, certain number of passengers may fail to board the right train.

To better address the first challenge, recent research interest has shifted from the conventional A/B skip-stop mode to Flexible Skip-Stop Schemes (FSSS). A summary and comparison of the existing skip-stop schemes for metro operation is provided in Table 1. Sun and Hickman [5], Zheng et al. [16], Freyss et al. [10], Lee et al. [13], and Cao et al. [15] used different approaches to optimize A/B skip-stop scheme for a single direction urban transit line based on assumed origin-destination (OD) information. Sogin et al. [17] developed a mixed integer program that could be solved by commercial solvers for small scale problems. A genetic algorithm (GA) based solution approach was developed to solve larger scale problems with more stations and trains considered. Niu et al. [19] optimized a skip-stop scheme based on predetermined train stop plans; a mixed integer nonlinear programming model was constructed and solved by a GAMS solver in their paper. Wang et al. [18] developed optimization models that offer real-time operations with Flexible Skip-Stop Schemes; that is, trains are allowed to adjust their skip-stop patterns in real-time to minimize passenger travel time and energy consumption. Jamili and Aghaee [14] built a model to determine the "uncertain skip patterns" which means that trains can flexibly skip or stop at stations along the line so that the train operation speed is optimized.

Another often overlooked challenge is the operation at turnaround terminals. The only work that considers turnaround operation is Wang et al. [18], which treated the turnaround terminal as an intermediate station. The authors believe it is more realistic to consider turnaround operation with departure time optimization and storage capacity. This enables more flexible bidirectional operation that can lead to more efficient line utilization. For the third challenge, recent research has shown that transit smart card data can be utilized to better interpret the temporal and spatial dynamics of passenger demand [20, 21]. Niu et al. [19] used time-varying passenger demand obtained from the high-speed rail system. For metro operation, researchers (e.g., Sogin et al. [17]; Wang et al. [18]; Jamili and Aghaee [14]) usually assume uniform passenger arrival rates. Little work has been done using dynamic passenger demand from real operation data. With respect to the last challenge, although it is widely recognized as a major drawback of skip-stop operation, the effect of "fail to board the right train" has often been overlooked and little relevant analysis has been reported in literature.

In this paper, we develop a FSSS for bidirectional metro operation during off-peak periods. The problem is formulated as a mixed integer nonlinear programming (MINLP) model that minimizes the average passenger travel time. The proposed scheme is able to consider the train operation at the turnaround terminal with departure time optimization and storage capacity. The departure times at the two terminal stations are constrained by three types of departure intervals, that is, the minimum safety headway, headway that meets the passenger demand, and the turnaround headway. A GA-based solution approach is then developed to efficiently solve the problem. The proposed scheme is tested using a real metro line in Shenzhen, China. The case study considers time-dependent demand information inferred from the smart card data of the metro system. Compared to the all-stop scheme, we find that the proposed FSSS can reduce the average passenger travel time. Transit agencies may also benefit from this scheme due to energy and operational cost savings. For the cases that certain percentage of passengers fail to board the right train, results show that FSSS always outperforms the all-stop scheme even when most passengers of the skipped OD pairs are confused and cannot get on the right train. The remainder of this paper is organized as follows. Section 2 presents the FSSS model formulation. Section 3 proposes a GA-based solution approach to efficiently solve the model. The experiment and numerical results are summarized in Section 4, followed by the concluding remarks in Section 5.

## 2. FSSS Timetabling

In this section, we first introduce the notations, followed by a brief description of bidirectional train operation problem. We then introduce the proposed FSSS model formulation.

*2.1. Notations*

$k$: train index;

$i$, $j$: station indices;

$g$: time period index.

*Input Parameters*

$N_s$: number of stations of a one-way line;

$N_{\text{tr}}$: number of trains;

$l$: length of each time period;

$t_{\text{tr}}^{i}$: section travel time between stations $i$ and $i+1$, in seconds;

$t_{\text{dw}}^{i}$: dwell time at station $i$, in seconds;

$r_{g}^{i,j}$: time-dependent passenger's OD from stations $i$ to $j$ during period $g$, in people per second;

$\delta$: time loss during a stop, in seconds;

Table 1: Existing skip-stop schemes for metro operation.

| Publication | Skip-stop mode | Input data type/source | Single/bidirectional line | Objective | Model | Solution approach | Results |
|---|---|---|---|---|---|---|---|
| Suh et al. [12] | Fixed A/B | Static OD | Single | Maximize the total time saving of a skip-stop schedule in both peak and off-peak period | Train operation simulation model | Calculate the total travel time of five passenger boarding types | Per the specific scenario, the skip-stop operation results in less total travel saving in the peak period compared to the off-peak period |
| Zheng et al. [16] | Fixed A/B | Static OD with 13 stations | Single | Minimize total passenger travel time | A 0-1 integer programming model by analyzing the passenger traveling time | Tabu search | It is better to assign a type AB train between a type A train and type B train |
| Freyss et al. [10] | Fixed A/B | Assumed static OD from Chilean metro data | Single | Minimize the cost function that is comprised of total passenger transfer time, waiting time, and total traction and energy cost | Consider five passenger types under different combination of A or B or AB stations | Continuous approximation approach | Short lines with fewer stations are less favorable for skip-stop operation; larger benefit can be obtained in the lines with smaller minimum headway |
| Lee et al. [13] | Fixed A/B | Manipulated static OD based on boarding and alighting ratio from the Seoul Metro line | Single | Minimize the passenger travel time | Consider three types (A, B, and AB) of passenger OD and collision constraints in four scenarios | Genetic algorithm | 17%–20% travel time saving for passengers and reduced operational cost for the operator |
| Cao et al. [15] | Fixed A/B | Simulated static OD based on Beijing Metro data | Single | Biobjective formulation that minimizes the passenger waiting time and trip time | A 0-1 mixed integer model that considers the constraints of two successive trains; train operation simulation | Tabu search | Smaller headway leads to better passenger waiting time and trip time |
| Sogin et al. [17] | Dynamic | Simulated static OD using gravity model/Chicago Metro lines | Single | Minimize the total passenger travel time | A mixed integer formulation that considers the tradeoff between passenger travel time and service frequency caused by skip-stop | GAMS and CPLEX for small scale problems; genetic algorithm for large-scale problems | Skip-stop operation can reduce passenger travel time by 9.5% |

Table 1: Continued.

| Publication | Skip-stop mode | Input data type/source | Single/bidirectional line | Objective | Model | Solution approach | Results |
|---|---|---|---|---|---|---|---|
| Wang et al. [18] | Dynamic | Static OD based on real data from the Beijing Metro (Yizhuang) Line | Bidirectional; consider turnaround terminal as an intermediate station without storage capacity | Minimize the total passenger travel time and energy consumption | Consider the impact of boarding passenger number on the train dwell time and train utilization at the departure terminal | Bilevel mixed integer programming model with rolling horizon | Time and energy saving up to 15% |
| Niu et al. [19] | Predetermined skip-stop | Time-dependent demand for nine stations of the Shanghai-Hangzhou High-Speed Rail Line in China | Single | Minimize the total passenger waiting time | A unified quadratic integer programming model with linear constraints; consider safe operation and train capacity | GAMS solver | The model can be solved by standard commercial optimization packages with good computation time |
| Jamili and Aghaee [14] | Uncertain skip-stop | Static OD based on an Iranian metro line | Single | Minimize the total passenger travel time | A nonlinear programming model that considers boarding passengers and train capacity | Decomposition-based algorithm and simulated annealing (SA) algorithm | Trip time is reduced from 92.3 minutes to 86.6 minutes; the average speed increases by 3.4 km/h |
| Our approach | Flexible | Time-dependent arrival rate based on smart card data from the Shenzhen Metro line | Bidirectional; consider departure time optimization and storage capacity at the turnaround terminal | Minimize the average travel time | Three types of headways are considered; train operation simulation; turnaround operation | Genetic algorithm | See numerical section |

$I_{\min}$: minimum headway for safety operation, in seconds;

$I_d$: headway that meets the passenger demand, in seconds;

$I_t$: minimum turnaround headway, in seconds;

$I_p$: maximum acceptable passenger waiting time, in seconds;

$c$: number of lines that allow trains to be stored at the turnaround terminal;

SN: set contains all the stations along a line;

$n_{\mathrm{ps}}$: number of the solutions selected at the first step of the GA;

$r_c$: crossover rate;

$r_m$: mutation rate;

$M$: maximum iteration number;

$n_c$: number of initial chromosomes;

$r_f$: confusion rate.

*Variables*

$a_{k,i}$: Operation state of train $k$ at the station $i$; it equals 1 if train $k$ stops at station $i$; otherwise it is zero;

$d_{k,i}$: departure time of train $k$ at station $i$, in seconds;

$c_i^k$: train stopped at station $i$ prior to the arrival of train $k$;

$C_i^k$: the set of trains between trains $k$ and $c_i^k$ that stop at station $i$ and skip station $j$;

$i_k$: the next stopped station of train $k$ after departure from station $i$;

$k_e$: the next available train after train $k$ that can transport the confused passengers alighted at $i_k$ to $j$;

$np_k^{i,j}$: number of passengers who board train $k$ at station $i$, with a destination $j$;

$\mathrm{tw}_k^{i,j}$: total waiting time associated with passengers $np_k^{i,j}$;

$\mathrm{tv}_k^{i,j}$: total in-vehicle time associated with passengers $np_k^{i,j}$;

$n_{\mathrm{pg}}^{k,i}$: total number of onboard passengers for train $k$ at station $i$;

$t_{\mathrm{wait}}^{k,i}$: total waiting time for passengers who board train $k$ at station $i$;

$t_{\mathrm{veh}}^{k,i}$: total in-vehicle time for passengers who board train $k$ at station $i$;

$n_{\mathrm{pg}}$: number of onboard passengers;

$t_{\mathrm{wait}}$: total passenger waiting time, in seconds;

$t_{\mathrm{veh}}$: total passenger in-vehicle time, in seconds;

SS: the candidate skipping station set that contains the stations that can be potentially skipped;

$S_k$: the set that records the skip-stop schemes of train $k$ for the forward direction;

$S_k'$: the set that records the skip-stop schemes of train $k$ for the backward direction;

$m_s^k$: starting station of arc $s$ of train $k$;

$n_s^k$: ending station of arc $s$ of train $k$;

$\mathrm{arc}_{m_s^k,n_s^k}$: skip arc that connects two stations ($m_s^k$ and $n_s^k$), between which at least one station is skipped;

$\Delta t_{m_s^k,n_s^k}$: maximum travel time difference between $m_s^k$ and $n_s^k$;

$\Delta_k$: maximum travel time difference between train $k-1$ and train $k$ for the forward direction;

$\Delta_k'$: maximum travel time difference between train $k-1$ and train $k$ for the backward direction.

*Outputs*

$t_{\mathrm{wait}}^{\mathrm{pg}}$: average waiting time per passenger, in minutes;

$t_{\mathrm{veh}}^{\mathrm{pg}}$: average in-vehicle time per passenger, in minutes;

$t_{\mathrm{tv}}^{\mathrm{pg}}$: average travel time per passenger which equals $t_{\mathrm{wait}}^{\mathrm{pg}} + t_{\mathrm{veh}}^{\mathrm{pg}}$, in minutes;

$Z$: objective value;

$n_{\mathrm{ss}}$: number of skipped stations;

$n_{\mathrm{bpg}}$: total number of the passengers whose in-vehicle time is reduced by skip-stop operation.

*2.2. Bidirectional Metro Operation Problem.* Consider a bidirectional metro line. The system is comprised of $N_{\mathrm{tr}}+1$ trains and $N_s$ stations, indexed from 0 to $N_{\mathrm{tr}}$ and 1 to $N_s$, respectively. Train 0 is an all-stop train that departs at time 0; this all-stop train is operated to ensure the system is in operation at the beginning of the optimization period. Station 1 is the departure terminal and station $N_s$ is the turnaround terminal. The line has only one track in each direction as shown in Figure 1. To avoid confusion, stations are denoted as $1, 2, \ldots, N_s, N_s+1, \ldots, 2N_s$, in which stations 1 ($2N_s$) indicate departure terminal; stations $N_s$ and $N_s+1$ represent the turnaround terminal; $\{1, 2N_s\}, \{2, 2N_s-1\}, \ldots, \{N_s, N_s+1\}$ are used to represent the same stations, which share the same dwell time. The section travel times between two adjacent stations are assumed to be the same for both directions. In real practices, there are usually more than one turnaround line at the turnaround terminal where trains are allowed to be stored on the lines. As illustrated in Figure 1, train $k-1$ up to train $k-c-1$ are allowed to wait at turnaround terminal before train $k-c$ departs from station $N_s+1$. A proof is given in Appendix to show that more turnaround lines (e.g., $c \geq 1$) can lead to better train utilization and therefore increase the line capacity. The solution of the metro operation problem is to make skip/stop decisions and determine the departure time at each station. Thereby a complete operation timetable can be obtained.

The authors notice that there is a debate regarding whether skip-stop is more suitable for off-peak or peak period operation. For example, Niu et al. [19] built mathematical models to solve the rail optimization problem during peak period. As reported in Suh et al. [12], the benefits received in off-peak period (due to skip-stop) are more significant compared to the peak period. The results are largely due to the fact that there are larger passenger demands at most stations during the peak period. The skip operation may lead to excessive waiting time at the skipped stations and it may also cause station congestion. In this work, the authors consider the off-peak operation. The train capacity is not considered as a constraint since our early study [22] found out that trains are far from crowded during the off-peak period.

*2.3. Assumptions.* The key assumptions used in this paper are summarized and discussed as below.

*Assumption 1.* The proposed method targets offline timetabling design and the passenger information (e.g., OD) is assumed to be known and can be accurately calibrated using smartcard data collected from their day-to-day travel activities.

*Assumption 2.* The number of the supplied trains and the storage capacity at the departure terminal are unlimited.

*Assumption 3.* All stations can only accommodate one train at a time for one direction; overtaking is prohibited at any point of the line. This assumption is consistent with the common infrastructure setup at most metro systems in China.

*Assumption 4.* The travel time between two adjacent stations and the dwell time at each station varies along the line but they are assumed to be fixed for all trains. The time loss associated with a stop (due to braking and acceleration) is assumed to be a constant for all trains and all stops.

*Assumption 5.* It is assumed that if train $k$ stops at station $i$ and is about to skip station $j$, certain percentage of passengers from OD pairs $i$ to $j$ will mistakenly board the train (this percentage is referred to as the confusion rate). It is assumed

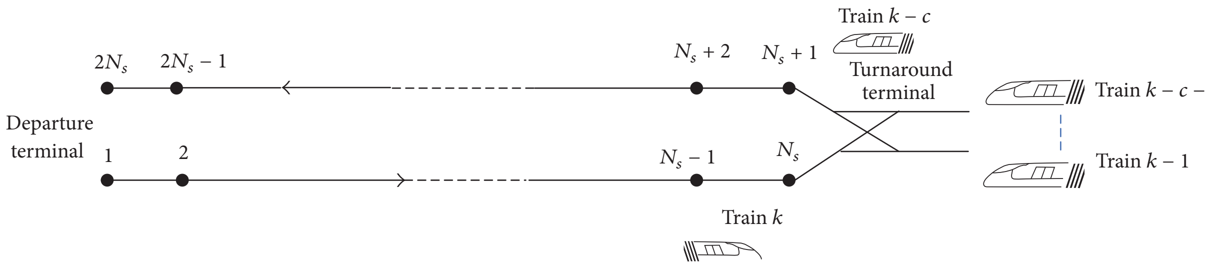



FIGURE 1: Illustration of a bidirectional metro system.

that these passengers shall notice their mistake (e.g., via train radio broadcast) once they get on the train, and they will alight at the next stopping station. These passengers will then wait at the station until being transported by the next available (correct) train to their destination $j$.

*2.4. Objective Function.* The objective of FSSS is to determine the operation scheme and departure times to minimize the average passenger travel time including the in-vehicle travel time and waiting time. The operation for train $k$ at station $i$ is defined as a binary variable $a_{k,i}$.

$$a_{k,i} = \begin{cases} 1, & \text{if train } k \text{ stops at the station } i \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

The total number of onboard passengers $n_{\mathrm{pg}}$, the total passenger waiting time $t_{\mathrm{wait}}$, and the total passenger in-vehicle time $t_{\mathrm{veh}}$ can be calculated using (2), in which $n_{\mathrm{pg}}^{k,i}$, $t_{\mathrm{wait}}^{k,i}$, and $t_{\mathrm{veh}}^{k,i}$ are the number of onboard passengers, the passenger waiting time, and the passenger in-vehicle time for train $k$ at station $i$, respectively.

$$\begin{aligned} n_{\mathrm{pg}} &= \sum_{k=1}^{N_{\mathrm{tr}}} \sum_{i=1}^{2N_s} n_{\mathrm{pg}}^{k,i}, \\ t_{\mathrm{wait}} &= \sum_{k=1}^{N_{\mathrm{tr}}} \sum_{i=1}^{2N_s} t_{\mathrm{wait}}^{k,i}, \\ t_{\mathrm{veh}} &= \sum_{k=1}^{N_{\mathrm{tr}}} \sum_{i=1}^{2N_s} t_{\mathrm{veh}}^{k,i}. \end{aligned} \tag{2}$$

We use $d_{k,i}$ to denote the departure time of train $k$ at station $i$. The dwell time at station $i$ is $t_{\mathrm{dw}}^{i}$, which is determined based on the real metro operation data. The arrival time at station $i$ can be easily acquired by subtracting $t_{\mathrm{dw}}^{i}$ from $d_{k,i}$. Variable $c_i^k$ is used to denote the train that stopped at station $i$ prior to the arrival of train $k$, which can be referred to as $c_i^k = \max\{k' \mid k' < k,\ a_{k,i} = a_{k',i} = 1\}$ according to Niu et al. [19].

The time-dependent passenger OD rate between stations $i$ and $j$ during time period $g$ is denoted as $r_g^{i,j}$ and the length of each period is $\tau$. To calculate the values of $n_{\mathrm{pg}}^{k,i}$, $t_{\mathrm{wait}}^{k,i}$, and $t_{\mathrm{veh}}^{k,i}$, two cases are considered; see Figure 2. The first case is when trains $k$ and $c_i^k$ depart in the same time period (i.e., $d_{k,i}, d_{c_i^k,i} \in [(g-1)\tau, g\tau)$). In this case, the passenger arrival rate is uniform. The resulting $n_{\mathrm{pg}}^{k,i}$, $t_{\mathrm{wait}}^{k,i}$, and $t_{\mathrm{veh}}^{k,i}$ can be calculated as in (3)–(5). Here $d_{k,i} - d_{c_i^k,i}$ represents the departure interval between trains $k$ and $c_i^k$ at station $i$; therefore $r_g^{i,j}(d_{k,i} - d_{c_i^k,i})$ represents the amount of passengers who arrive at station $i$ and plan to travel to station $j$ during this time interval. $d_{k,j} - d_{k,i} - t_{\mathrm{dw}}^{j}$ refers to the in-vehicle travel time for passengers travel from $i$ to $j$. The second case is when trains $k$ and $c_i^k$ depart in different time periods (i.e., $d_{c_i^k,j} \in [(g-1)\tau, g\tau) \wedge d_{k,i} \in [g\tau, (g+s)\tau)\ \forall s \geq 1$). The corresponding $n_{\mathrm{pg}}^{k,i}$, $t_{\mathrm{wait}}^{k,i}$, and $t_{\mathrm{veh}}^{k,i}$ can be calculated as in (6)–(8). The objective function is then formulated to minimize the average passenger travel time; see formulation (9).

$$n_{\mathrm{pg}}^{k,i} = \sum_{j=i+1}^{2N_s} a_{k,i} a_{k,j} r_g^{i,j} \left(d_{k,i} - d_{c_i^k,i}\right), \tag{3}$$

$$t_{\mathrm{wait}}^{k,i} = \frac{1}{2} n_{\mathrm{pg}}^{k,i} \left(d_{k,i} - d_{c_i^k,i}\right)^2, \tag{4}$$

$$t_{\mathrm{veh}}^{k,i} = \sum_{j=i+1}^{2N_s} a_{k,i} a_{k,j} r_g^{i,j} \left(d_{k,i} - d_{c_i^k,i}\right)\left(d_{k,j} - d_{k,i} - t_{\mathrm{dw}}^{j}\right), \tag{5}$$

$$\begin{aligned} n_{\mathrm{pg}}^{k,i} = \sum_{j=i+1}^{2N_s} a_{k,i} a_{k,j} \Bigg[ & r_g^{i,j}\left(g\tau - d_{c_i^k,i}\right) + \sum_{t=g+1}^{g+s-1} \tau r_t^{i,j} \\ & + r_{g+s}^{i,j}\left(d_{k,i} - (g+s-1)\tau\right) \Bigg], \end{aligned} \tag{6}$$

$$t_{\mathrm{wait}}^{k,i} = \frac{1}{2} n_{\mathrm{pg}}^{k,i} \left(d_{k,i} - d_{c_i^k,i}\right)^2, \tag{7}$$

$$\begin{aligned} t_{\mathrm{veh}}^{k,i} = \sum_{j=i+1}^{2N_s} a_{k,i} a_{k,j} \Bigg[ & r_g^{i,j}\left(g\tau - d_{c_i^k,i}\right) + \sum_{t=g+1}^{g+s-1} \tau r_t^{i,j} \\ & + r_{g+s}^{i,j}\left(d_{k,i} - (g+s-1)\tau\right) \Bigg] \left(d_{k,j} - d_{k,i} - t_{\mathrm{dw}}^{j}\right), \end{aligned} \tag{8}$$

$$\min Z = \frac{\left(t_{\mathrm{wait}} + t_{\mathrm{veh}}\right)}{n_{\mathrm{pg}}}. \tag{9}$$

*2.5. Model Constraints.* The model is subject to a few safety, operational, and infrastructure constraints. Train headway is an important concept that is directly related to operation

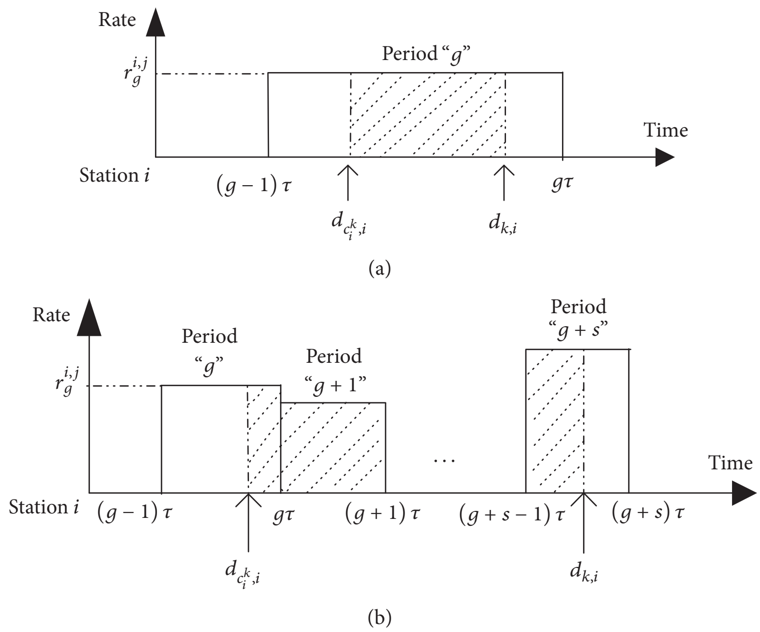



FIGURE 2: The departure of two consecutive trains; (a) trains $k$ and $c_i^k$ depart in the same time period; (b) trains $k$ and $c_i^k$ depart in different time periods.

safety and the capacity of a metro line. Three types of train headways are considered in this study, which are the minimum safety headway $I_{\min}$, the headway that meets the passenger demand $I_d$, and the minimum turnaround headway $I_t$. Different from the work in Wang et al. [18] who took the turnaround terminal as a normal station, we specifically consider the departure time optimization and storage capacity at the turnaround station. In this subsection, the headway-related constraints are discussed below, followed by the operational constraints.

*2.5.1. Minimum Safety Headway.* Equations (10) and (11) represent that train 0 departures at time 0; after arriving at the turnaround terminal, the new departure time of train 0 is calculated by adding the section travel time, station dwell time, and the turnaround time. For safe operation, the departure interval of two consecutive trains should satisfy formulation (12) at all stations.

$$d_{0,1} = 0, \tag{10}$$

$$d_{0,N_s+1} = \sum_{r=1}^{N_s} \left(t_{\mathrm{tr}}^i + t_{\mathrm{dw}}^i\right) + I_t, \tag{11}$$

$$d_{k,i} - d_{k-1,i} \geq I_{\min} \quad \forall 1 \leq i \leq 2N_s,\ 1 \leq k \leq N_{\mathrm{tr}}. \tag{12}$$

Note that skip operation leads to shorter travel time and it may cause some safety concerns. An illustrative example is provided in Figure 3. Suppose train 1 stops at all stations and train 2 skips station 2 and station 3. If train 2 only keeps the minimum safety headway $I_{\min}$ at the departure terminal, the minimum safety headway will not suffice at station 3. This situation poses safety issues and requires further intervention control. To avoid this, an extra amount of time $\Delta_k$ is added to the departure time (of train $k$) at the departure terminal; see the trajectory of train 2.

To calculate the value of $\Delta_k$, set $S_k$ and set $S_k'$ are defined as the skip-stop plans of train $k$ for the forward direction and backward direction, respectively. Each set is comprised of a series of the "skip arcs"; each "skip arc" connects two stations where train $k$ stops, between which at least one station has been skipped. For example, as shown in Figure 4, $\mathrm{arc}_{m_1^k,n_1^k}$ denotes the fact that train $k$ stops at station $m_1^k = 1$ and stops again at station $n_1^k = 4$, between which station 2 and station 3 are skipped. Here $m_s^k$ refers to the starting station of arc $s$; $n_s^k$ refers to the ending station of arc $s$. The last "skip arc" in set $S_k$ is denoted as $\mathrm{arc}_{m_q^k,n_q^k}$.

$$S_k = \left\{\mathrm{arc}_{m_1^k,n_1^k}, \ldots, \mathrm{arc}_{m_s^k,n_s^k}, \ldots, \mathrm{arc}_{m_q^k,n_q^k}\right\}. \tag{13}$$

Under this definition, $\sum_{i=m_s^k}^{n_s^k}(1-a_{k,i})$ can be used to record the total number of skip operations of train $k$ from stations $m_s^k$ to $n_s^k$ with respected to $\mathrm{arc}_{m_s^k,n_s^k}$. The travel time saving of train $k$ from stations $m_s^k$ to $n_s^k$ can be expressed as $\sum_{i=m_s^k}^{n_s^k}(1-a_{k,i})(t_{\mathrm{dw}}^i+\delta)$, where $t_{\mathrm{dw}}^i$ is the dwell time at station $i$ and $\delta$ is the time loss due to the acceleration and deceleration associated with a stop at a station. Therefore, the maximum travel time difference between train $k-1$ and train $k$ from stations $m_s^k$ and $n_s^k$, denoted as $\Delta t_{m_s^k,n_s^k}$, can be represented as

$$\Delta t_{m_s^k,n_s^k} = \sum_{i=m_s^k}^{n_s^k} \left(1-a_{k,i}\right)\left(t_{\mathrm{dw}}^i + \delta\right)$$

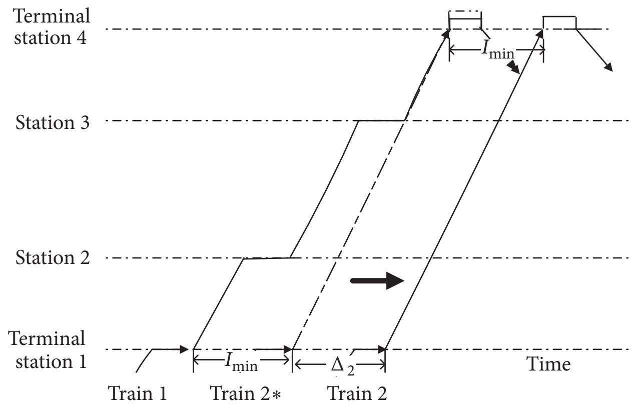



Figure 3: The illustration of the method to obtain the train safety headway.

$$- \sum_{i=m_s^k}^{n_s^k} \left(1 - a_{k-1,i}\right)\left(t_{\mathrm{dw}}^i + \delta\right)$$

$$= \sum_{i=m_s^k}^{n_s^k} \left(a_{k-1,i} - a_{k,i}\right)\left(t_{\mathrm{dw}}^i + \delta\right). \tag{14}$$

The next step is to calculate the maximum travel time difference between train $k-1$ and train $k$ at any point along the metro line, which can be expressed as $\max\{\sum_{u=1}^{s}\sum_{i=m_u^k}^{n_u^k}(a_{k-1,i} - a_{k,i})(t_{\mathrm{dw}}^i + \delta) \mid s = 1, \ldots, q\}$. In the cases that this maximum travel time difference is larger than zero (i.e., train $k$ travels faster than train $k-1$), an extra amount of time is added to the departure time of train $k$ to avoid collision. In the case that the maximum travel time difference is smaller than zero, no adjustment is needed; see formulation (15). Similarly, we use $\Delta_k'$ to represent the maximum travel time difference for the backward direction.

$$\Delta_k = \max\left\{\max\left\{\sum_{u=1}^{s}\sum_{i=m_u^k}^{n_u^k}\left(a_{k-1,i} - a_{k,i}\right)\left(t_{\mathrm{dw}}^i + \delta\right) \mid s = 1, \ldots, q\right\}, 0\right\}. \tag{15}$$

To ensure minimum safety headway between two consecutive trains at any point of the line, the departure time of train $k$ is constrained at terminal stations 1 and $N_s + 1$ by formulation (16a) and formulation (16b), respectively.

$$d_{k,1} - d_{k-1,1} \geq I_{\min} + \Delta_k, \quad \forall 1 \leq k \leq N_{\mathrm{tr}}, \tag{16a}$$

$$d_{k,N_s+1} - d_{k-1,N_s+1} \geq I_{\min} + \Delta_k', \quad \forall 1 \leq k \leq N_{\mathrm{tr}}. \tag{16b}$$

*2.5.2. Headway That Meets the Passenger Demand.* A strict constraint that satisfies the passenger demand should take the form of constraints (17a) and (17b). Here $I_d$ is the headway that meets the passenger demand.

$$d_{k,1} \leq kI_d, \quad \forall 1 \leq k \leq N_{\mathrm{tr}}, \tag{17a}$$

$$d_{k,N_s+1} \leq d_{0,N_s+1} + kI_d, \quad \forall 1 \leq k \leq N_{\mathrm{tr}}. \tag{17b}$$

This constraint, however, is too strict for skip-stop operation. In the cases that the safety operation headway is larger than the demand headway (due to skip operation), we cannot guarantee that the demand headway is satisfied at all times. Therefore we consider constraints (18a) and (18b) as a compromise to constraints (17a) and (17b).

$$d_{k,1} \leq \max\left\{kI_d, I_{\min} + \Delta_k + d_{k-1,1}\right\}, \quad \forall 1 \leq k \leq N_{\mathrm{tr}}, \tag{18a}$$

$$d_{k,N_{s+1}} \leq \max\left\{d_{0,N_s+1} + kI_d, I_{\min} + \Delta_k + d_{k-1,N_s+1}\right\}, \quad \forall 1 \leq k \leq N_{\mathrm{tr}}. \tag{18b}$$

*2.5.3. Turnaround Headway.* In our model, we specifically consider $c$ ($c \geq 1$) turnaround lines at the turnaround terminal. Under the infrastructure geometry, train $k-c$ should depart from station $N_s + 1$ before train $k$ departs from station $N_s$. That is, formulations (18a) and (18b) should be satisfied. Here $I_t$ refers to the minimum turnaround headway at the turnaround station. The formulation can better utilize the storage capacity at the turnaround station.

$$d_{k,N_s} + I_t - d_{k-c,N_s+1} \geq 0, \quad \forall c \leq k \leq N_{\mathrm{tr}}. \tag{19}$$

*2.5.4. Operational Constraint.* To avoid intolerable waiting time at the trip origins, we assume that each viable OD pair should be served in an acceptable time period which is marked as $I_p$; that is, constraint (20) needs to be satisfied. Here $x = \max\{k' \mid k' < k,\ a_{x,i}a_{x,j}np_x^{i,j} > 0 \wedge a_{k,i}a_{k,j}np_k^{i,j} > 0\}$, $\forall 1 \leq k \leq N_{\mathrm{tr}}$, $0 \leq i \leq N_s$, $i < j \leq N_s$.

$$d_{k,i} - d_{x,i} \leq I_p. \tag{20}$$

In practice, some stations such as terminal stations and transfer stations should not be skipped. For this, we further define a set SS that contains the stations that can be skipped. For the stations that do not belong to this set, they should not be skipped; see constraint (21). Here $\mathrm{SS} \cup \overline{\mathrm{SS}} = \mathrm{SN}$, where SN is the set that contains all metro stations.

$$a_{k,i} = 0, \quad \forall i \in \overline{\mathrm{SS}}. \tag{21}$$

## 3. Solution Approach

In the previous section, the FSSS model is formulated as a mixed integer nonlinear programming (MINLP) problem, which is comprised of the objective function (9), constraint (1), constraint (10)-constraint (11), constraints (16a) and (16b), and constraints (18a) and (18b)–constraint (21). Note that the model requires departure time optimization at each station, which imposes high computational complexity. To simplify

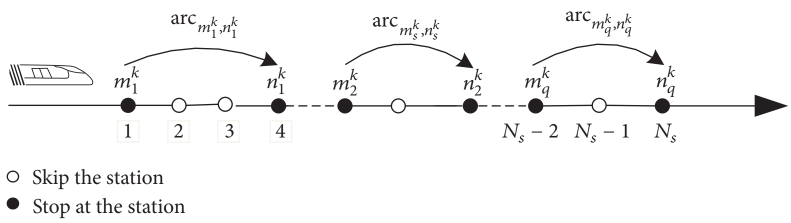



Figure 4: Definition of $S_k$ and "skip arcs."

the problem, we apply a rule-based simulation ((22)–(24)) that generates departure times to satisfy constraints (16a) and (16b) and constraints (18a) and (18b)-constraint (19). A similar rule-based computation of train departure times was implemented by Salzborn [23].

First, we define the terminal departure time as (22); the departure time at the terminal departure takes the maximum allowed value that satisfies the safety and passenger demand constraints; therefore it should result in the smallest required number of trains. Similarly, the departure time at the turnaround terminal should take (23). Equation (24) is then used to update the departure times at the intermediate stations.

$$d_{k,1} = d_{k-1,1} + \max\left\{kI_d - d_{k-1,1}, I_{\min} + \Delta_k\right\} + \max\left\{d_{k-c,N_s+1} - I_t - d_{k,N_s}, 0\right\}, \quad \forall c \le k \le N_{\mathrm{tr}}, \tag{22}$$

$$d_{k,N_s+1} = \max\left\{d_{k,N_s} + I_t, d_{k-1,N_s+1} + I_{\min} + \Delta'_k\right\}, \quad \forall 1 \le k \le N_{\mathrm{tr}}, \tag{23}$$

$$d_{k,i} = \begin{cases} d_{k,1} + \displaystyle\sum_{v=0}^{i}\left(t_{\mathrm{tr}}^{i} + a_{k,i}t_{\mathrm{dw}}^{i} - \left(v - \sum_{0}^{v} a_{k,i}\right)\delta\right), & i \le N_s, \ \forall 1 \le k \le N_{\mathrm{tr}} \\ d_{k,N_s+1} + \displaystyle\sum_{v=N_s+1}^{i}\left(t_{\mathrm{tr}}^{i} + a_{k,i}t_{\mathrm{dw}}^{i} - \left(v - \sum_{0}^{v} a_{k,i}\right)\delta\right), & i > N_s, \ \forall 1 \le k \le N_{\mathrm{tr}}. \end{cases} \tag{24}$$

It is particularly important to note that if the turnaround terminal is considered as an intermediate station (as in most existing literature), an extra amount of the maximum travel time difference, which takes the form of $\max\{\Delta_k, \Delta'_k\}$, needs to be added to the departure terminal. In the cases that $\Delta'_k$ is larger than $\Delta_k$, the departure time at the departure terminal can be expressed by (25). And this leads to delayed departure time compared to (22), which results in line capacity loss.

$$\begin{aligned} \tilde{d}_{k,1} &= d_{k-1,1} + \max\left\{kI_d - d_{k-1,1}, I_{\min} + \Delta'_k\right\} \\ &\quad + \max\left\{d_{k-c,N_s+1} - I_t - d_{k,N_s}, 0\right\}, \\ &\qquad \forall c \le k \le N_{\mathrm{tr}} \end{aligned} \tag{25}$$

Furthermore, it is widely known that NP-hard problems suffer from the "Curse of Dimensionality"; separating the forward and backward operation scheme can significantly reduce the dimension of the problem; that is, the computational complexity reduces from $O(2^{2N_sN_{\mathrm{tr}}})$ to $O(2^{N_sN_{\mathrm{tr}}+1})$.

Using the simulation scheme, the departure times can be determined, and the problem is thus reduced to the optimization of the binary variables $a_{k,i}$. The problem is NP-hard and an exhaustive search of the skip-stop schemes has an exponential time complexity $O(2^{N_sN_{\mathrm{tr}}})$. To avoid enumerating all possible outcomes, we develop a genetic algorithm (GA) approach to guide the searching process. GA is a metaheuristic method that imitates the process of natural evolution. GA is motivated by the principles of natural selection and survival-of-the-fittest individuals [24]. It is proved to be efficient in solving NP-hard optimization problems in the transportation field [25–28]. This method was also used by Lee et al. [13] to find a solution to the A/B mode-based skip-stop operation problem. In their work, gene is defined as the station type and it can take the choices of A, B, or AB. The time complexity of their problem is therefore $3^{N_s}$. Compared to Lee et al. [13], the proposed FSSS model offers flexible schemes and it can further consider the operation at the turnaround terminal. Realistic operational and infrastructure constraints are also considered. Therefore the proposed method is different from the one in Lee et al. [13]. The solution procedure of the developed GA algorithm is shown in Figure 5.

In our solution approach, a gene is defined as the operation choice at the specific station (0 or 1 which denotes skip or stop) and therefore a chromosome is defined as a binary string that embodies $\{a_{k,i} \mid i \in SS\}$. And this ensures that constraint (21) is satisfied throughout the searching process. In the initialization step, a number of $n_c$ chromosomes are randomly generated. Next, we check the feasibility of the generated chromosomes, using constraint (20). An arbitrarily large objective (fitness) value is assigned to the infeasible chromosomes to penalize the infeasible predecessors; for the feasible predecessors, their fitness values need to be evaluated using the objective function. We then apply the GA operators

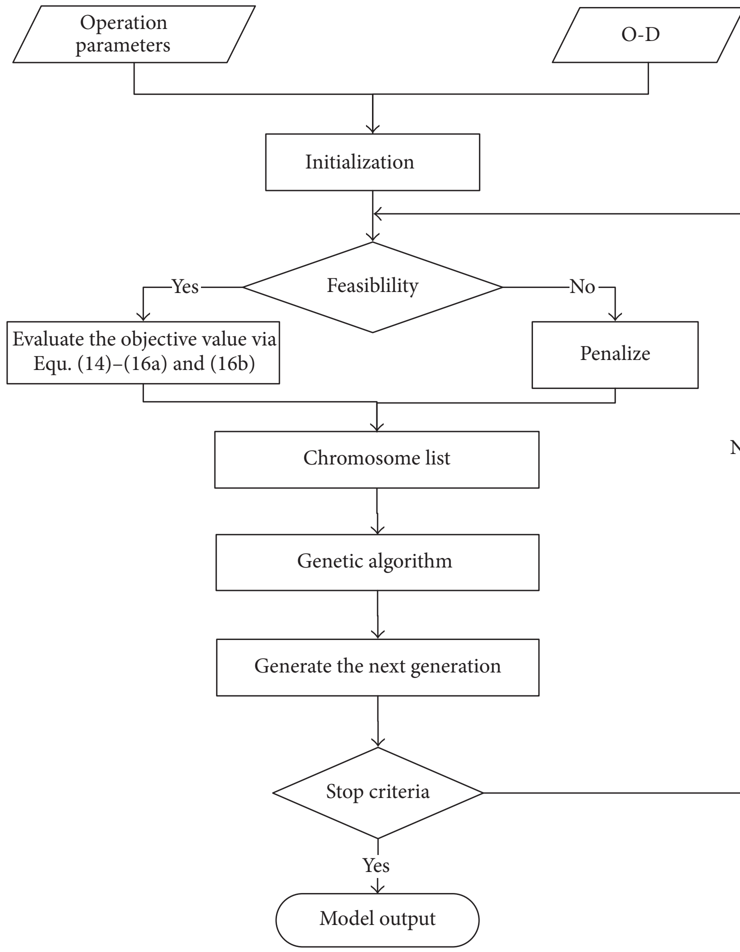



FIGURE 5: A GA-based solution approach.

to the list of chromosomes to generate a list of new solutions. The GA operators include (i) a selection operator, which is used to select $n_{\text{ps}}$ fitter solutions based on the Roulette Wheel Selection Algorithm [29]; (ii) a crossover operator, which is used to generate the child solutions from the selected solutions through a recombination process (with crossover rate $r_c$; see Back [30]); (iii) a mutation operator, which uses a small probability (i.e., $r_m$) to model the genetic mutation process. We further interpret the binary strings as Gray-coded integers, which is a proven practice that can be used to avoid the hamming cliff problem (see Back [30]). We then check the generated new solutions to see if they satisfy the stopping criteria; that is, the number of iterations exceeds a predefined value ($M$) or convergence is reached. If so, the final solution is recorded; otherwise, return to check the feasibility. Readers can refer to Jong [24] for more detailed information of GA.

## 4. Case Study and Numerical Results

*4.1. Data Processing and Experiment Setup.* Smart card data collected from urban transit systems have been widely used in transportation applications. Pelletier et al. [31] reviewed the application of smart card data in public transit field and categorized the data usage into three levels: the strategic level for long-term planning, the tactical level for service adjustment, and the operational level for performance measurement. Munizaga and Palma [32] proposed a method to estimate metro passenger OD using GPS data and smart card data in Santiago. Hong et al. [33] developed an approach that can match passenger trips to trains. Zhang et al. [22] formulated a model to optimize the urban rail operation based on the detailed passenger transaction data during weekdays. It is revealed by the literature that smart card data can better reflect the passenger OD pattern.

The proposed FSSS model was tested using the smart card data and operation parameters from a real world bidirectional metro line in Shenzhen, China. The metro system is shown in Figure 6. We selected Line 1 (with 28 intermediate stations and two terminals; highlighted in green) to test the proposed FSSS model.

The dynamic passenger OD rates were obtained from the smart card data using the following steps: (i) a transfer rule was defined based on the least number of stations; that is, for a trip that starts and ends at different lines, the trip will use the transfer station that yields the least number of stations

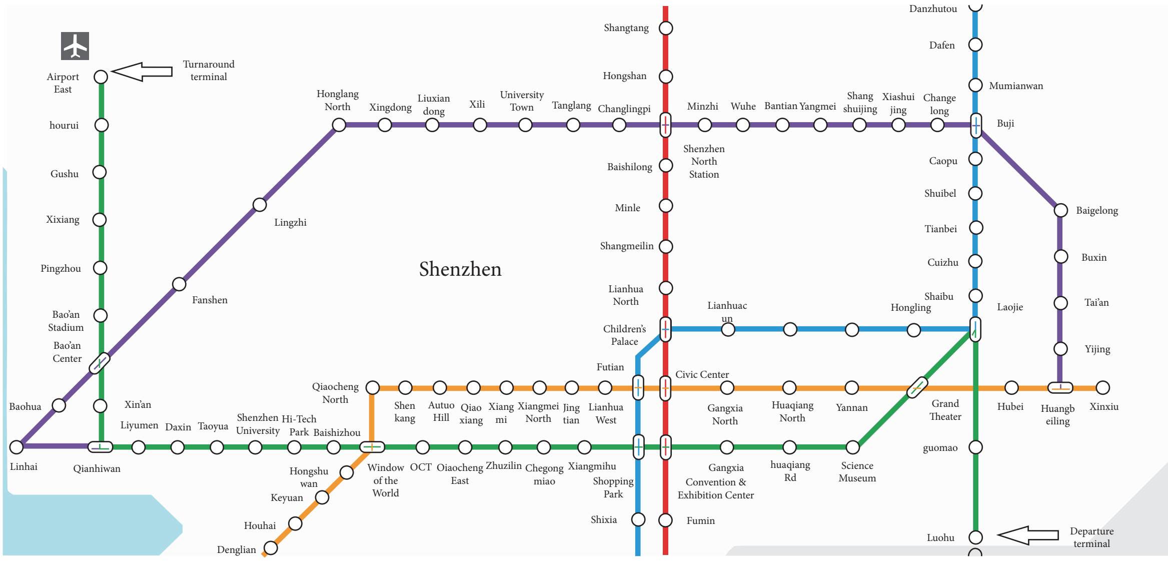



Figure 6: Shenzhen Metro map.

during the trip; (ii) the transfer rule was applied to the records that either start or terminate at some station of Line 1; (iii) the origins and destinations of all the Line 1 trips were then identified; (iv) by aggregating the OD data for the weekday off-peak period (13:30–17:00) during one month, the dynamic passenger OD rates ($r_g^{i,j}$) were calculated for each time period $g$, which was set to be 15 minutes in the experiment.

This metro line currently operates under an all-stop scheme. The round trip travel time is about 138.8 minutes. $N_{\mathrm{tr}}$ was set to be 10. The minimum safety headway, the demand headway, and the minimum turnaround headway were set to be 2 minutes, 6 minutes, and 3 minutes, respectively. According to the real world practice of Chicago Metro [11], the maximum acceptable waiting time $I_p$ was 12 minutes during the peak hours. As our study targets off-peak operation, $I_p$ was set to be 15 minutes. The time loss $\delta$ was calculated based on the reachable maximum speed $V_{\max}$ using formulation (26). Here $\alpha$ is the maximum acceleration rate and $\beta$ is the maximum deceleration rate. Parameters $V_{\max}$, $\alpha$, and $\beta$ were set to be 80 km/h, 0.8 m/s$^2$, and 1 m/s$^2$, respectively; the calculated $\delta$ was 25 seconds, using (26). Further discussions of this formulation can be found in Zheng et al. [16].

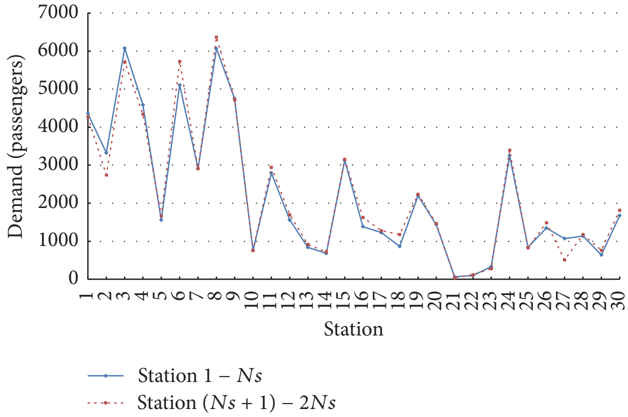
$$\delta = \frac{(\alpha + \beta) V_{\max}}{432\alpha\beta}. \tag{26}$$

The off-peak boarding and alighting counts at each station are shown in Figure 7(a). It is found that the number of passengers is relatively low at station 21, station 22, and station 23. An example of the time-dependent passenger demand from station 8 to station 15 is illustrated in Figure 7(b). The section travel time between two consecutive stations and the dwell time at each station are shown in Table 2. The dwell time at transfer stations $\{3, 4, 8, 9, 15, 24\}$ is 45 seconds, which is higher compared to the 35-second dwell time at other intermediate stations. The minimum turnaround time was set to be 120 seconds. The section travel times and dwell times were set according to the real world operating conditions. Parameter $c$ was set to be 2 which means that one train is allowed to be stored at the turnaround line.

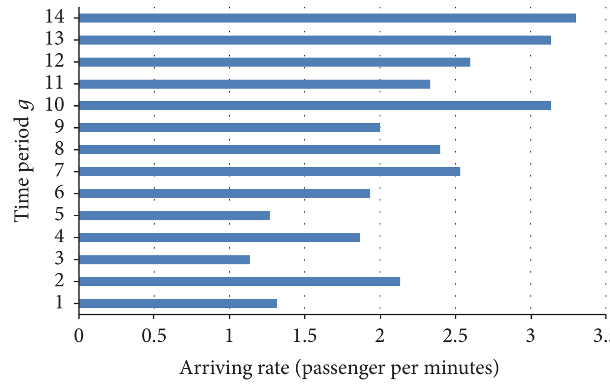
The all-stop operation scheme was simulated by setting parameter $a_{k,i}$ to 1 for all the trains and stations. We then use formulation (1)–formulation (11) and formulation (22)–formulation (24) to calculate the performance of the system under the all-stop scheme. It is found that the average waiting time $t_{\mathrm{wait}}^{\mathrm{pg}}$ and the average in-vehicle time $t_{\mathrm{tv}}^{\mathrm{pg}}$ are 3.0 minutes and 13.5 minutes, respectively. The average travel time under the all-stop scheme is 16.5 minutes. The total number of onboard passengers $n_{\mathrm{pg}}$ is 26146.

*4.2. Skip-Stop Optimization Results.* The proposed FSSS model is then solved using the GA-based solution approach. There are five major parameters used in GA, including the size of the chromosome list ($n_c$), population's size ($n_{\mathrm{ps}}$), maximum number of iterations ($M$), crossover rate ($r_c$), and mutation rate ($r_m$). The first three parameters were set to be 200, 500, and 2000, respectively. The crossover rate and the mutation rate are particularly important and are often found to affect the modeling performance (Back [30]). In the experiment, these parameters were set to be 0.8 and 0.005. The selection of parameters is based on a standard sensitivity analysis approach; details are not presented here.

For the FSSS, we manually set SS to include 20 low demand stations, that is, $\mathrm{SS} = \{10, 13, 14, 18, 21, 22, 23, 27, 28, 29, 32, 33, 34, 38, 39, 40, 43, 47, 48, 51\}$. The corresponding chromosome length is 200 (20 binary variables multiple 10 trains). The problem is solved on a computer with 2.0 GB RAM on a 2.0 GHz Intel Core i7 processor. The GA runtime is around 2 hours. However, as the FSSS model is proposed for offline operations, the computation time is not a big concern. By solving the MINLP, the optimized FSSS is listed in Table 3, only for the stations that belong to set SS. For the stations that belong to $\overline{\mathrm{SS}}$, they cannot be skipped. It is found that the stations that have the lower passenger demand (e.g., station




(a) Passenger demand at each station




(b) Time-dependent passenger OD rates from station 8 to station 15

Figure 7

Table 2: Train operation parameters.

| Station | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $t_{\mathrm{tr}}^{i}$ (seconds) | 0 | 95 | 68 | 85 | 103 | 77 | 145 | 95 | 62 | 132 | 97 | 93 | 122 | 84 | 85 |
| $t_{\mathrm{dw}}^{i}$ (seconds) | 35 | 35 | 45 | 45 | 35 | 35 | 35 | 45 | 45 | 35 | 35 | 35 | 35 | 35 | 45 |
| Station | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 |
| $t_{\mathrm{tr}}^{i}$ (seconds) | 69 | 101 | 90 | 171 | 83 | 90 | 78 | 100 | 81 | 72 | 104 | 84 | 204 | 210 | 156 |
| $t_{\mathrm{dw}}^{i}$ (seconds) | 35 | 35 | 35 | 35 | 35 | 35 | 35 | 35 | 45 | 35 | 35 | 35 | 35 | 35 | 35 |

10, station 14, station 21–station 23, station 38–station 40, and station 51) are frequently skipped under the optimized scheme. A total number of 35 stations are skipped, which can lead to energy and operational cost savings for operating agencies. The performance of FSSS is then compared to the all-stop operation, as shown in Table 3.

The results present in Table 4 show that the objective value of FSSS decreases by 5%, from 990 seconds to 940 seconds, which indicates a travel time saving of about 50 seconds per passenger. It is found that the average in-vehicle time is reduced by 6.5% with slightly increased (less than 4 seconds) passenger waiting time. As indicated by the value of $n_{\mathrm{bpg}}$, about 26.9% of total passengers (6604 out of 24576) benefited from FSSS. Compared to the all-stop operation, the number of onboard passengers under FSSS slightly decreases by 6%. This is mainly because the last train skips some stations and some passengers are not served by this train. These passengers are assumed to be picked up by the subsequent trains [18].

The timetables of the two operation schemes are illustrated in Figure 8. The light line denotes the all-stop operation, while the dark line represents the FSSS operation. It is found that FSSS has a slightly higher travel speed (the dark lines travel faster than the light lines). In specific, the average round trip travel time under FSSS is 135.7 minutes, which is more than 3 minutes shorter compared to the all-stop operation. The average departure interval is about 6 minutes which is similar to the all-stop operation.

Figure 9 indicates the seat occupancy and maximum section passenger load under the two schemes. Seat occupancy is defined as the passenger load over the seat numbers. It is found that, compared with all-stop scheme, passengers who board the skip-stop trains have a better chance to find empty seats, meaning FSSS is more favorable to the passengers as it enhances their level of service. By analyzing the maximum section passenger load, we find that the highest passenger load under FSSS is 558, which is far below the train capacity (1860). Therefore FSSS does not cause congestion on the trains.

*4.3. Analysis of the Passenger Confusion Rate.* It is widely recognized that the major problem in applying skip-stop operation is to keep the passengers informed regarding the skip-stop plans. The solution to it relies on information dissemination such as radio broadcast in stations/trains. Nonetheless, the authors believe the effect of certain percentage of passengers being confused by the skip operation and failing to find the right train is worth studying. The percentage is referred to as the passenger confusion rate hereafter in the paper. As noted in Assumption 5, we assume that these confused passengers shall notice their mistake (e.g., via train radio broadcast) once they board the train, and they will get off at the next stopping station. They will then wait at the station until being transported by the next available (correct) train to their destination. Variable $r_f$ is introduced here to denote the confusion rate, and a sensitivity analysis was conducted to see the effects of $r_f$ on the modeling performance. The variable $r_f$ was tested by taking values from 0 to 100%, with an increment of 20%.

The calculation of the boarded passengers (on train $k$ which stops at station $i$, with a destination $j$) can be divided

Table 3: Optimized chromosome.

| Train | S10 | S13 | S14 | S18 | S21 | S22 | S23 | S27 | S28 | S29 | S32 | S33 | S34 | S38 | S39 | S40 | S43 | S47 | S48 | S51 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 0 |
| 2 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 3 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 |
| 4 | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 1 |
| 5 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 6 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 0 |
| 7 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 8 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 1 |
| 9 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 10 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| *Skipped times* | *4* | *0* | *3* | *1* | *3* | *3* | *3* | *1* | *2* | *2* | *0* | *0* | *0* | *2* | *4* | *4* | *0* | *0* | *0* | *3* |

Table 4: Performance of all-stop operation and FSSS.

| Operation scheme | $t_{\mathrm{wait}}^{\mathrm{pg}}$ (min) | $t_{\mathrm{veh}}^{\mathrm{pg}}$ (min) | $t_{\mathrm{tv}}^{\mathrm{pg}}$ (min) | $n_{\mathrm{pg}}$ | $Z$ | $n_{\mathrm{bpg}}$ | $n_{\mathrm{ss}}$ |
|---|---|---|---|---|---|---|---|
| All-stop | 2.98 | 13.50 | 16.48 | 26,146 | 990 | 0 | 0 |
| FSSS | 3.04 | 12.62 | 15.66 | 24,576 | 940 | 6604 | 21 |
| | +0.06 (2%) | −0.88 (6.5%) | −0.82 (5%) | −1570 (6%) | −50 (5%) | — | — |

into two cases. Case 1 is that train $k$ will stop at station $j$; therefore passengers (from OD $i$ to $j$) board the right train. In Case 2, train $k$ will skip station $j$; people who board the train are the confused passengers. Further define set $C_i^k$ to represent the set of trains between trains $c_i^k$ and $k$ that stop at station $i$ and skip station $j$; that is, $C_i^k = \{x \mid a_{x,i} = 1 \wedge a_{x,j} = 0,\ c_i^k \le x < k\}$; the cardinality of the set is denoted as $n$; $c^h$ is used to denote the value of the $h$-th element in set $C_i^k$.

For Case 1, the number of onboard passengers from station $i$ to station $j$ via train $k$ can be calculated using (27). The corresponding passenger waiting time $\mathrm{tw}_k^{i,j}$ and passenger in-vehicle time $\mathrm{tv}_k^{i,j}$ can be calculated using (28) and (29), respectively.

$$\mathrm{np}_k^{i,j} = \sum_{h=1}^{n-1} \left(1 - r_f\right) * r_g^{i,j} \left(d_{c^h+1,i} - d_{c^h,i}\right) + r_g^{i,j} \left(d_{k,i} - d_{c^n,i}\right), \tag{27}$$

$$\begin{aligned}\mathrm{tw}_k^{i,j} &= \sum_{h=1}^{n-2} \left(1 - r_f\right) \left[\frac{1}{2} r_g^{i,j} \left(d_{c^h+1,i} - d_{c^h,i}\right)^2 \right. \\ &\quad \left. + r_g^{i,j} \left(d_{c^h+1,i} - d_{c^h,i}\right)\left(d_{k,i} - d_{c^h+1,i}\right)\right] + r_g^{i,j} \left(d_{k,i} - d_{c^n,i}\right)^2,\end{aligned} \tag{28}$$

$$\mathrm{tv}_k^{i,j} = \mathrm{np}_k^{i,j} * \left(d_{k,j} - d_{k,i} - t_{\mathrm{dw}}^{j}\right). \tag{29}$$

For Case 2, the number of confused passengers can be calculated using (30). They will get off at the nearest stop denoted as $i_k = \min\{i' \mid i' > i,\ a_{k,i'} = 1\}$ after departure from station $i$. If $i_k < j$ (destination has not been passed), these passengers shall wait until $k_e = \min\{k' \mid k' > k,\ a_{k',i_k} a_{k',j} = 1\}$ arrives. If $i_k > j$, these passengers need to take a train in the opposite direction, noted as $k_e = \min\{k' \mid d_{k',(2N_s+1-i_k)} > d_{k,i_k} \wedge a_{k',(2N_s+1-i_k)} a_{k',(2N_s+1-j)} = 1\}$, where $N_s$ is the total number of stations. The associated waiting time can be obtained using (31), and the in-vehicle time can be calculated using (32) for $i_k < j$ and (33) for $i_k > j$.

$$\mathrm{np}_k^{i,j} = r_g^{i,j} \left(d_{k,i} - d_{c^n,i}\right) * r_f, \tag{30}$$

$$\mathrm{tw}_k^{i,j} = \frac{1}{2} \left(d_{k,i} - d_{c^n,i}\right) \mathrm{np}_k^{i,j} + \left(d_{k_e,i_k} - d_{k,i_k} + \mathrm{tw}_{i_k}\right) \mathrm{np}_k^{i,j}, \tag{31}$$

$$\begin{aligned}\mathrm{tv}_k^{i,j} &= \left(d_{k,i_k} - d_{k,i} - \mathrm{tw}_{i_k}\right) \mathrm{np}_k^{i,j} \\ &\quad + \left(d_{k_e,j} - d_{k_e,i_k} - \mathrm{tw}_j\right) \mathrm{np}_k^{i,j},\end{aligned} \tag{32}$$

$$\begin{aligned}\mathrm{tv}_k^{i,j} &= \left(d_{k,i_k} - d_{k,i} - \mathrm{tw}_{i_k}\right) \mathrm{np}_k^{i,j} \\ &\quad + \left(d_{k_e,(2N_s+1-j)} - d_{k_e,(2N_s+1-i_k)} - \mathrm{tw}_{(2N_s+1-j)}\right) \mathrm{np}_k^{i,j}.\end{aligned} \tag{33}$$

By taking the summation of the above variables for all OD pairs of all trains, the total boarded passenger number, total waiting time, and in-vehicle time could be calculated, and the objective value could then be obtained using (9). The results are provided in Table 5.

It is observed that as the confusion rate increases (e.g., poorer information dissemination), the number of confused

Table 5: Sensitive analysis of the passenger confusion rate.

| Scheduling schemes | $r_f$ | Number of confused passengers | Ave. waiting time of confused/unconfused passengers | Ave. in-vehicle time of confused/unconfused passengers | Ave. waiting time of confused/unconfused passengers | $t_{\text{wait}}^{\text{pg}}$ (min) | $t_{\text{veh}}^{\text{pg}}$ (min) | $t_{\text{tv}}^{\text{pg}}$ (min) |
|---|---|---|---|---|---|---|---|---|
| All-stop | 0 | 0 | NA | NA | NA | 2.98 | 13.50 | 16.48 |
| FSSS | 0 | 0 | NA | NA | NA | 3.04 | 12.62 | 15.66 |
| | 0.2 | 79 | 5.75/4.42 | 8.74/8.19 | 14.49/12.61 | 3.05 | 12.63 | 15.68 |
| | 0.4 | 158 | | | | 3.06 | 12.67 | 15.73 |
| | 0.6 | 234 | | | | 3.08 | 12.70 | 15.78 |
| | 0.8 | 312 | | | | 3.09 | 12.74 | 15.83 |
| | 1.0 | 390 | | | | 3.10 | 12.77 | 15.87 |

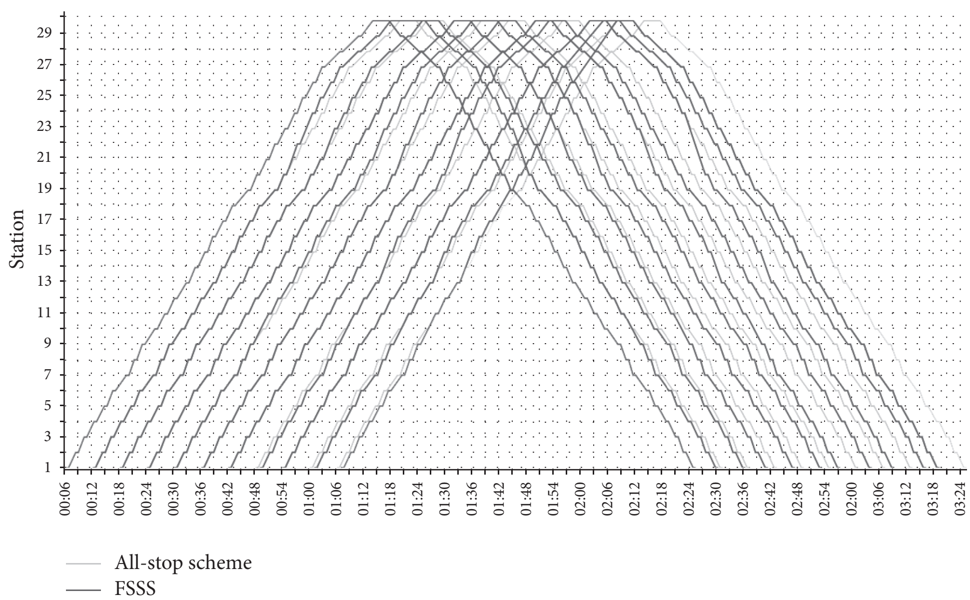



Figure 8: Timetables of all-stop scheme and FSSS.

passengers increases proportionally. However, this number only represents a small proportion of the some 25,000 passengers in total. This is because only the OD pairs with skipped destinations are affected, and the skipped stations have comparably low volume. For the confused passengers, it is found that their in-vehicle time and especially their waiting time increase. The increased in-vehicle time is because some passengers missed their destination and have to take the trains to the inverse direction (as in Case 2, $i_k > j$). And the increased waiting time is mainly because the correct train has skipped station $i_k$ and the confused passengers are not able to get on the correct train at that station. Per the specific experiment, we found that FSSS always outperforms the skip-stop scheme even when all the passengers of the skipped OD pairs get confused (i.e., $r_f = 1$). Therefore the system performance is stable under FSSS.

## 5. Conclusion

In this study, we developed a FSSS approach for bidirectional metro line operation during off-peak periods to better accommodate the spatially and temporally varied passenger demand. The problem was formulated as a mixed integer nonlinear programming (MINLP) model that minimizes the average passenger travel time. The proposed scheme is able to optimize the skip-stop scheme and it also models the train departure time at the departure and turnaround terminals. The departure intervals are constrained by three types of headways, that is, the minimum safety headway, headway that meets the passenger demand, and the minimum turnaround headway. A GA-based solution approach was then developed to efficiently solve the problem. The proposed scheme was tested using the smart card data collected at a real metro line in Shenzhen, China. Time-dependent passenger demands were considered in the experiment and the passenger confusion rate was also analyzed in the case study part.

The model performance was assessed with respect to the passenger benefits and the operational benefits. Overall, the FSSS is effective in timetable design. It was found that the FSSS is able to reduce the average passenger travel time by 0.82 minutes, which corresponds to 5% travel time saving. About 26.9% of total passengers benefit from the skip-stop scheme. It was also observed that the average round trip travel time under the skip-stop scheme is 135.7 minutes, which is more than 3 minutes shorter compared to the all-stop operation. The average train departure interval is 6 minutes which is the same compared to the all-stop scheme. From the timetable, it was found that 35 stations are skipped. Transit agencies may benefit from this scheme due to the savings of energy and operational costs. Through the sensitivity analysis of passenger confusion rate, we found that the proposed scheme always outperforms the all-stop scheme even when most passengers of the skipped OD pairs are confused and fail to get on the right train.

This paper mainly considers the skip-stop operation during off-peak scenarios in which the demands for some stations are quite low. As shown by the results, the skip operation does not lead to excessive waiting time at these stations. For peak-hour operations, since the passenger demands at most stations are relatively large, the skip operation may lead to longer waiting time at the skipped stations. In extreme cases, the number of waiting passengers may exceed the design capacity of the station. In future work, the tradeoff between in-vehicle travel time and passenger waiting time needs to be properly leveraged. The FSSS model currently requires a rule-based simulation to calculate the train departure time. Under the rule-based simulation, the terminal departure times need to be adjusted by the maximum travel time difference between two consecutive trains. This adjustment, however, may correspond to suboptimal solutions since departure choice can be made more wisely at each station including the intermediate stations. The GA-based solution approach requires a predetermined set (SS) that specifies the stations that can be skipped. This set needs to be determined using a more rigorous approach. The authors are also interested in extending the current methodology to optimize the skip-stop operation for a network-wide metro system. More results and findings will be provided in our subsequent studies.

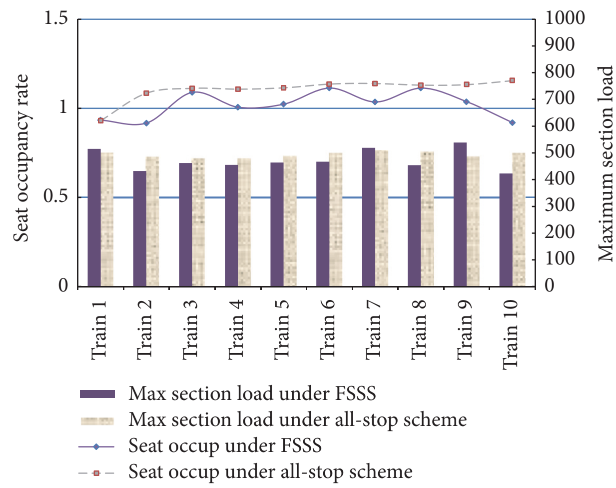



Figure 9: Seat occupancy and onboard passengers per train.

## Appendix

We consider $c$ turnaround lines that can be utilized to store trains at the turnaround terminal. Below we prove that, under certain circumstances, more turnaround lines (i.e., $c > 1$) can lead to reduced departure time at the turnaround station.

Let us first consider single turnaround line (i.e., $c = 1$). According to constraint (19), the following relationships hold.

When $c = 1 \wedge c \leq x \leq N_{\mathrm{tr}} - k$,

$$\begin{gathered} d_{k+1,N_s} - d_{k,N_s+1} \geq I_t \\ d_{k+2,N_s} - d_{k+1,N_s+1} \geq I_t \\ \vdots \\ d_{k+x,N_s} - d_{k+x-1,N_s+1} \geq I_t. \end{gathered} \tag{A.1}$$

For train $k + x$, formulation (A.2) also holds:

$$d_{k+x,N_s+1} - d_{k+x,N_s} \geq I_t. \tag{A.2}$$

By summing up the formulations in (A.1) and (A.2) above, it can be shown that

$$d_{k+x,N_s+1} - d_{k,N_s+1} \geq 2xI_t. \tag{A.3}$$

Based on constraint (12), (A.4) can be derived:

$$d_{k+x,N_s+1} - d_{k,N_s+1} \geq xI_{\min}. \tag{A.4}$$

Therefore, given the turnaround departure time of train $k$, the turnaround departure time of train $k + x$ should satisfy

$$\begin{gathered} d_{k+x,N_s+1} - d_{k,N_s+1} \geq \max\left\{2xI_t, xI_{\min}\right\} \\ \forall c = 1 \wedge c \leq x \leq N_{\mathrm{tr}} - k. \end{gathered} \tag{A.5}$$

For the cases with more than one turnaround line, similar derivations are made as follows.

When $c > 1 \wedge c \leq x \leq N_{\mathrm{tr}} - k$, we have

$$d_{k+x,N_s+1} - d_{k,N_s+1} \geq 2I_t. \tag{A.6}$$

Since (A.4) also holds for this case, the turnaround departure time of train $k + x$ should satisfy (A.6).

$$\begin{gathered} d_{k+x,N_s+1} - d_{k,N_s+1} \geq \max\left\{2I_t, xI_{\min}\right\} \\ \forall c > 1 \wedge c \leq x \leq N_{\mathrm{tr}} - k. \end{gathered} \tag{A.7}$$

Since $xI_{\min} \geq 2I_t \geq I_{\min}$ is usually true in real practices, the following relationships hold:

$$\begin{gathered} d_{k+x,N_s+1} - d_{k,N_s+1} \geq 2xI_t \\ \forall c = 1 \wedge c \leq x \leq N_{\mathrm{tr}} - k, \\ d_{k+x,N_s+1} - d_{k,N_s+1} \geq xI_{\min} \\ \forall c > 1 \wedge c \leq x \leq N_{\mathrm{tr}} - k. \end{gathered} \tag{A.8}$$

Since $2xI_t \geq xI_{\min}$, it can be concluded that the turnaround departure time can be reduced if they are more than one turnaround line and $2I_t \geq I_{\min}$.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This research was financially supported by the National Natural Science Foundation of China under Grant no. 71671147.

## References

[1] C. E. Cortés, V. Burgos, and R. Fernández, "Modelling passengers, buses and stops in traffic microsimulation: Review and extensions," *Journal of Advanced Transportation*, vol. 44, no. 2, pp. 72–88, 2010.

[2] Vuchic V. R., "Skip-stop operation: high speed with good area coverage," *Revue de I'UITP*, vol. 2, pp. 105–128, 1976, (French and German).

[3] V. R. Vuchic, *Urban transit: Operations, Planning and Economics*, John Wiley Sons, Hoboken, NJ, USA, 2005.

[4] X. J. Eberlein, N. H. M. Wilson, C. Barnhart, and D. Bernstein, "The real-time deadheading problem in transit operations control," *Transportation Research Part B: Methodological*, vol. 32, no. 2, pp. 77–100, 1998.

[5] A. Sun and M. Hickman, "The real-time stop-skipping problem," *Journal of Intelligent Transportation Systems: Technology, Planning, and Operations*, vol. 9, no. 2, pp. 91–109, 2005.

[6] B. Yu, Z. Z. Yang, and S. Li, "Real-time partway deadheading strategy based on transit service reliability assessment," *Transportation Research Part A: Policy and Practice*, vol. 46, no. 8, pp. 1265–1279, 2012.

[7] L. Fu, Q. Liu, and P. Calamai, "Real-time optimization model for dynamic scheduling of transit operations," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1857, pp. 48–55, 2003.

[8] National Research Council, *TCRP Report 100: Transit Capacity and Quality of Service Manual*, Transportation Research Board, Washington, DC, USA, 2nd edition, 2003.

[9] C. Leiva, J. C. Muñoz, R. Giesen, and H. Larrain, "Design of limited-stop services for an urban bus corridor with capacity constraints," *Transportation Research Part B: Methodological*, vol. 44, no. 10, pp. 1186–1201, 2010.

[10] M. Freyss, R. Giesen, and J. C. Muñoz, "Continuous approximation for skip-stop operation in rail transit," *Transportation Research Part C: Emerging Technologies*, vol. 36, pp. 419–433, 2013.

[11] Chicago-L.org. A/B Skip-Stop Express Service, https://www.chicago-l.org/operations/lines/route_ops/A-B.html.

[12] W. Suh, K.-S. Chon, and S.-M. Rhee, "Effect of skip-stop policy on a Korean subway system," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1793, pp. 33–39, 2002.

[13] Y.-J. Lee, S. Shariat, and K. Choi, "Optimizing skip-stop rail transit stopping strategy using a genetic algorithm," *Journal of Public Transportation*, vol. 17, no. 2, pp. 135–164, 2014.

[14] A. Jamili and M. P. Aghaee, "Robust stop-skipping patterns in urban railway operations under traffic alteration situation," *Transportation Research Part C: Emerging Technologies*, vol. 61, pp. 63–74, 2015.

[15] Z. Cao, Z. Yuan, and D. Li, "Estimation method for a skip-stop operation strategy for urban rail transit in China," *Journal of Modern Transportation*, vol. 22, no. 3, pp. 174–182, 2014.

[16] L. Zheng, R. Song, S. W. He, and H. D. Li, "Optimization model and algorithm of skip stop strategy for urban rail transit," *Journal of the China Railway Society*, vol. 6, pp. 135–164, 2009.

[17] S. L. Sogin, B. M. Caughron, and S. G. Chadwick, "Optimizing skip stop service in passenger rail transportation," in *Proceedings of the 2012 Joint Rail Conference, JRC 2012*, pp. 501–512, April 2012.

[18] Y. Wang, B. De Schutter, T. J. J. van den Boom, B. Ning, and T. Tang, "Efficient Bilevel approach for urban rail transit operation with stop-skipping," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 6, pp. 2658–2670, 2014.

[19] H. Niu, X. Zhou, and R. Gao, "Train scheduling for minimizing passenger waiting time with time-dependent demand and skip-stop patterns: nonlinear integer programming models with linear constraints," *Transportation Research Part B: Methodological*, vol. 76, pp. 117–135, 2015.

[20] B. Agard, C. Morency, and M. Trépanier, "Mining public transport user behaviour from smart card data," *IFAC Proceedings Volumes*, vol. 39, no. 3, pp. 399–404, 2006.

[21] M. Bagchi and P. R. White, "The potential of public transport smart card data," *Transport Policy*, vol. 12, no. 5, pp. 464–474, 2005.

[22] P. Zhang, X. Liu, and M. Chen, "Optimal train scheduling under a flexible skip-stop scheme for urban rail transit based on smartcard data," in *Proceedings of the 2016 IEEE International Conference on Intelligent Rail Transportation, ICIRT 2016*, pp. 13–22, August 2016.

[23] F. J. M. Salzborn, "Timetables for a suburban rail transit system," *Transportation Science*, vol. 3, no. 4, pp. 297–316, 1969.

[24] J. C. Jong, *Optimizing highway alignments with genetic algorithms [Doctoral dissertation]*, Department of Civil Engineering, University of Maryland, College Park, MD, USA, 1998.

[25] D. Arenas, R. Chevrier, S. Hanafi, and J. Rodriguez, "Solving the train timetabling problem: A mathematical model and a genetic algorithm solution approach," in *Proceedings of the In Proceedings of the 6th International Conference on Railway Operations Modelling and Analysis*, Tokyo, Japan, 2015.

[26] Y. Huang, X. Ma, S. Su, and T. Tang, "Optimization of train operation in multiple interstations with multi-population genetic algorithm," *Energies*, vol. 8, no. 12, pp. 14311–14329, 2015.

[27] Y. Yin, "Genetic-algorithms-based approach for bilevel programming models," *Journal of Transportation Engineering*, vol. 126, no. 2, pp. 115–119, 2000.

[28] Y. Yin, "Multi-objective bilevel optimization for transportation planning and management problems," *Journal of Advanced Transportation*, vol. 36, no. 1, pp. 93–105, 2002.

[29] D. E. Goldberg and K. Deb, "A comparative analysis of selection schemes used in genetic algorithms," in *Foundations of Genetic Algorithms*, pp. 69–93, Morgan Kaufmann, 1991.

[30] T. Back, "Optimal Mutation Rates in Genetic Search," in *Proceedings of the In Proceedings of the 5th International Conference on Genetic Algorithms*, pp. 2–8, 1993.

[31] M.-P. Pelletier, M. Trépanier, and C. Morency, "Smart card data use in public transit: a literature review," *Transportation Research Part C: Emerging Technologies*, vol. 19, no. 4, pp. 557–568, 2011.

[32] M. A. Munizaga and C. Palma, "Estimation of a disaggregate multimodal public transport Origin-Destination matrix from passive smartcard data from Santiago, Chile," *Transportation Research Part C: Emerging Technologies*, vol. 24, pp. 9–18, 2012.

[33] S.-P. Hong, Y.-H. Min, M.-J. Park, K. M. Kim, and S. M. Oh, "Precise estimation of connections of metro passengers from Smart Card data," *Transportation*, vol. 43, no. 5, pp. 749–769, 2016.