6-2023

Optimization of station-skip in a cyclic express subway service

Jingfeng YANG
Singapore Management University, jfyang.2018@phdcs.smu.edu.sg

Hai WANG
Singapore Management University, haiwang@smu.edu.sg

Jiangang JIN
Shanghai Jiaotong University

Citation

Optimization of Station-Skip in a Cyclic Express Subway Service

Jingfeng Yang[1] · **Hai Wang**[1] · **Jiangang Jin**[2]

Abstract

With rapid population growth and increasing demand for urban mobility, metropolitan areas such as Singapore, Tokyo, and Shanghai are increasingly dependent on public transport systems. Various strategies are proposed to improve the service quality and capacity of bus and subway systems. Express trains—i.e., trains that skip certain stations—are commonly used because they can travel at higher speeds, potentially reduce travel time, and serve more passengers. In this paper, we study cyclic express subway service (CESS), in which express trains provide routine transport service with cyclic (periodic) station-skip patterns that can be used in daily service. We propose an exact Mixed Integer Programming (MIP) model to optimize cyclic station-skip patterns for express trains operating in a single-track subway system. The objective is to reduce passengers' total travel time—i.e., the sum of waiting time and riding time—while considering demand intensity and distribution and train headway, frequency, and capacity. We implement the model in a set of numerical experiments using real data from Singapore. To solve the optimization problem more efficiently, we also propose a heuristic to solve large-scale problems. We observe that the exact MIP model for CESS provides optimal cyclic express service patterns within a reasonable computational time, and the heuristic method can significantly reduce the computational time and provide a good solution. The case study demonstrates that passengers' average travel time could be significantly reduced compared to local train service. We also discuss the potential transfer of passengers between express trains and evaluate its effects using numerical experiments.

✉ Jiangang Jin
jiangang.jin@sjtu.edu.cn

1 School of Computing and Information Systems, Singapore Management University,
Singapore, Singapore

2 School of Naval Architecture, Ocean & Civil Engineering, Shanghai Jiao Tong University,
Shanghai, China

1 Introduction

With rapid population growth and increasing demand for urban mobility, metropolitan areas such as Shanghai, Tokyo, and Singapore are increasingly dependent on public transport, such as bus and subway systems. For example, Singapore's population of 5.6 million is served by a metro system with five lines and 106 stations, and more than 2 million commuters use the subway every day. In Shanghai, the average daily metro system ridership is nearly 10 million. With such high demand, various strategies are proposed to improve the quality and capacity of subway systems. At the planning level, massive investment is required to improve and extend subway infrastructure—e.g., by building more lines and increasing the number of service trains. However, the construction of new infrastructure cannot keep up with the increase in passenger demand, especially in cities with fast-growing populations. Many strategies have been implemented at the operational level, including limiting the number of passengers who enter certain stations during peak hours (which, essentially, reduces demand by rejecting passengers) and skipping specific stations under high demand, even without advance notice to passengers.

Despite the merits of urban subway systems—large capacity, high reliability, and high efficiency—they also have disadvantages that cannot be ignored. One is long riding time due to frequent stops at many stations. This paper studies cyclic express subway service (CESS), which provides routine train service with cyclic and periodic station-skip patterns. CESS express trains follow well-optimized cyclic station-skip patterns and selectively skip certain stations. Express trains reduce travel time for passengers due to fewer speed decelerations and accelerations and fewer passengers alighting and boarding at skipped stations. We aim to reduce passengers' average travel time.

An example of CESS is depicted schematically in Fig. 1. The subway service has three station-skip patterns—the red node indicates the station is skipped, while the green node indicates the station should be stopped—and express trains follow the three patterns periodically. Each train serves a subset of all of the stations along the subway line. Using information about express train service displayed on the platform, passengers can decide which train to board according to their destination stations.

Once the CESS implemented in urban subway systems, the main challenge in practical daily operation is how to accurately convey information of trains arrival
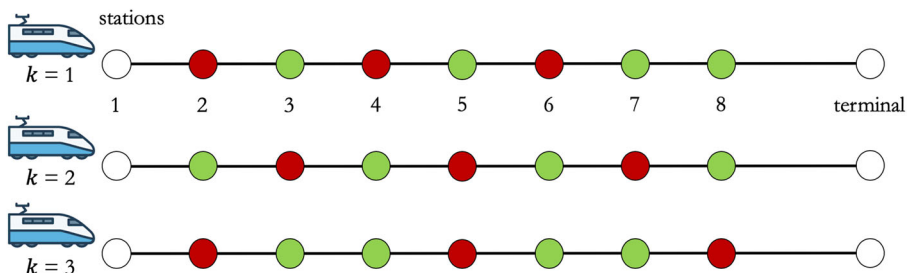

Fig. 1 Example of cyclic express subway service

and departure to travelers, especially for subway lines under high demand. While with the rapid development of IoT (The Internet of Things), information sharing platforms of urban subway system have been widely used in many cities, such as Singapore's "Mass Rapid Transit" (MRT) and Shanghai metro transit system. In Singapore, plasma displays with Station Travel Information System (STIS) programmed by Closed-loop Technology are located in stations at the concourse and on platforms, as shown in Fig. 2. At the concourse, displays show the destination and the estimated arrival times of the next two trains at each platform. At the platforms, the displays show the arrival times and destination of the next two trains. At terminal stations, the departure time is shown too. Some advanced sensors are also used in new subway lines (e.g., Downtown Line), and passengers can even see the load level of each car through the plasma displays. Another most used and effective real-time train information sharing tool is mobile application. As shown in Fig. 2c, "MyTransport" application is developed by Singapore government can provide bus and train travel guidance, which helps passengers easily obtain subway lines timetable and make travel plans accordingly by their mobile phones.

The main contribution of this paper is to propose an exasct Mixed Integer Programming (MIP) model to optimize cyclic station-skip patterns for express trains operating in a single-track subway system. The problem is challenging. First, CESS aims to provide routine express train service with cyclic station-skip patterns that
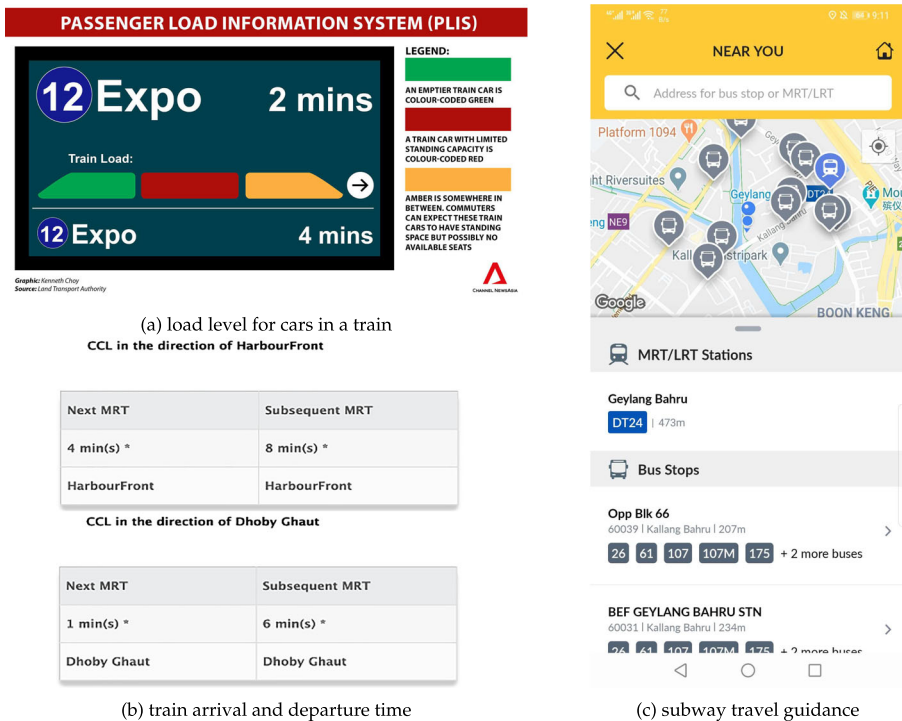

(a) load level for cars in a train

(b) train arrival and departure time

(c) subway travel guidance

Fig. 2 Communications between passengers and subway: information sharing and guidance

can be used daily—instead of temporary express train service under disruptions or extraordinary events—and therefore the combination of periodic station-skip patterns is critical. If, we assume that there have a line with 15 stations and the number of trains in one cycle equals 3, then the total number of possible station-skip patterns is more than 3×2^{15}. Second, many cities, such as Singapore and Shanghai, have single-track subway systems; as a result, we must consider headway between consecutive express trains to avoid train overtaking. In this paper, we consider implementing the routine express trains with cyclic station-skip patterns at certain times of the day, e.g., off-peak hours. Express trains with headway constraints make the problem complicated and hard to solve. We do not combine the regular trains and express trains in one model. Otherwise, train overtaking, which is not allowed in a single-track, is hard to avoid. Third, there are complicated trade-offs between passengers' in-vehicle riding time and out-of-vehicle waiting time under cyclic station-skip patterns. While passengers in the express subway system tend to have less in-vehicle riding time, they will suffer longer waiting time for the train that serves their destination station. During off-peak hours, we can skip low demand stations and save passengers travel time by providing express subway services. While for peak hours, express subway service may make passengers waiting at platform and cause congestion. The trade-off will be much more complicated if we consider the potential transfer of passengers between express trains.

2 Literature Review

Literature on the design and operation of urban transit services (e.g., buses and subways) can be classified into two groups. The first group focuses on the design of public transit network—for example, determining a set of lines and stations given demand distribution and city topology, such as work by Ceder and Wilson (1986), Yang and Bell (1998), Melkote and Daskin (2001), Guihaire and Hao (2008), and Yin et al. (2009) and Farahani et al. (2013). The second group focuses on the timetable of public transit service, e.g., optimizing timetables to reduce passenger waiting time in the context of either no congestion or oversaturation, such as Liebchen (2008) and Kaspi and Raviv (2013), and Sun et al. (2014).

Much research has been conducted on express service for bus systems. For example, Leiva et al. (2010) develop optimization models for the design of limited-stop services for an urban bus corridor to minimize social costs. Chiraphadhanakul and Barnhart (2013) propose an optimization model to determine a limited-stop express service route to operate in parallel with local bus service to maximize total user welfare. Chen et al. (2015) develop a mathematical model for the optimal stopping strategy on a bus line given stochastic travel time and vehicle capacity. Larrain et al. (2010) conduct experimental simulations and find that average trip length is a crucial parameter for the potential benefits of express service with high demand. As an extension, Larrain et al. (2015) propose heuristics for express service configuration of a bus corridor with and without congestion. Furth and Rahbee (2000) propose a discrete approach to study the impacts of changing bus-stop spacing, and identify the

optimal bus-stop location through dynamic programming. Stewart and El-Geneidy (2016) propose a new bus stop consolidation method to improve the quality of a transit agency. Cao and Ceder (2019) also use the skip-stop tactic on autonomous shuttle bus service in Auckland, New Zealand, which shows a reduction of 1.83% passengers' travel time and 8.11% number of vehicles compared to no skip-stop operation.

Literature on express service for subway systems is limited. The work of Suh et al. (2002) is among the earliest studies on express service for subway systems. They consider three typical train-stopping patterns. (1) All-stopping: Trains stop at all stations along the subway line. It is essentially a local train service. (2) Skip-stopping: Trains visit a subset of stations along the subway line, and some passengers must transfer between trains if no direct service is provided to their destination. (3) A combination of all-stopping and skip-stopping for multiple track systems: Trains on some tracks stop at a subset of stations, while trains on other tracks serve all stations.

Other work approaches the problem from diverse perspectives. Ulusoy et al. (2010) develop a model for cost-efficient operation that optimizes all-stopping short-term express rail services under heterogeneous demand. Gao et al. (2016) propose an optimization model to deal with the express subway rescheduling problem in the case of overcrowding after disruptions, and show that express service outperforms common service in that context. Gao et al. (2018) propose an optimization model for combined service with express trains and local trains on double tracks, in which express trains can overtake local trains. They present numerical results from Beijing's Metro Line to demonstrate the reduction in energy consumption and travel time. Jamili and Aghaee (2015) propose an operation mode for express subway service under uncertainty; a robust mathematical model and two heuristic algorithms are developed with the objective of increasing train speed and reducing operational costs. Freyss et al. (2013) study subway express service in a one-way single-track system given two types of stop patterns. Niu et al. (2015) aim to minimize passengers' waiting time given predetermined station-skip patterns with temporal demand. Parbo et al. (2018) formulate the skip-stop problem as a bi-level optimization model and implement it on the suburban railway network in the Greater Copenhagen Region. Results show the skip-stop strategy achieves 5.5% passengers' in-vehicle time and 3.2% travel cost reduction. Dong et al. (2020) study the integrated train stop planning and timetabling problem, which is modeled as a mixed-integer nonlinear programming problem and further solved using an extended adaptive large neighborhood search algorithm.

Some gaps are identified in the literature. Many previous studies focus on temporary express subway service under specific circumstances (e.g., disruption or overcrowding) instead of routine express subway service with cyclic station-skip patterns. In addition, in this paper we focus on a subway system with a single track, which imposes strong constraints on station-skip patterns due to the infeasibility of train overtaking. We also discuss the potential transfer of passengers—which is ignored in the literature—between express trains and evaluate its effects using numerical experiments.

3 The Cyclic Express Subway Service Problem

This section first describes the CESS problem for a single-track urban subway system, then proposes an exact MIP model to optimize cyclic station-skip patterns for routine express subway service.

3.1 Problem Description and Notation

We aim to optimize the cyclic station-skip patterns for express trains with the objective of minimizing passengers' average travel time. We consider a single-track subway system in which $S = \{1, 2, ..., |S|\}$ is the set of subway stations, and $K = \{1, 2, ..., |K|\}$ is the set of express trains with cyclic station-skip patterns. If we are given the number of station-skip patterns K in one service cycle and the number of skipped stations for each express train, the problem is to determine which stations $s \in S$ should be skipped by each train k. Figure 3 presents an example of a service cycle with 3 station-skip patterns ($|K| = 3$). Green nodes denote the stations that are served by a train, and the red nodes denote the stations that are skipped by a train. In this example, train 1 skips stations 2, 4, and 6; train 2 skips stations 3, 5, and 7; and train 3 skips stations 2, 5, and 8. With more stations skipped, the entire travel time from the first station to the last station on the subway line could be reduced, which could reduce passengers' travel time but increase passengers' waiting time.

The CESS problem is a combinatorial optimization problem. Before presenting the exact optimization model, we introduce the mathematical notation as follows:

3.2 An Exact Mixed Integer Programming Model

We now propose an exact MIP model that optimizes the cyclic station-skip patterns for express trains operating in a single-track subway system We introduce notation in Table 1.

3.2.1 Objective Function

In general, passengers using express service will have less in-vehicle riding time due to time reduction at the skipped stations. However, they will suffer longer out-of-vehicle waiting time for the train that serves both their origin and destination
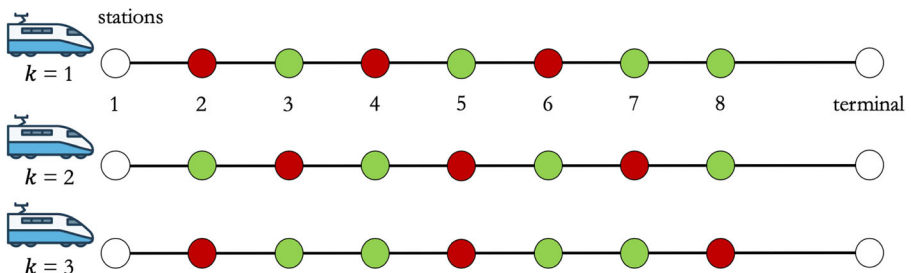

Fig. 3 Example of cyclic express subway service with three trains ($|K| = 3$)

Table 1 Notation for the exact MIP model

Sets	
S	$= \{1, 2, ..., |S|\}$, set of subway stations
K	$= \{1, 2, ..., |K|\}$, set of express trains in one service cycle
Parameters and Variables for Trains	
Parameters	
t_i	travel time from station i to station $i + 1$, $i \in S \setminus \{|S|\}$
p	service time for trains in each served station, i.e., passenger boarding and alighting time
h_{min}	minimum allowed safety headway between express service trains
h_{max}	maximum allowed headway between express service trains
C	train capacity
β	input parameter, represents the estimated average waiting time for passengers waiting for the next train which can take them to their destination
Decision Variables	
x_i^k	binary variable, 1 if station $i \in S$ is served by train $k \in K$, 0 otherwise
y_{ij}^k	binary variable, 1 if both stations i and j are served by train $k \in K$, 0 otherwise
Intermediate Variables	
h_L	headway between local trains service
A_i^k	arrival time of train k at station i in one service cycle
c_i^k	maximum number of passengers allowed to board train k at station i due to capacity
Parameters and Variables for Passengers	
Parameters	
λ_{ij}	number of passengers from station i to station j per unit of time
Intermediate Variables	
d_{ij}^k	number of passengers entering station i with destination j who want to board train k
r_{ij}^k	number of passengers at station i with destination j before the arrival of train k
r_i^k	total number of passengers wait at station i before the arrival of train k
R_{ij}^k	number of passengers at station i with destination j who have chance to board train k
R_{ij}^k	total number of passengers at station i who have chance to board train k
b_{ij}^k	number of boarding passengers at station i with destination j for train k
b_i^k	total number of boarding passengers at station i for train k
w_{ij}^k	number of passengers at station at station i with destination j who fail to board k
w_i^k	total number of passengers at station i who fail to board train k
n_i^k	total number of passengers in train k after passing station i

stations. To minimize passengers' total travel time, we choose an objective function as follows:

$$\text{maximize} \quad \Delta T = \Delta T_{riding} - \beta \cdot \Delta T_{waiting}$$

$$= \sum_{k=1}^{|K|} \sum_{i=2}^{|S|-1} n_{i-1}^{k} \cdot p \cdot (1 - x_i^k) - \beta \cdot \sum_{k=1}^{|K|} \sum_{i=2}^{|S|-1} w_i^k \quad (1)$$

Objective function (1) maximizes total travel time reduction for passengers using express service in one service cycle, which equivalently minimizes total travel time for passengers in one service cycle. In the first component, n_{i-1}^k denotes the number of onboard passengers in train k after its departure from station $i-1$, and x_i^k denotes whether train k serves station i. If $x_i^k = 1$, train k serves station i and there is no riding time reduction at station i for passengers on train k; if $x_i^k = 0$, train k skips station i and the corresponding riding time reduction is $n_{i-1}^k \cdot p$. In the second component, w_i^k denotes the total number of passengers at station i after the departure of train k, β denotes the estimated average waiting time at each station. So $\beta \cdot w_i^k$ is the increment of out-of-vehicle waiting time for passengers that cannot board train k at station i.

If we denote ΔT^* as the optimal value of equation (1) (i.e., the optimal travel time reduction for passengers in one service cycle), then the benefits of the express service can be evaluated using the average travel time reduction for each passenger:

$$\Delta T_{avg}^* = \frac{\Delta T^*}{|K| \cdot |T| \cdot \sum_{i=1}^{|S|} \sum_{j=1}^{|S|} \lambda_{ij}} \quad (2)$$

where $|K| \cdot |T| \cdot \sum_{i=1}^{|S|} \sum_{j=1}^{|S|} \lambda_{ij}$ is the total number of passengers requesting train service during one service cycle.

3.2.2 Constraints

There are four types of constraints in the exact MIP model, which are express service constraints, train headway constraints, train capacity constraints and passenger boarding and alighting constraints. We present all constraints as follows:

Express Service Constraints

$$\sum_{k=1}^{|K|} x_i^k \geq 1 \quad \forall i \in S \quad (3)$$

$$x_1^k = x_{|S|}^k = 1 \quad \forall k \in K \quad (4)$$

$$y_{ij}^k = x_i^k \cdot x_j^k \quad \forall i, j \in S, k \in K \quad (5)$$

$$\sum_{k=1}^{|K|} y_{ij}^k \geq 1 \quad \forall i, j \in S \quad (6)$$

$$x_i^k \in \{0, 1\} \quad (7)$$

$$y_{ij}^k \in \{0, 1\} \quad (8)$$

Constraint (3) ensures that for each station $i \in S$, at least one train stops and serves it in one service cycle. Constraint (4) ensures that all trains start from and serve the first station ($i = 1$) and terminate service at the last station ($i = |S|$). For each train k, constraint (5) ensures that it can bring passengers from origin i to destination j if and only if it serves both station i and station j. Constraint (6) ensures that at least one train will stop at both station i and station j at one cycle, which guarantee that the system can serve passengers with any origins and destinations pairs. Constraints (7) and (8) define the domain of decision variables.

Train Headway Constraints

$$A_1^1 = 0 \tag{9}$$

$$A_i^k = A_1^k + \sum_{i=1}^{i-1} t_{i-1} + p \cdot \sum_{i=1}^{i-1} x_i^k \quad \forall i \in S \setminus \{1\}, k \in K \tag{10}$$

$$h_{min} \leq A_i^{k+1} - A_i^k \leq h_{max} \quad \forall i \in S, k \in K \setminus \{|K|\} \tag{11}$$

Equation (9) indicates the arrival time of the first train ($k = 1$) at the first station is set to 0. Constraint (10) computes the arrival time of each express train at each station, and constraint (11) guarantees a minimum safety headway between successive express trains, and no train over-taking is allowed in the single-track subway system.

Train Capacity Constraints

$$c_1^k = C \quad \forall k \in K \tag{12}$$

$$c_i^k = C - n_{i-1}^k + \sum_{j<i}^{i-1} b_{ji}^k \quad \forall i \in S \setminus \{1, |S|\}, j \in S, j > i, \forall k \in K \tag{13}$$

$$n_1^k = b_1^k \quad \forall k \in K \tag{14}$$

$$n_i^k = n_{i-1}^k - \sum_{j=1}^{i-1} b_{ji}^k + \sum_{j=i+1}^{|S|} b_{ij}^k \quad \forall i \in S \setminus \{1, |S|\}, \forall k \in K \tag{15}$$

Constraints (12) and (13) represent the maximum number of passengers allowed to board at station i for train k due to train capacity. n_i^k in constraints (14) and (15) represents the total number of onboard passengers on train k after its departure from station i, which equals the number of passengers on train k at station $i - 1$ minus the total number of alighting passengers, then plus the total number of boarding passengers at station i

Passenger Boarding and Alighting Constraint

$$d_{ij}^k = \lambda_{ij} \cdot (A_i^k - A_i^{k-1}) \quad \forall i, j \in S, j > i, \forall k \in K \setminus \{1\} \tag{16}$$

$$r_{ij}^k = d_{ij}^k + w_{ij}^{k-1} \quad \forall i, j \in S, j > i, \forall k \in K \setminus \{1\} \tag{17}$$

$$r_i^k = \sum_{j=i+1}^{|S|} r_{ij}^k \quad \forall i \in S, \forall k \in K \tag{18}$$

Equation (16) obtains the number of passengers entering origin station i with destination station j who want to board train k. Term $(A_i^k - A_i^{k-1})$ denotes the departure time interval between successive express trains. Constraint (17) obtains the total number of passengers at station i with destination j before the arrival of train k, which equals the number of new passengers entering the station i during time interval $[A_i^{k-1}, A_i^k]$, plus the number of passengers failed to board train $k - 1$. Constraint (18) obtains the total number of passengers at station i before the arrival of train k.

$$R_{ij}^k = r_{ij}^k \cdot y_{ij}^k \quad \forall i, j \in S, j > i, \forall k \in K \tag{19}$$

$$R_i^k = \sum_{j=i+1}^{|S|} R_{ij}^k \quad \forall i \in S, \forall k \in K \tag{20}$$

In constraints (19) and (20), R_{ij}^k denotes the number of passengers at station i with destination j who have chance to board the express train k. If variable $y_{ij}^k = 1$, the value of R_{ij}^k equals r_{ij}^k. While if $y_{ij}^k = 0$, we have $y_{ij}^k = 0$, so passengers waiting at station i who want to board train k must wait for next train $k + 1$. Variable R_i^k is the total number of passengers in station i that has opportunity to board train k.

The total number of passengers who can board the train k at station i is represented by variable b_i^k. Since the capacity in a train is limited, over-saturated conditions are often observed during peak-hours, some passengers cannot board the train they want (e.g., express train k), and need to wait for next service (e.g., express train $k + 1$). Hence, variable b_i^k must satisfy following constraint:

$$b_i^k = \min \left\{ R_i^k, C - n_{i-1}^k + \sum_{j<i}^{i-1} b_{ji}^k \right\} \quad \text{for } i \in S \setminus \{1\}, \text{ and } k \in K \tag{21}$$

Term $\sum_{j<i}^{i-1} b_{ji}^k$ calculates the total number of passengers alighting at station i from train k. Variable b_{ij}^k is the number of passengers at station i with destination j who can board train k. When the capacity constraint is not violated, which means train k has enough capacity, we have $b_i^k = R_{ij}^k$. If available capacity of train k is less than R_i^k, then we have $b_i^k = C - n_{i-1}^k + \sum_{j<i}^{i-1} b_{ji}^k$. While to calculate variable b_i^k, we need to calculate variable b_{ij}^k first. In this paper, we follow the well-mixed assumption Gao et al. (2016), which assumes passengers with different destinations randomly arrive and are well mixed in each station $i \in S$. Hence, we have following equation:

$$\frac{b_{ij}^k}{b_i^k} = \frac{R_{ij}^k}{R_i^k} \implies b_{ij}^k = \frac{R_{ij}^k}{R_i^k} \cdot b_i^k \quad \text{for } i, j \in S, j > i, \text{ for } k \in K \tag{22}$$

Once the available capacity of train $k \in K$ is less than the number of passengers who want to board the train, some passengers are left behind at station i and need to wait for next train $k + 1$. Constraints (23) and (24) obtain the values of variable w_{ij}^k and w_i^k.

$$w_{ij}^k = R_{ij}^k - b_{ij}^k \quad \text{for } i, j \in S, j > i, \text{ for } k \in K \tag{23}$$

$$w_i^k = \sum_{j=i+1}^{|S|} w_{ij}^k \quad \text{for } i \in S, \text{ for } k \in K \tag{24}$$

3.2.3 Model Linearization

To reduce the computational requirements of the problem, we must linearize objective function (1) and constraints (5), (19), (21), and (22).

For the first component in objective function (1), we introduce a nonnegative variable $z_i^k = n_{i-1}^k \cdot (1 - x_i^k)$ to represent the in-vehicle riding time reduction for passengers on train k at station i; we can then linearize the component with the following auxiliary linear constraints (note that M is a large number):

$$z_i^k \leq n_{i-1}^k \quad \text{for } i \in S \setminus \{1, |S|\}, \text{ for } k \in K$$
$$z_i^k \leq M \cdot (1 - x_i^k) \quad \text{for } i \in S \setminus \{1, |S|\}, \text{ for } k \in K$$
$$z_i^k \geq n_{i-1}^k - M \cdot x_i^k \quad \text{for } i \in S \setminus \{1, |S|\}, \text{ for } k \in K$$

If decision variable $x_i^k = 0$, we have $z_i^k = n_{i-1}^k$; while if $x_i^k = 1$, the value of z_i^k equals 0.

For constraint (5), it can be linearized using the following auxiliary linear constraints:

$$y_{ij}^k \leq x_i^k \quad \text{for } i, j \in S, j > i, \text{ and } k \in K$$
$$y_{ij}^k \leq x_j^k \quad \text{for } i, j \in S, j > i, \text{ and } k \in K$$
$$y_{ij}^k \geq x_i^k + x_j^k - 1 \quad \text{for } i, j \in S, j > i \text{ and } k \in K$$

Constraint (19) can be linearized using the following auxiliary linear constraints:

$$R_{ij}^k \leq r_{ij}^k \quad \text{for } i, j \in S, j > i, \text{ and } k \in K$$
$$R_{ij}^k \leq M \cdot y_{ij}^k \quad \text{for } i, j \in S, j > i, \text{ and } k \in K$$
$$R_{ij}^k \geq r_{ij}^k - M \cdot (1 - y_{ij}^k) \quad \text{for } i, j \in S, j > i, \text{ and } k \in K$$

Constraint (21) can also be linearized using the following auxiliary linear constraints:

$$b_i^k \leq R_i^k \quad \text{for } i, j \in S \setminus \{1\}, j > i, \text{ and } k \in K$$

$$b_i^k \leq C - n_{i-1}^k + \sum_{\substack{j < i}}^{i-1} b_{ji}^k \quad \text{for } i, j \in S \setminus \{1\}, j > i, \text{ and } k \in K$$

$$b_i^k \geq R_i^k - M \cdot (1 - \theta_i^k) \quad \text{for } i, j \in S \setminus \{1\}, j > i, \text{ and } k \in K$$

$$b_i^k \geq C - n_{i-1}^k + \sum_{\substack{j < i}}^{i-1} b_{ji}^k - M \cdot \theta_i^k \quad \text{for } i, j \in S \setminus \{1\}, j > i, \text{ and } k \in K$$

Here we introduce new a binary variable θ_i^k to denote the relation between variable R_{ij}^k and $C - n_{i-1}^k + \sum_{j<i}^{i-1} b_{ji}^k$. The value of θ_i^k is determined by following rule:

$$\theta_i^k = \begin{cases} 1 & \text{if } R_i^k \leq C - n_{i-1}^k + \sum_{j<i}^{i-1} b_{ji}^k \\ 0 & \text{if } R_i^k > C - n_{i-1}^k + \sum_{j<i}^{i-1} b_{ji}^k \end{cases} \tag{25}$$

If $\theta_i^k = 1$, we have $b_i^k = R_i^k$; while if $\theta_i^k = 0$, the number of passengers who can board the express train (b_i^k) equals $C - n_{i-1}^k + \sum_{j<i}^{i-1} b_{ji}^k$. We can further linearize constraint (25) as following:

$$\theta_i^k > \frac{(C - n_{i-1}^k + \sum_{j<i}^{i-1} b_{ji}^k) - R_i^k}{M}$$

$$\theta_i^k \leq 1 + \frac{(C - n_{i-1}^k + \sum_{j<i}^{i-1} b_{ji}^k) - R_i^k}{M}$$

$$\theta_i^k \in \{0, 1\}$$

Finally, we linearize constraint (22) by using linear functions to approximate the quadratic term:

$$b_{ij}^k \leq R_{ij}^k \quad \text{for } i, j \in S, j > i, \text{ and } k \in K$$

$$b_{ij}^k = \min \left\{ \frac{\lambda_{ij}}{\sum_{j=i+1} \lambda_{ij}} \right\} \quad \text{for } i, j \in S, j > i, \text{ and } k \in K$$

Parameter η_{ij}^k is the ratio of passenger demand for each origin-destination pair (i, j) and the sum of passengers at station i. This is a reasonable assumption which is also used in Gao et al. (2016) to deal with the skip-stop train scheduling under over-crowded situation after disruption.

4 Discussion of CESS with Potential Passenger Transfer

Note that the exact MIP model proposed in Section 3 aims to optimize cyclic station-skip patterns while assuming that passengers will not transfer between express trains. In practice, however, some passengers may transfer between express trains. Instead of

boarding a train that serves their destination station directly, they could board a train that stops at any intermediate station (i.e., transfer station) along the subway line and transfer to another train that serves destination station. Transfer behavior and passengers' corresponding travel time are notoriously difficult to analyze. To the best of our knowledge, no previous work addresses the behavior and effects of passenger transfer. In this section, we discuss potential passenger transfers and their effects on passenger travel time. Specifically, we provide a simple approximated component to the objective function to capture the travel time reduction due to passenger transfer, and evaluate CESS performance with passenger transfer using numerical experiments.

Passenger transfer imposes great complexity on the problem. An example is illustrated in Fig. 4. For passengers entering origin station $i = 3$ with destination station $j = 6$ before arrival of express train $k = 1$, there are two route choices: either (1) waiting at platform until the arrival of train $k = 3$, which will take them to destination station $j = 6$ directly, or (2) taking train $k = 1$ to station 4 firstly, and then transferring to train $k = 2$, which will finally take them to destination station $j = 6$.

4.1 Simple Approximated Revision to the Objective Function

For a passenger with specific origin and destination stations, express service with cyclic station-skip patterns can be represented as a network graph. In the example illustrated in Fig. 4, three station-skip patterns ($|K| = 3$) operate in every service cycle. Nodes (red represent skipped stations) denote all stations between origin i and destination j for the three station-skip patterns. Solid lines denote passengers riding, and dashed lines denote passengers transferring or waiting.

To obtain the best transfer route for a passenger with any specific origin and destination stations, we must solve a shortest path problem on the network graph in Fig. 5. However, the network graph is unknown before we determine the cyclic station-skip patterns, and it is difficult or even impossible to obtain an exact closed-form formula for passengers' travel time on the best route under unknown station-skip patterns. In this subsection, we provide a simple approximated component to revise the objective function to capture passenger travel time due to passenger transfer. We intend to evaluate the additional benefits due to passenger transfer using numerical experiments, rather than providing an exact formula theoretically.
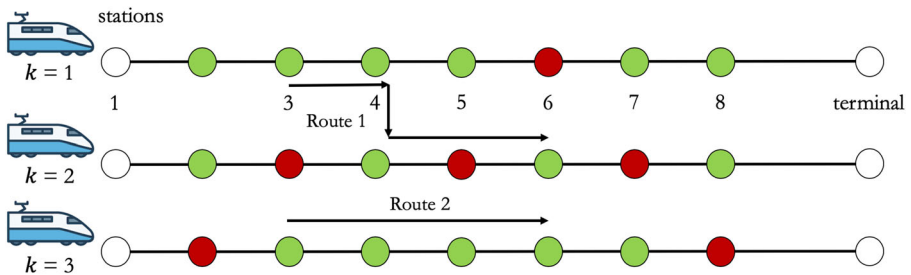

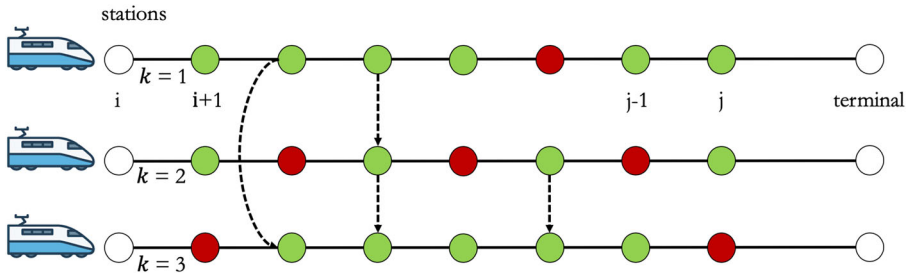**Fig. 4** Example of passenger transfer between express trains

Fig. 5 Network graph representation for passenger riding and transfer

As described in Section 3, the objective function contains two components: in-vehicle riding time ΔT_{riding} and out-of-vehicle waiting time $\Delta T_{waiting}$. When we take passenger transfer into consideration, in-vehicle riding time:

$$\Delta T_{riding} = \sum_{k=1}^{|K|} \sum_{i=2}^{|S|-1} n_{i-1}^k \cdot p \cdot (1 - x_i^k)$$

is still valid, so we only need to revise the waiting time function $\Delta T_{waiting}$. If passengers can transfer between express trains, waiting time will be further reduced compared to the case in which passengers are limited to using direct service. Specifically, with passenger transfer, w_i^k, the number of waiting passengers at each station i after the departure of each train k will be reduced, which leads to less out-of-vehicle waiting time.

Intuitively, with a larger K, passengers will have more feasible transfer routes, so out-of-vehicle waiting time will decrease; with many stations be skipped, passengers will have fewer feasible transfer routes, so out-of-vehicle waiting time will increase. Hence, we select four most representative features in CESS, which are list as follows:

- average train headway in CESS: h_{avg};
- average number of stations skipped for each express train: m;
- total number of trains operations: $|K|$;
- total number of waiting passengers: $\sum_{i \in S} w_i^k$

What needs to be emphasized here is that the variable m is a given parameter, not a decision variable, it equals $\frac{\sum_{i,k} x_i^k}{|K|}$, where the value of x_i^k is obtained by solving the CESS without passenger transfer. Besides, we do not set any constraints on the number of stations we need to skipped in CESS without passenger transfer. So variable m is only used in passenger transfer case. We test different formulas through numerical experiments, and obtain a simple approximated term:

$$\Delta_{transfer} = h_{avg} \cdot |K| \cdot \left(\sum_{i \in S} w_i^k + \frac{1}{m} \right) \tag{26}$$

Equation (26) will be added as a third component in the objective function to capture waiting time reduction due to passenger transfer. We evaluate the approximated formula using numerical experiments and find that it works well.

4.2 Heuristic

Although the CESS is a planning problem for which we do not need to solve the MIP model in a short time, for subway lines with a lot of stations ($|S|$ is large) or number of trains (value of $|K|$) in one cycle is large, it is hard for the commercial solvers to find a optimal solution within hours. Thus we propose a heuristic to reduce the computational time. The main idea of this heuristic arises from the observation that the optimal express subway service pattern with number of $|K|$ trains in one cycle is almost identical to the optimal express subway service pattern with number of $|K|-1$ trains in one cycle. In other words, we assume that increasing the value of $|K|$ in the MIP model will affect the station-skip plan of the last train $|K|$, while the optimal station-skip plan for trains 1 to $|K|-1$ will not change too much. Then an optimal solution with $|K|-1$ trains can be a valuable input to the MIP model with $|K|$ trains. Here we describe the main steps of this heuristic to solve the large CESS problem as follows:

- **Input**: number of trains $|K| = n$, where n is large and the CESS is hard to be solved to optimal; k_{max} is the maximum number of trains in one cycle which the CESS can be solved exactly in 60 minutes; we have $k_{max} \leq |K|$;
- **Step 1.** let variable k' denotes the number of trains in one cycle in CESS, initialize the value of $k' = k_{max}$;
- **Step 2.** solve the exact CESS model with the number of trains in one cycle with k', get the optimal solution x_i^k, for $k \leq k'$;
- **Step 3.** check the solution obtained in step 2, if $x_i^k = 0$, add the constraint $x_i^k = 0$ into the exact CESS model with the number of trains in one cycle $k' = k' + 1$;
- **Step 4.** repeat steps 2 and 3 until we have $k' = |K| = n$.

In this paper, we first try to solve the exact MIP model with small number of stations $|S|$ and trains $|K|$ in one cycle. With the value of $|K|$ increases and it cannot be solved by solvers within hours, we use the heuristic method to solve it.

5 Case Study

In this section, we implement the proposed exact MIP model for both CESS with and without passenger transfer in a set of numerical experiments using real data on passenger transactions from Singapore's MRT (Mass Rapid Transit) system; data are recorded by a smart card-based automated fare collection system. We choose two subway lines, as is shown in Fig. 6: (1) a section of Downtown Line consists of 10 stations (from MacPherson MRT station to Expo MRT station), and (2) the whole North East Line consists of 15 stations (from HarbountFront MRT station to Punggol MRT station).

Fig. 6 Singapore metro map

5.1 Inputs and Parameters

Table 2 summarizes input and parameters. We implement the MIP model to optimize cyclic express train service during the evening off-peak hour (2:00 p.m.–3:00 p.m.). The arrival rate of passengers can be obtained from the automatic fare collection system, which records the passenger ID, boarding station, alighting station, ride start time, and ride end time. Sun et al. (2014) analyze smart card data from Singapore's MRT system for metro service timetable design, and distinguish two types of metro trips: full trips and partial trips. In this paper, we use the same method to reconstruct

Table 2 Input and parameters in case study

| Inputs & Parameters	North East Line	Downtown Line
p: service time at each station	1 min	1 min
h_L: headway for local service train	3 min	3 min
h_{min}: minimum safety headway	30 sec	30 sec
h_{max}: maximum allowed headway	4 min	4 min
C: capacity	1920 per train	931 per train
β: average waiting time in CESS	30, 60 sec	30, 60 sec
$|K|$: number of patterns in a CESS	2 to 4 trains	2 to 4 trains
D: demand distribution with different values of λ_{ij} (average travel distance)	1 (6.54km), 2 (7.00km), 3 (8.35km)	1 (9.65km)
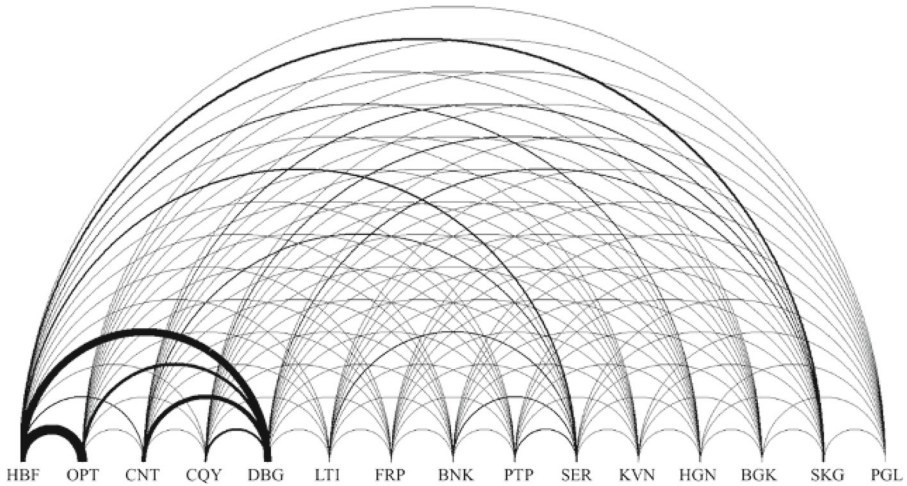
Fig. 7 Origin and destination distribution along the subway line

travel trips and estimate passenger demand λ_{ij}. Here, in the last row of Table 2, we use set $D = \{1, 2, 3\}$ to denote tested three demand distributions with different values of λ_{ij} for the North East Line with the average travel distance for each demand distribution in parentheses. Same for the Downtown Line. Figure 7 shows the origin and destination distribution along the North East Line, and Fig. 8 shows the boarding and alighting passengers demand at each station. The width of each arc indicates the intensity of demand on each origin and destination pair. We can see from the figure that demand distribution is unbalanced.

The exact MIP model for the CESS problem is mathematically similar to the facility location problem, with many additional constraints. Model linearization has brought a larger number of auxiliary decision variables and constraints. The total
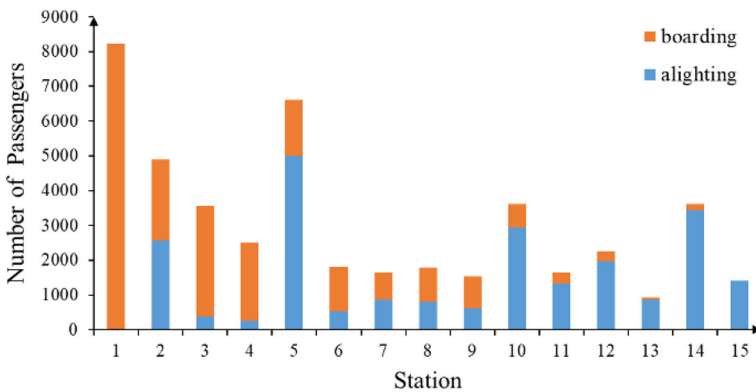

Fig. 8 Number of passengers boarding and alighting at each station

number of possible station-skip patterns increases exponentially with the total number of stations in one cyclic express service. So the CESS problems with or without passenger transfers are both NP-hard problems. In this paper, we first employ standard solvers (e.g., Cplex and Gurobi) to solve the CESS problem with two express service trains ($|K| = 2$) on the section of Downtown Line with total number of 10 MRT stations ($|S| = 10$). For larger instances with more trains ($|K| \geq 3$) and more stations ($|S| = 15$), it is hard for the commercial sovlers to find an optimal solution within 1 hour, thus a heuristic is further proposed to reduce the computational time. All experiments were conducted on a PC with an Intel Core i7 3.40 GHz with 16 GB RAM. The optimization model is coded in Python and Gurobi solver 8.1, with the standard configuration.

5.2 Computational Results

In this section, we implement the exact MIP model with heuristic (if needed) and evaluate the performance of the cyclic express subway service with and without passenger transfer. Note that parameter $|K|$ denotes the number of station-skip patterns in one service cycle. Because we aim to provide the CESS be operated with regular trains as daily service in the real world urban subway system, the total number of trains in one service cycle should not be too large. As mentioned in Section 1, with the help of plasma displays in MRT station, passengers can know the running schedule of next two trains. So with more trains in one cycle may also cause complexity in practical train scheduling and communication. So in this paper, we evaluate the cases with $|K| \in [2, 4]$, most of which can be solved exactly within 60 minutes. Since the CESS problem is a design problem at the planning level, this computational time is acceptable. Table 3 lists all instances we test and the methods we used to solve them.

By solving instances No.1 to No.3, corresponding optimal cyclic station-skip patterns for certain values of $|K|$ are shown in Fig. 9. As can be seen in the figure, the red nodes represent stations will be skipped while the green nodes represent stations must be stopped. Each train k will skip different sets of stations. Stations with low boarding and alighting demand are more likely to be skipped. For example, both stations 7 and 9 have low demand, and are skipped by some train(s) (e.g., $k = 1$ in Fig. 9a, $k = 1$ in Fig. 9b and $k = 1, 4$ in Fig. 9c) in all of the example cases; both stations 2 and 3 have high demand, and are rarely skipped by any train in all of the example cases.

To further explore the benefits of CESS, we calculate the average travel time reduction for passengers with different values of passenger average waiting time β on Downtown Line with instances No.11 to No.15. The results are shown in Table 4. we find that the percentage of average passenger travel time reduction is 10.09%, 12.70% and 12.72% for $|K| = 2, 3$ and 4, respectively. With the given number of trains in one cycle increases, the average passenger travel time will also increase, but with a slower growth rate (2.70% from $|K| = 2$ to $|K| = 3$, while 0.02% from $|K| = 3$ to $|K| = 4$), which means that a more flexible CESS (i.e., a larger $|K|$) will bring limited benefits in general. Of course, with higher passenger average waiting time ($\beta = 60$ sec), the average passenger travel time reduction will then decrease or

Table 3 Case study instances

Instance	Line	Trains	Avg Waiting Time	Avg Travel Distance	Transfer	Method
No.1	NE	$|K| = 2$	$\beta = 30$ sec	6.54 km	×	exact
No.2	NE	$|K| = 3$	$\beta = 30$ sec	6.54 km	×	exact
No.3	NE	$|K| = 4$	$\beta = 30$ sec	6.54 km	×	heuristic
No.4	NE	$|K| = 2$	$\beta = 30$ sec	7.00 km	×	exact
No.5	NE	$|K| = 3$	$\beta = 30$ sec	7.00 km	×	exact
No.6	NE	$|K| = 4$	$\beta = 30$ sec	7.00 km	×	heuristic
No.7	NE	$|K| = 2$	$\beta = 30$ sec	8.35 km	×	exact
No.8	NE	$|K| = 3$	$\beta = 30$ sec	8.35 km	×	exact
No.9	NE	$|K| = 4$	$\beta = 30$ sec	8.35 km	×	heuristic
No.10	DT	$|K| = 2$	$\beta = 30$ sec	9.65km	×	exact
No.11	DT	$|K| = 3$	$\beta = 30$ sec	9.65km	×	exact
No.12	DT	$|K| = 4$	$\beta = 30$ sec	9.65km	×	exact
No.13	DT	$|K| = 2$	$\beta = 60$ sec	9.65km	×	exact
No.14	DT	$|K| = 3$	$\beta = 60$ sec	9.65km	×	exact
No.15	DT	$|K| = 4$	$\beta = 60$ sec	9.65km	×	exact
No.16	NE	$|K| = 2$	$\beta = 30$ sec	9.65km	✓	exact
No.17	NE	$|K| = 3$	$\beta = 30$ sec	9.65km	✓	exact
No.18	NE	$|K| = 4$	$\beta = 30$ sec	9.65km	✓	heuristic

even drop below 0, in which case the increment of waiting time caused by the express train cannot be compensated by the riding time reduction.

5.3 Sensitivity Analysis

The numerical results presented in the previous subsection are based on real passenger demand during the evening off-peak hour (2:00 p.m.–3:00 p.m.). In this subsection, we examine CESS performance when passenger demand changes—e.g., when passenger origin and destination demand distributions changes. To evaluate CESS performance with different passenger origin and destination demand distributions, we implement the model in the cases of different demand distributions, especially different average passenger travel distance in the subway system. In the original demand, the average travel distance of passengers in the subway system is 6.54 km. In this subsection, we construct two distribution scenarios with longer average passenger travel distances. Figure 10 shows a hypothetical demand distribution with average passenger travel distance equals to 7.00 km and 8.35 km. Table 5 shows the corresponding performance of the CESS.

Typically, with a longer average passenger travel distance in the subway system, the in-vehicle riding time reduction will become larger, and hence the average passenger travel time reduction will increase. In the case of 6.54 km average passenger
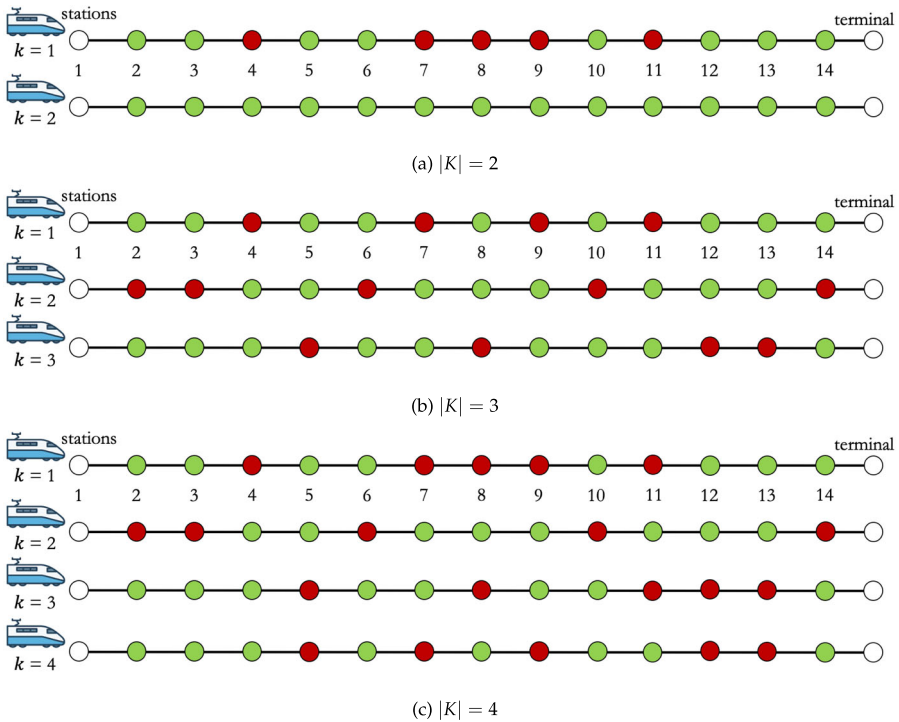
(a) $|K| = 2$

(b) $|K| = 3$

(c) $|K| = 4$

**Fig. 9** Optimal cyclic station-skip patterns for instances No.1, No.2 and No.3

travel distance, the percentage of average passenger travel time reduction is 5.95%, 8.28% and 9.15% for $|K| = 2, 3$ and 4; in the case of 7.00 km average passenger travel distance, the percentage of average passenger travel time reduction is 7.51%, 8.58% and 9.31%; and in the case of 8.35 km average passenger travel distance, the percentage of average passenger travel time reduction is 8.90%, 10.50% and 11.59%. In general, as the average passenger travel distance increases, express train service brings comparatively more value to passengers.

Table 4 Average passenger travel time reduction for instances No.10 to No.15

Instance	$|K|$	β	Avg Time Reduction (percent)
No.10	2	30 sec	70.69 sec (10.09%)
No.11	3	30 sec	88.93 sec (12.70%)
No.12	4	30 sec	89.07 sec (12.72%)
No.13	2	60 sec	58.06 sec (8.290%)
No.14	3	60 sec	70.30 sec (10.04%)
No.15	4	60 sec	74.59 sec (10.65%)
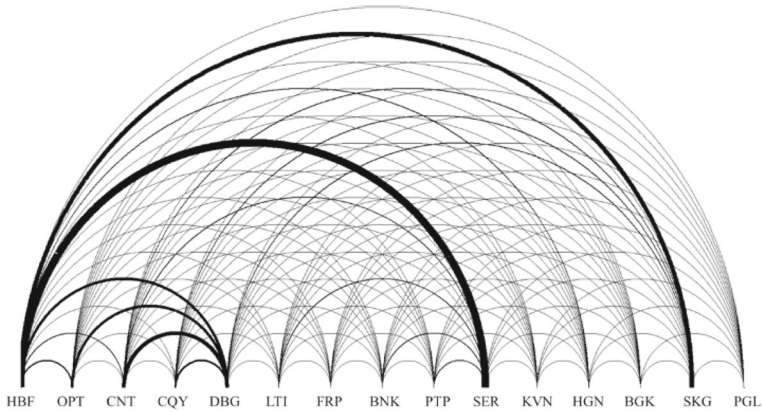
(a) Average Travel Distance 7.00 km
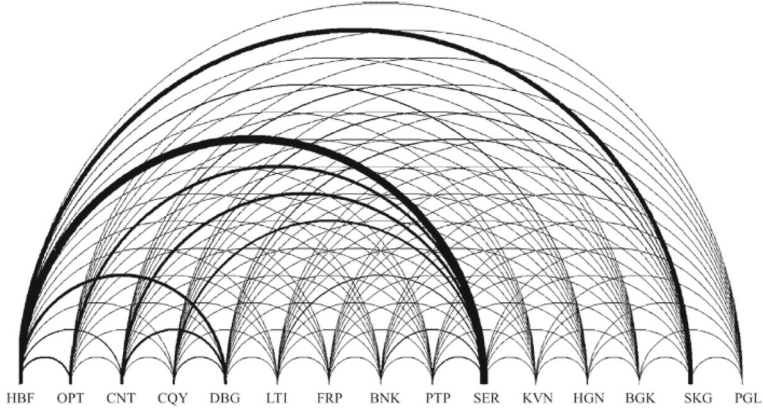
(b) Average Travel Distance 8.35 km

Fig. 10 Demand origin destination distribution with different average travel distance

Table 5 Average passenger travel time reduction for instances No.1 to No.9

Instance	$|K|$	Avg Travel Distance	Avg Time Reduction (percent)
No.1	2	6.54 km	51.50 sec (5.95%)
No.2	3	6.54 km	71.70 sec (8.28%)
No.3	4	6.54 km	79.20 sec (9.15%)
No.4	2	7.00 km	64.97 sec (7.51%)
No.5	3	7.00 km	74.21 sec (8.58%)
No.6	4	7.00 km	80.52 sec (9.31%)
No.7	2	8.35 km	77.02 sec (8.90%)
No.8	3	8.35 km	90.85 sec (10.50%)
No.9	4	8.35 km	100.3 sec (11.59%)

Table 6 Average passenger travel time reduction for instances No.16 to No.18

Instance	$|K|$	β	Avg Time Reduction (percent)
No.16	2	30 sec	56.63 sec (6.54%)
No.17	3	30 sec	79.31 sec (9.17%)
No.18	4	30 sec	89.69 sec (10.36%)

5.4 Passenger Transfer

We implement the model with the revised objective function in a set of numerical experiments including instances No.16 to No.18, and evaluate the performance of CESS considering passenger transfer. The average passenger travel time reduction in CESS with passenger transfer is greater compared to that in CESS without passenger transfer. In the case of the original demand with average passenger travel distance equals to 6.54km. As the results shown in Table 6, the percentage of average passenger travel time reduction for CESS with passenger transfer is 6.54%, 9.17% and 10.36% for $|K| = 2$, 3 and 4, respectively (compared to 5.95%, 8.28% and 9.15% for CESS without passenger transfer). Potential passenger transfer between express trains will enhance the benefits of express subway service to a larger extent.

6 Conclusion

This paper addresses the cyclic express subway service (CESS) problem. We study routine express subway service with cyclic station-skip patterns and propose an optimization model to determine cyclic station-skip patterns for express trains operating in a single-track subway system. The aim is to reduce passenger travel time, given demand intensity and distribution and train headway, frequency, and capacity. An exact Mixed Integer Programming (MIP) model is formulated.

We implement the model in numerical experiments using real-world data from Singapore MRT, the North East Line and the Downtown Line. We find that the proposed exact MIP can provide optimal cyclic express service patterns for the Downtown Line with 10 stations within a reasonable computational time. For larger instances of the North East Line with 15 stations, a heuristic is proposed to improve computational efficiency further. Average travel time for passengers could be significantly reduced compared to local train service. We also discuss the potential transfer of passengers between express trains and evaluate its effects using numerical experiments.

A natural extension for future research would be extending the model from one subway line to a subway network with multiple lines, in which train synchronization and passenger transfer between lines at transfer stations are important. Another extension would be to include the stochastic passenger demand in MIP modeling. In this paper, we model the stochastic passenger demand by doing experiments with many instances with different demand distributions. Robust or stochastic optimization methodology could be valuable in future research.

References

Cao Z, Ceder AA (2019) Autonomous shuttle bus service timetabling and vehicle scheduling using skip-stop tactic. Transp Res Part C Emerg Technol 102:370–395

Ceder A, Wilson NHM (1986) Bus network design. Transport Res B-Meth 20(4):331–344

Chen J, et al. (2015) Design of limited-stop bus service with capacity constraint and stochastic travel time. Transp Res E Logist Transp Rev 83:1–15

Chiraphadhanakul V, Barnhart C (2013) Incremental bus service design: combining limited-stop and local bus services. Public Transport 5(1–2):53–78

Dong X, et al. (2020) Integrated optimization of train stop planning and timetabling for commuter railways with an extended adaptive large neighborhood search metaheuristic approach. Transp Res Part C Emerg Technol 117:102681

Farahani RZ, et al. (2013) A review of urban transportation network design problems. European J Oper Res 229(2):281–302

Freyss M, Giesen R, Muñoz JC (2013) Continuous approximation for skip-stop operation in rail transit. Procedia Soc Behav Sci 80:186–210

Furth PG, Rahbee AB (2000) Optimal bus stop spacing through dynamic programming and geographic modeling. Transp Res Rec 1731(1):15–22

Gao Y, Yang L, Gao Z (2018) Energy consumption and travel time analysis for metro lines with express/local mode. Transp Res D Transp Environ 60:7–27

Gao Y, et al. (2016) Rescheduling a metro line in an over-crowded situation after disruptions. Transport Res B-Meth 93:425–449

Guihaire V, Hao J-K (2008) Transit network design and scheduling: A global review. Transp Res Part A Policy Pract 42(10):1251–1273

Jamili A, Aghaee MP (2015) Robust stop-skipping patterns in urban railway operations under traffic alteration situation. Transp Res Part C Emerg Technol 61:63–74

Kaspi M, Raviv T (2013) Service-oriented line planning and timetabling for passenger trains. Transp Sci 47(3):295–311

Larrain H, Giesen R, Muñoz JC (2010) Choosing the right express services for bus corridor with capacity restrictions. Transp Res Rec 2197(1):63–70

Larrain H, Muñoz JC, Giesen R (2015) Generation and design heuristics for zonal express services. Transp Res E Logist Transp Rev 79:201–212

Leiva C, et al. (2010) Design of limited-stop services for an urban bus corridor with capacity constraints. Transport Res B-Meth 44(10):1186–1201

Liebchen C (2008) The first optimized railway timetable in practice. Transp Sci 42(4):420–435

Melkote S, Daskin MS (2001) An integrated model of facility location and transportation network design. Transp Res Part A Policy Pract 35(6):515–538

Niu H, Zhou X, Gao R (2015) Train scheduling for minimizing passenger waiting time with time-dependent demand and skip-stop patterns: Nonlinear integer programming models with linear constraints. Transport Res B-Meth 76:117–135

Parbo J, Nielsen OA, Prato CG (2018) Reducing passengers' travel time by optimising stopping patterns in a large-scale network: A case-study in the Copenhagen Region. Transp Res Part A Policy Pract 113:197–212

Stewart C, El-Geneidy A (2016) Don't stop just yet! A simple, effective, and socially responsible approach to bus-stop consolidation. Public Transport 8(1):1–23

Suh W, Chon K-S, Rhee S-M (2002) Effect of skip-stop policy on a Korean subway system. Transp Res Rec 1793(1):33–39

Sun L, et al. (2014) Demand-driven timetable design for metro services. Transp Res Part C Emerg Technol 46:284–299

Ulusoy YY, Chien SI-J, Wei C-H (2010) Optimal all-stop, short-turn, and express transit services under heterogeneous demand. Transp Res Rec 2197(1):8–18

Yang H, Bell MGH (1998) Models and algorithms for road network design: a review and some new developments. Transp Rev 18(3):257–278

Yin Y, Madanat SM, Lu X-Y (2009) Robust improvement schemes for road networks under demand uncertainty. European J Oper Res 198(2):470–479

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.