Article:

![White Rose university consortium logo - Universities of Leeds, Sheffield & York]

Designing Robust Schedule Coordination Scheme for Transit Networks with Safety Control Margins

WeitiaoWu[ab], Ronghui Liu[b1] and Wenzhou Jin[a]

a. Department of Transportation Engineering, South China University of Technology, Guangzhou 510641, China

b. Institute for Transport Studies, University of Leeds, Leeds, LS29JT, United Kingdom

E-mail addresses: ctwshen@163.com; R.Liu@its.leeds.ac.uk; ctwzhjin@scut.edu.cn

Abstract

We propose a robust schedule coordination scheme which combines timetable planning with a semi-flexible departure delayed control strategy in case of disruptions. The flexibility is provided by allowing holding for the late incoming bus within a safety control margin (SCM). In this way, the stochastic travel time is addressed by the integration of real-time control and slacks at the planning phase. The schedule coordination problem then jointly optimises the planning headways and slack times in the timetable subject to SCM. Analytical formulations of costs functions are derived for three types of operating modes: uncoordinated operation, departure punctual control and departure delayed control. The problem is formulated as a stochastic mixed integer programming model and solved by a branch-and-bound algorithm. Numerical results provide an insight into the interaction between SCM and slack times, and demonstrate that the proposed model leads to cost saving and higher efficiency when SCM is considered. Compared to the conventional operating modes, the proposed method also presents advantages in transfer reliability and robustness to delay and demand variation.

Keywords: Schedule coordination; Robust; Stochastic travel time; Holding control; Branch-and-bound algorithm.

---

[1] Corresponding Author: Tel: +44(0)113 3435338; Fax: +44(0) 113 3435334; Email : r.liu@its.leeds.ac.uk

1. Introduction

Traffic congestion and air pollution are major concerns in urban development. It is widely anticipated that the shift from auto to public transport can significantly mitigate these problems. With large capacity and ease to penetrate urban networks, bus transit system plays an important role in developing sustainable urban transportation systems. A good transit network design involves several hierarchically-related procedures including: line planning, timetable generation, vehicle scheduling, and crew schedule (Ceder and Wilson, 1986; Ceder, 2007). This paper is concerned with the timetable coordination problem, which aims to maximize the number of simultaneous arrivals at a transfer station and minimize the transfer waiting time cost to passengers.

A well-designed timetable can greatly improve transit service reliability by providing users with seamless transfers. However, due to the stochastic nature of public transit attributes, such as variable travel time, fluctuating demand and changeable weather, vehicles may experience unplanned riding time and as such are difficult to adhere to the scheduled arrival time at transfer nodes (Hadas and Ceder, 2010; Wu et al., 2015). This results in transfer failure and passengers are forced to wait for the next bus, which may lead to late arrival at work. It has been shown that bus passengers value their waiting time twice as much as their time on board travelling (Quarmy, 1967), and that they value late arrival four times higher than mean travel time (Hollander and Liu, 2008). Therefore, it is important to design robust timetable and schedule coordination taking into account the above mentioned uncertainty.

The present paper explores ways to improve the transfer reliability and schedule coordination by integration of strategies currently employed separately at the planning level and the operation level. This is done by applying probabilistic approximations of generalized cost to the network with the information about delay distribution and demand. We show that when combining operational level controls with timetable planning, the results are reduced planned slack time in (and therefore more efficient) timetable, and more reliable schedule coordination and transfer success. We derive the analytical formulations for three types of operation modes. The relationships developed in this study can be used by policy planners and transit agencies to help determine the schedule design for given desired service level.

2. Literature Review and Contributions of the Present Paper

There is a wealth of studies on timetabling and scheduling problems for public transport. Ceder et al. (2001) proposed that the timetable generation problem can be considered as maximizing the number of simultaneous bus arrivals at transfer nodes, and formulated the problem as a mixed integer linear programming (MIP) model. The work is extended by Eranki (2004), where the condition of synchronization is relaxed and re-defined as the arrival of vehicles at a transfer station within a small window of separation time. Following field observation, Ibarra-Rojas and Rios-Solis (2012, 2015) developed a flexible timetable synchronization model for a typical bus network, in which different arrival time windows are set for different buses. The objectives of the problem are to maximize the number of synchronizations for transferring passengers and avoid bus bunching along the common line. Wong et al.

(2008) presented a MIP model for schedule coordination problem in a railway system with the objective to minimize the total interchange waiting time, by adjusting trains' departure time, ride time and dwell time.

The above mentioned models are mostly developed to deal with deterministic scenarios at the planning stage. In reality, passenger demand varies over time and at different stops which can lead to variations in bus dwell times (Fonzone et al., 2015). There is also uncertainty in bus journey times due to traffic incidents, weather effects, etc. To improve the reliability of schedule coordination, there have been studies to impose a buffer time in the schedule to absorb the travel time variability and to ensure certain degrees of connectivity. Ting and Schonfeld (2005) investigated a schedule coordination problem by jointly optimizing the headways and slack times at transfer nodes to minimize the total cost of the transit network. Wu et al. (2015) developed a timetabling model with stochastic travel time by adding a slack time. The model is formulated to optimize the departure times of all buses from all lines, with the goal of minimizing the total waiting time cost for three types of passengers: transferring passengers, boarding passengers and through passengers. The authors reported that the model is relatively effective if the ratio of the through-passengers to transfer passengers is less than a critical value. Parbo et al. (2014) addressed a timetable optimization problem to minimize the transfer waiting time while considering passengers' route choice. The problem is formulated as a bi-level programming model, in which the lower level is a transit assignment model and the upper level is to minimize the weighted transfer waiting time. There are also literatures that investigate the schedule coordination optimization for intermodal transport system. Lee and Schonfeld (1991) developed an analytical model that optimizes the slack time for simple systems with transfer between one bus route and one rail line at the planning level. Chien and Schonfeld (1998) presented a model to jointly optimize the layout of a rail line with feeder bus lines, and the operational parameters (e.g., bus headways) in an urban corridor. Later, Chien and Schonfeld (2002) formulated another model to optimize the intermodal transit coordination problem considering the stochastic feeder vehicle arrivals and deterministic train arrivals. The decision variables are the buses and rail headways and slack times on bus routes.

At the operation stage, there have also been studies which focus on improving the reliability of transit coordination through real-time control strategies. Among which, holding strategy is most commonly used. The strategy can be applied both for regular stops in a bus corridor and at transfer stations (Ibarra-Rojas et al., 2015). The former is used to delay bus movement deliberately when a bus is ahead of the schedule and maintain regularity (Daganzo, 2009; Xuan et al., 2011; Dalgado et al., 2012; Eberlein et al., 2001; Hickman, 2001; Hernandez et al., 2015); while the latter involves holding the ready-to-depart vehicle to wait for the delayed vehicle(s) in order to ensure planned connection. Such models require real-time data and accurate prediction of vehicle arrival times.

Chowdhury and Chien (2001) developed a dynamic vehicle dispatching model at an intermodal transfer station by making use of real-time information (e.g., estimated vehicle arrival times and transfer demand) provided by advanced public transportation systems. Ting and Schonfeld (2007) proposed a dispatch control model to optimize the dispatch decision for a multi-route and multi-hub transit network. Chung and

Shalaby (2007) developed a holding simulation model for connection protection for intermodal transfer, and applied the model to evaluate the bus system in the city of Brampton, Canada. There are also studies addressing the holding strategy at transfer nodes when stochastic disturbances occur (e.g. Yu et al., 2012; Dessouky et al., 1999, 2003; Chowdhury and Chien, 2001). To mitigate the uncertainty about simultaneous arrivals of different vehicles at a transfer station, Hadas and Ceder (2010) developed a simulation model incorporating operational tactics such as hold, skip stop and short-turn, and a dynamic programming was developed for increasing the number of direct transfers and minimizing the total travel time. Nesheli and Ceder (2014) enhanced the simulation model by introducing the concept of skip-a-segment in addition to skip-a-stop. Recently, Sun and Schonfeld (2016) proposed a vehicle holding method for intermodal freight operations by considering decision risks and correlations among vehicle arrivals.

Most of the existing literatures deal with stochastic travel time aspect of schedule coordination problem at either the planning stage or the operation stage, not both. The solutions adopted at the two stages are different: adding a slack time in the schedule at the planning stage, while the operation stage relying on holding control. However, each has its own drawbacks. Real-time holding control could propagate disturbances to the subsequent downstream trips, as such impose adverse impacts to other vehicles and lead to bus bunching problems for example. Thus pure dynamic corrective (holding) actions without sufficient slack may adversely affect the operation stability. On the other hand, although a well-designed schedule plan considering uncertainty at the planning stage can reduce the need to frequently relying on control tactics, it is achieved at the expense of increased travel time and cost (with the addition of slack times), which reduce the commercial speed of buses and thus reduce the efficiency (Daganzo, 2009). In addition, the pre-designed timetable may not be feasible any more once vehicle disruptions arise without any dynamic control tactics, since there still exists unexpected lateness that cannot be prevented at the planning stage. One of the greatest problems facing transit agencies is maintaining robustness while achieving high efficiency (Berrebi, et al., 2015). To the best of the authors' knowledge, there is so far no research on schedule coordination that combines timetable synchronization methods at the planning stage with holding strategies at the operation stage, and explores the 'communication' between what a planned schedule entails and what real-time operations do. The schedule scheme may be more robust against uncertainty if we combine the advantages of both levels. The challenge is how to improve robustness and service quality without significantly increasing additional operation cost.

Our study explores this idea. In this paper, we propose an innovative robust schedule coordination scheme for a transit network, in which the stochastic travel time aspect of the problem is addressed at both the planning stage and the operation stage. The former is by adding a slack time into schedule while the latter is through a safety control margin. The concept of safety control margin is defined as follows: if the lateness of an incoming bus at a transfer station is not sufficiently compensated by the slack time (which had already been built in the planning stage) but the predicted arrival time is before a maximum allowable holding time (termed as the safety control margin), the ready-to-depart bus will be held for transfer. The safety control margin is set such that the resulting departure tardiness caused by holding can be totally absorbed by driver's schedule recovery effort before it reaches the next transfer station. Essentially, the

dynamic characteristics of the flexible control strategy requires operators to make difficult trade-off between economic viability of the system and maintaining good service for passengers, and trade-off between different groups of passengers (e.g., through passengers and transferring passengers).

Our main contributions in this paper are therefor as follows: (1) we propose a co-optimisation model of timetable generation which combines operational-level control strategies into planning-level schedule coordination; (2) we obtain analytical formulations in the forms of integrable functions which feature high computational efficiency compared to previous studies with numerical integration for unintegrable functions; (3) we derive an analytical formulation to calculate transfer failure rates for a given planning and control policy, as a measure of transfer reliability; (4) we compare the performance of three operational control strategies in timetable planning: uncoordinated operation, departure punctual control and departure delayed control in two typical bus networks; (5) we demonstrate that by including the real-time control strategy (i.e. the safety control margin setting) in the planning stage of setting the slack times, more cost-effective timetables with smaller slack times can be achieved. As far as we are aware, this is the first time the schedule coordination problem is addressed at both the planning and the operation stage.

The rest of the paper is organized as follows. In Section 3, we introduce the model. In Section 4, the solution algorithms are developed. In Section 5, numerical examples are performed. Finally, we conclude in Section 6.

3. Modelling Approach

3.1 Problem description
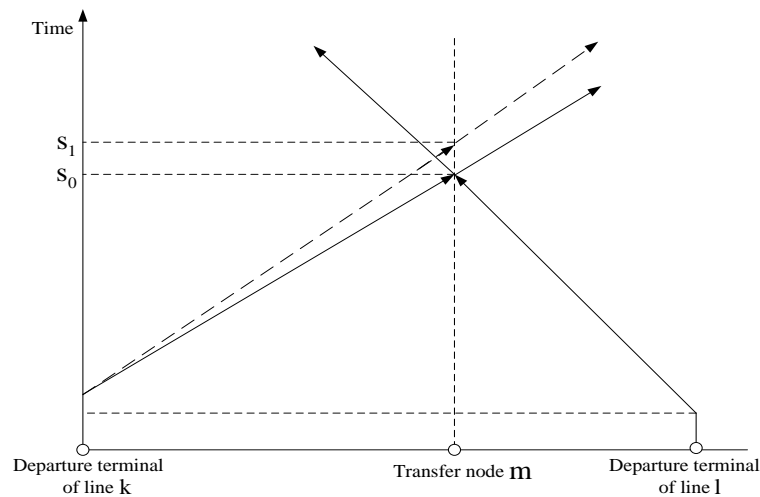
Fig.1 A diagrammatic illustration of the interchange process between two transit lines.

In this section, we revisit the schedule coordination problem and introduce the rational behind our proposed model. The modelled bus network consists of a set of lines and nodes. The nodes are shared by multiple bus lines for passenger interchange, commonly termed as transfer nodes. (We consider the terminals and transfer stations in the network only, and ignore the intermediate bus stops between

consecutive transfer stations.) The transferring process of passengers between two lines is shown in a time-space diagram in Fig.1. In a deterministic scenario, the optimal departure time of line k and line l is scheduled at time s_0. However, the travel time of buses can be unpredictable due to many external factors, such as traffic condition. In case of delay, the vehicle of line k may reach transfer node m at a later time s_1. If vehicle of line l departs on time at s_0, then the transfer passengers will miss the scheduled connection and have to wait for the next bus. In order to improve the reliability of schedule coordination, a slack time is usually built in the timetable as an effective strategy to absorb such travel time uncertainty (Randolph, 2001; Wu et al., 2015). In reality, however, there may still be delayed arrival that goes beyond the slack time.

To solve this problem, we propose a robust schedule coordination method that combines the operational strategies intro timetable planning. At the planning stage, a slack time is added onto the schedule to ensure higher probability of scheduled transfer; while at the operation stage, a dynamic holding time is applied to realize a connection in case of delays beyond the slack time. Our schedule coordination method optimizes the planning headways and slack times in the timetable, subjective to a given holding control specification, with an overall objective to minimize monetary cost and transfer failure.

As soon as the transfer is realized, the holding bus can depart even if it is before a maximum allowable holding time. This maximum allowable holding time is termed as safety control margin (SCM). Such a system aims to inject flexibility by providing 'additional' time for connection protection, which can be used in a real-time holding framework to allow deviation from scheduled departure time within a time window.

The SCM is set so that the resulting tardiness or loss time can be 'absorbed' before the next transfer node. Here, we highlight that the value of SCM should be set to ensure full schedule recovery. In other words, the driver will manage to maintain/recover the schedule at the next transfer node. In practice, schedule recovery effort can be made section by section along the bus route with the information of schedule deviation provided by the intermediate bus stops or time points. Since there are a number of sections between contiguous transfers, the distance is generally long enough to realize full schedule recovery. To this end, we suggest that the value of SCM should be not greater than a predetermined threshold in view of traffic safety requirement (e.g., valid speed limit). While the valid speed limits vary period by period, engineering experience shows that the variable speed limit is closely related to the traffic conditions (Lu, 2003). Therefore, in practice, the threshold can be tuned or given by rule-of-thumb in specified periods of time and be flexibly adjusted based on real situations.

To illustrate the concept of SCM, we consider a scenario whereby a ready-to-depart bus (let's call it Bus B) is waiting for a connecting bus (Bus A) which is behind the schedule. If Bus A is predicted to arrive within the SCM (for example within 30 seconds after the slack time), then Bus B will be held for transfer connection and depart as soon as the transfer process finishes. Otherwise, bus B will depart on time. Once the holding strategy is adopted, a schedule recovery effort is injected into the bus B, such that bus B driver

should devote his/her effort to catch up the scheduled arrival time within limited bus stops along the route and before the next transfer node. For example, if the bus is held for 30s, and the maximal recovery time for one bus stop is 10 seconds, then only the passengers of the next two stops downstream will be affected, and the extra waiting time for passengers in the first stop and the second stop is 20 seconds and 10 seconds, respectively. Such schedule recovery effort is widely adopted in the context of advanced public transport systems (APTS) (e.g. Yan et al., 2012; Lin and Bertini, 2004; Chen et al., 2005; Ji et al., 2014). However, to simplify the model, in this study we use a continuum approach to develop a closed-form expression to capture such interaction; we describe the approach in the model formulation section. Such a scheme, which we call semi-flexible control to underline the fact that the holding strategy is only used within SCM, is analogous to a hard time window. This tactic requires the APTS to predict the accurate and reliable arrival time of the connecting bus and convey that to the ready-to-depart bus, which is attainable in cases where buses are isolated through dedicated lanes or exclusive right of ways. In the cases of normal lanes, prediction errors may exist in disseminated arrival times since buses are exposed to the general traffic. However, since the prediction arrival time can be updated continuously in a real-time manner with the vehicle positioning data, higher accuracy can be expected when the bus position is approaching to the transfer point. For the bus predicted to arrive within SCM, its position is generally quite close to the transfer point upon making the dispatch-or-not decision, thus the prediction error at that instance is small enough to be negligible.

Although imposing the slack time and holding strategy into schedule improves the robustness of coordination and benefits the transferring passengers, it increases the through-passengers' on-board waiting time and bus companies' operating cost. Therefore, the problem involves a trade-off pertained to economic viability of the system versus maintenance of good service for passengers (i.e., extra operating cost versus reduced waiting time cost), as well as trade-offs among various types of waiting time components (i.e., on-board passengers and transferring passengers). This paper develops a stochastic programming model for robust coordination schedule scheme with stochastic travel time (RCSS-STT). In the model, the headways of bus routes and slack times at transfer nodes are jointly optimized with a given SCM to minimize the network-wide total cost, including the user cost and operating cost.

3.2 Assumptions and notations

In the following section, the fundamental elements of RCSS-STT are described, and the system cost components are derived. To facilitate the model development, some assumptions are made and presented as follows:

(A1) Passenger demand and bus travel time information is given as an input, which is assumed to be available from historical data.

(A2) The line headways in the network are usually large (generally longer than 10 min) in the planning stage, and the SCM is relatively small compared to headway.

(A3) Bus (delay) arrival times of different lines are statistically independent with given continuous distributions, and the parameters can be gathered from historic data.

(A4) There is no capacity limitation on buses so that they can satisfy all demands.

(A5) The maximal recovery time for each bus stop is an equal constant for simplicity.

(A6) The transfer time between buses is assumed to be zero, which is reasonable because the minimum transfer time for passengers can be simply added into the travel time as input.

(A7) The SCM value is set to ensure the full schedule recovery, as explained in Section 3.1, and is subject to an upper bound.

Before formulating RCSS-STT, we first define the following notations in Table 1.

Table 1 List of primary symbols, definitions and units

Symbol	Definition	Unit
Indices and Sets		
k, l	Indexes of lines in the bus network	—
m	Indexes of transfer nodes	—
L	Set of connection lines at a transfer node	—
Model parameters		
B_k	The unit vehicle operating cost for line k, for simplicity we assume $B_k = \bar{B}$	\$/veh/min
μ_w	Passenger unit waiting time value	\$/pax/min
λ_{mk}	The parameter of delayed arrival time distribution for line k connecting to transfer node m	1/min
ρ	The unit monetary incentive to the drivers per holding action	\$
t_w	An arrival time window under schedule-dependent arrival behaviour	min
α	Discount factor of the waiting time under schedule-dependent arrival behaviour	—
Auxiliary variables		
Q_k	The boarding demand on line k	pax/min
Q_{mk}	The through passenger demand at transfer station m on line k	pax/min
q_k^l	The transfer passenger flow from line $l \in L$ to line k	pax/min
q_l^k	The transfer passenger flow from line k to line $l \in L$	pax/min
q_{mk}	The passengers on the link immediate downstream of station m on line k which experienced extra delay from holding control	pax/min
R_k	The total transfer demand to line k	pax/min
H	The safety control margin	min
C_o	The vehicle operating cost	\$/min

C_w	The passenger waiting time cost	\$/min
C_t	The total transfer waiting time cost for uncoordinated operation	\$/min
C_c	The missed connection cost	\$/min
C_f	The delayed connection cost	\$/min
C_s	The induced tardiness cost, as the result of introducing slack time (and SCM) in the planned timetable	\$/min
C_r	The monetary incentives for schedule recovery effort	\$/min
C_d	The extra waiting time cost for downstream passengers	\$/min
C	The network-wide total cost	\$/min
Decision variables		
s_{mk}	The slack time on line k at transfer node m	min
h_k	The headway of line k	min

1　Scheduled departure time in previous station
2　Scheduled arrival time (SAT)
3　Actual delayed arrival time (T)
4　Scheduled departure time (SDT)
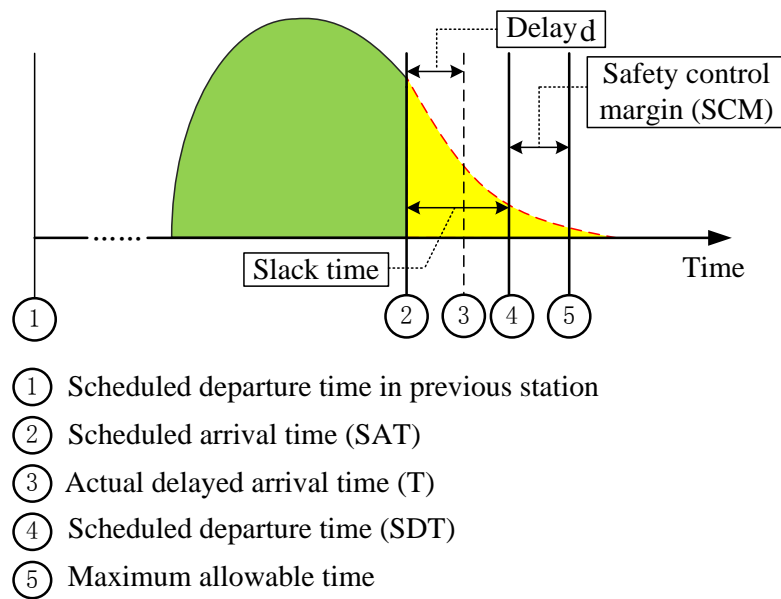5　Maximum allowable time

Fig.2 Definition of concepts.

We consider a schedule-based transit system in which buses either depart on schedule or after the schedule if they arrive late to the checkpoint. In this system, the lateness of bus arrival times follows a known probability distribution. As Fig.2 shows, let SAT be the scheduled arrival time for a bus on a feeder line, T the actual (random) delayed arrival time for the line, and SDT the scheduled departure time. If a vehicle arrives at a transfer node earlier, it has to wait until SDT.

The delay arrival distribution is associated with bus arrival time distribution and SAT setting. According to field observations conducted by Guenthner and Hamat (1988) and Strathman and Hopper (1993), bus

arrival time at a designated stop follows a right-skewed unimodal distribution and can be approximated by a known theoretical distribution, such as a lognormal or gamma distribution. Moreover, in the practice of timetable planning, the SAT is often decided by rule-of-thumb, commonly set as 85-percentile observed arrival time to achieve a reliable timetable (Muller and Furth, 2000), in which 85% of buses will arrive early (before SAT) and depart on time as illustrated by the dark (green) shaded area in Fig.2. Therefore, given the right-skewed unimodal pattern of bus arrival time distribution and high percentile value (85%) of SAT, the delay arrival distribution beyond SAT must be a monotonically decreasing function, as shown by dashed curve at the edge of the light (yellow) shaded area in Fig.2. For analytical tractability, it is reasonable to approximate the delay distribution with an exponential distribution that is a monotonically decreasing function. This is consistent with Randolph (2001) and Bookbinder and Desilets (1992).

Thus the delay time $d = T - SAT$ follows the exponential distribution given below:

$$p(T - SAT > t) = p(d > t) = e^{-\lambda t} \tag{1}$$

and the probability density function of the delay time is as follows:

$$f(t) = \lambda e^{-\lambda t} \tag{2}$$

where t is the general delay, which can be partially absorbed by the slack time; λ is the parameter of the exponential distribution, and $1/\lambda$ and $1/\lambda^2$ are the mean delay and the variation, respectively.

The distribution parameter λ can be calibrated or estimated from historical data collected from the automatic vehicle location system (Yan et al., 2012; Wu et al., 2015). By an exponential delay arrival assumption, we avoid arrival time distributions such as normal distributions with both negative and positive infinity. Meanwhile, we can obtain analytical formulations in the forms of integrable functions with high computational efficiency. It should be noted that delay $T - SAT$ must be truncated since infinite arrival time seldom exists in the real world (Wu et al., 2013; 2014). Therefore, truncated exponential distribution is more suitable. However, the truncated exponential distribution brings the difficulties in analytical derivation. For the sake of simplicity, non-truncated exponential distribution is assumed in this paper. The models with truncated distribution will be explored in our future works.

Lemma1. At the planning level, the same scheduled departure time (SDT) for different bus lines sharing a transfer station ensures the maximum transfer probability.

Proof. If the connecting bus lines have the same SDT, i.e., $t_{dep}^1 = t_{dep}^2$, the transferring passengers in the buses of both lines arriving in the scheduled slack time can make a successful connection to each other. Otherwise, if the SDT of bus line 1 is earlier than that of bus line 2, i.e, $t_{dep}^1 < t_{dep}^2$, the transferring passengers in the bus of line 2 arriving between the SDT of two bus lines, i.e, $t \in (t_{dep}^1, t_{dep}^2)$ will miss connections and have to wait for the next bus of line 1.

Q.E.D

Lemma1 is similar to the rationale of exploring the maximum encounter probability as discussed by Hadas and Ceder (2010), which focuses on exploring maximum overlapping of time-space trajectories. However, in our study, this condition is enhanced by the same SDT, which ensures maximum overlapping of time-space trajectories as well as maximum transfer probability.

3.3 Operation modes and systematic cost components

In this paper, we analyse three operation modes: (a) uncoordinated operation; (b) departure punctual control (DPC); and (c) departure delayed control (DDC). In the uncoordinated operation, vehicles arrive at transfer nodes based on their independently scheduled headway (Ting and Schonfeld, 2005). Under DPC, all buses in the system are operated under no-holding policy, i.e., vehicles are not allowed to leave after the SDT. Different to DPC, our proposed DDC policy operates in a semi-flexible way whereby vehicles are allowed to hold for late incoming connection buses. Both DPC and DDC are coordinated operations, for which we explore simultaneously optimized headways of different lines and slack times at transfer nodes. Note that DPC is a special case of DDC where the allowed holding time is zero. For simplicity, the bidirectional headways in a bus route are set equal for vehicle turnaround periodically without internal disruption.

Since DPC and DDC are coordinated modes, the line headways should be synchronized so as to reduce transfer waiting time. Given that the bus lines should have the same SDT in a transfer node as discussed in Lemma 1, the line headways should be set as common or inter-ratio so that the schedule can be operated periodically. However, to simplify the model, only common headway is considered in this paper. Inter-ratio scenario will be investigated in the future work.

For transit networks, the systematic costs are typically determined during the planning phase and include: operating cost C_o, passenger waiting cost C_w, and passenger transfer waiting cost C_t. Note that in-vehicle waiting time cost is not included in the overall model throughout the paper since it is not related to the decision variables.

For coordinated operations, there is an addition cost associated with the introduction of the slack time, and SCM for DDC mode, which we term it as the induced tardiness cost C_s. Furthermore, we separate the transfer waiting cost into two components: missed connection cost C_c, and delayed connection cost C_f.

Table 2 lists the cost components for the different operation modes. In the following sections, we derive the analytical formulations for each cost component, for the three different operation modes. The objective of bus scheduling is to minimize the systematic network-wide total cost.

Table 2. Characteristics and key cost components for different operation modes

Operation mode	Characteristics			Components for total cost C					
	Slack time	Headway coordination	SCM	C_o	C_w	C_t	C_s	C_f	C_c

Uncoordinated				√	√	√			
DPC	√	√		√	√		√	√	√
DDC	√	√	√	√	√		√	√	√

3.4 Cost formulations for uncoordinated operation

Following Ting and Schonfeld (2005), the total operating cost is the summation of product of required fleet size T_k/h_k and the unit vehicle operating cost for each link k, as:

$$C_o = \sum_k \frac{B_k T_k}{h_k} \tag{3}$$

When the headway is relatively short, passengers arrive at bus stops randomly (i.e. schedule-blind), Osuna and Newell (1972) show that their expected waiting time $E_k\{w\}$ can be given as:

$$E_k\{w\} = \frac{1}{2} E\{h_k\}(1 + \frac{Var\{h_k\}}{E^2\{h_k\}}) \tag{4}$$

where $E\{h_k\}$ is the expected headway and can be approximated by the target headway, i.e., $E\{h_k\} = h_k$. $Var\{h_k\}$ is the variance of headways.

When the headway becomes longer, passengers tend to coordinate their arrivals with the published departure times. According to Moccia and Laporte (2016) and Tirachini et al (2010), as the headway exceeds a critical value (12 min), passengers will follow timetables and arrive at stops t_w minutes before the service time. The expected waiting time saved by such passenger behaviour still has a cost but is discounted by a factor α less than one, and is expressed as follows:

$$E_k\{w\} = t_w + \frac{1}{2}\alpha E\{h_k\}(1 + \frac{Var\{h_k\}}{E^2\{h_k\}}) \tag{5}$$

Zhao et al. (2006) verified by simulations that the variance of headway can be approximated by the variance of delay as:

$$Var\{h_k\} \approx 2Var\{d_k\} \tag{6}$$

where $Var\{d_k\}$ represents the variance of delays measured at checkpoints for line k.

On this basis, the expected waiting time can be estimated from the headways and the variance of link delays. The total waiting time is associated with the demand and the individual expected waiting time. Thus, the total waiting time serving a total boarding demand Q_k can be approximated by substituting (5) and (6) into (4) as follows:

$$C_w = \mu_w \sum_k Q_k E_k \{w\} = \mu_w \sum_k Q_k [\beta + \varphi \frac{h_k}{2}(1 + \frac{2Var\{d_k\}}{h_k^2})] \qquad (7)$$

$$\beta = \begin{cases} 0 & \text{for } h_k \leq 12\text{min} \\ t_w & \text{otherwise} \end{cases} \quad \text{and} \quad \varphi = \begin{cases} 1 & \text{for } h_k \leq 12\text{min} \\ \alpha & \text{otherwise} \end{cases}$$

Assuming that the delayed arrival time is exponentially distributed (Eq.(2)), the variance of exponential distribution is $\frac{1}{\lambda_k^2}$, then the total waiting time becomes:

$$C_w = \mu_w \sum_k Q_k [\beta + \varphi(\frac{h_k}{2} + \frac{1}{\lambda_k^2 h_k})] \qquad (8)$$

Eqs. (3) and (8) apply also to coordinated operations.

The transfer time, incurred by passengers transferring from one route to another, depends on the coordination status between routes. For uncoordinated operations, the vehicles from different routes encounter at the transfer station randomly, thus the average transfer waiting time can be estimated from inter-arrival times at transfer stops, i.e., the mean and variance of headways, which can also be obtained by Osuna and Newell (1972) in Eq. (4). Consequently, based on exponential distribution, the formula for total transfer cost is calculated as the product of the average transfer waiting time and the total number of transfer passengers and expressed as follows:

$$C_t = \mu_w \sum_k R_k (\frac{h_k}{2} + \frac{1}{\lambda_k^2 h_k}) \qquad (9)$$

where $R_k = \sum_{l \in L} q_k^l$ is the total transfer passengers from connecting lines to line k.

In summary, the network-wide total cost for uncoordinated operation in the system is the summation of operating cost, waiting time cost and transfer waiting cost, and that is,

$$C = C_o + C_w + C_t = B \sum_k \frac{T_k}{h_k} + \mu_w \sum_k (\varphi Q_k + R_k)(\frac{h_k}{2} + \frac{1}{\lambda_k^2 h_k}) + \mu_w \beta \sum_k Q_k \qquad (10)$$

According to the first order condition, the optimal headway of route k can be derived from $dC/dh_k = 0$, and is given as below:

$$h_k^* = \sqrt{\frac{2BT_k}{\mu_w(\varphi Q_k + R_k)} + \frac{2}{\lambda_k^2}} \qquad (11)$$

3.5 Cost formulations for DPC operation

Compared to the uncoordinated mode, the system cost for DPC operation contains three different components: the induced tardiness cost, the missed connection cost and delayed connection cost (see Table 2). We provide the analytical formulations for each below.

3.5.1 Induced tardiness cost C_s

The tardiness cost includes the induced cost for through (non-transfer) passengers already on board and the extra operating cost of adding the slack time. The tardiness cost is formulated as follows:

$$C_s = \sum_m \sum_k (\mu_w Q_{mk} + \frac{B}{h_k}) s_{mk} \tag{12}$$

In this case, the affected passengers are those on board the holding bus for departure. For simplicity, we assume that the waiting time value is identical for both inside and outside the vehicle. This assumption holds throughout the paper.

3.5.2 Missed connection cost at transfer stations C_c

When a passenger on line k misses a connection to line l, he/she has to wait $h_l - (d_{ml} - s_{ml})$. Following Vansteenwegen and Qudheusden (2007), since in most cases $d_{ml} - s_{ml}$ is relatively small compared to h_l for large headways, the delay term $(d_{ml} - s_{ml})$ can be ignored. Therefore, we can further assume that if a transfer passenger misses a connection, he/she has to wait for the duration of one headway. This is reasonable for a relatively large headway (A2). The missed connection cost might be slightly overestimated when the designed headway is small. However, small headway is undesirable and uneconomical for schedule coordination, this will be examined in the numerical example in Section 5.

By assumption (A3), the joint probabilities of arrivals may be obtained by simply multiplying the probabilities obtained separately from the delay distributions of different bus lines. The possible conditions are given as follows:

a) The feeder vehicle on line k arrives at node m late while its connecting vehicle on line l is not late, i.e. $d_{mk} > s_{mk}$ and $d_{ml} < s_{ml}$.

The corresponding probability is:

$$\begin{aligned} p^a &= p(d_{mk} > s_{mk} \cap d_{ml} < s_{ml}) \\ &= p(d_{mk} > s_{mk}) p(d_{ml} < s_{ml}) \\ &= e^{-\lambda_{mk} s_{mk}} (1 - e^{-\lambda_{ml} s_{ml}}) \end{aligned} \tag{13}$$

where d_{mk} and d_{ml} denote the delay time (beyond SAT) of the feeder vehicle on line k and the connecting vehicle on line l at transfer node m, respectively.

b) Both vehicles are late, but the feeder vehicle on line k arrives after vehicle on line l

The corresponding probability is:

$$\begin{aligned} p^b &= p(d_{mk} > s_{mk} \cap d_{ml} > s_{ml} \cap d_{mk} - s_{mk} > d_{ml} - s_{ml}) \\ &= p(d_{mk} > s_{mk}) p(d_{ml} > s_{ml}) p(d_{mk} - s_{mk} > d_{ml} - s_{ml}) \end{aligned} \tag{14}$$

The probability $p(d_{mk} - s_{mk} > d_{ml} - s_{ml})$ should be calculated according to the relationship between the slack time of line k and that of its connecting line l, which is shown in the following piecewise function, see detailed derivation in Appendix 2:

$$p(d_{mk} - s_{mk} > d_{ml} - s_{ml}) = \begin{cases} \dfrac{\lambda_{ml}}{\lambda_{mk} + \lambda_{ml}} e^{-\lambda_{mk}(s_{mk} - s_{ml})}(1 - e^{-\lambda_{mk}h_k}), & \text{for } s_{mk} \geq s_{ml} \\ \dfrac{\lambda_{mk}}{\lambda_{mk} + \lambda_{ml}}(1 - e^{-\lambda_{ml}(s_{ml} - s_{mk})}), & \text{for } s_{mk} < s_{ml} \end{cases} \tag{15}$$

The overall probability for missed connection is then:

$$\begin{aligned} p^c &= p^a + p^b \\ &= \begin{cases} \dfrac{\lambda_{ml}}{\lambda_{mk} + \lambda_{ml}} e^{-\lambda_{mk}s_{mk} - \lambda_{ml}s_{ml} - \lambda_{mk}(s_{mk} - s_{ml})}(1 - e^{-\lambda_{mk}h_k}) + e^{-\lambda_{mk}s_{mk}}(1 - e^{-\lambda_{ml}s_{ml}}) & \text{for } s_{mk} \geq s_{ml} \\ \dfrac{\lambda_{ml}}{\lambda_{mk} + \lambda_{ml}} e^{-\lambda_{mk}s_{mk} - \lambda_{ml}s_{ml}}(1 - e^{-\lambda_{ml}(s_{ml} - s_{mk})}) + e^{-\lambda_{mk}s_{mk}}(1 - e^{-\lambda_{ml}s_{ml}}) & \text{for } s_{mk} < s_{ml} \end{cases} \\ &= \begin{cases} Ae^{-\lambda_{mk}(s_{mk} - s_{ml})}(1 - e^{-\lambda_{mk}h_k}) + B & \text{for } s_{mk} \geq s_{ml} \\ A(1 - e^{-\lambda_{ml}(s_{ml} - s_{mk})}) + B & \text{for } s_{mk} < s_{ml} \end{cases} \end{aligned} \tag{16}$$

where $A = \dfrac{\lambda_{ml}}{\lambda_{mk} + \lambda_{ml}} e^{-\lambda_{mk}s_{mk} - \lambda_{ml}s_{ml}}$, and $B = e^{-\lambda_{mk}s_{mk}}(1 - e^{-\lambda_{ml}s_{ml}})$.

The total missed connection costs are also calculated with the following piecewise function:

$$\begin{aligned} C_c &= \mu_w \sum_m \sum_l \sum_k q_l^k h_l \, p^c \\ &= \mu_w \sum_m \sum_l \sum_k q_l^k h_l \cdot \begin{cases} [Ae^{-\lambda_{mk}(s_{mk} - s_{ml})}(1 - e^{-\lambda_{mk}h_k}) + B] & \text{for } s_{mk} \geq s_{ml} \\ [A(1 - e^{-\lambda_{ml}(s_{ml} - s_{mk})}) + B] & \text{for } s_{mk} < s_{ml} \end{cases} \end{aligned} \tag{17}$$

3.5.3 Delayed connection cost at transfer station C_f

In contrast to missed connection, transfer passengers can make a successful connection but have to wait for relatively shorter time. Delayed connection cost of transfer passengers from the feeder vehicle on line k to connecting vehicle on line l at transfer station m should account for the following three possible conditions:

Case A: Both vehicles are not late, no matter whether the connecting bus on line l arrives after or before the feeder bus on line k

Case B: The feeder vehicle on line k arrives early while the connecting vehicle on line l is late

Case C: Both vehicles are late, but the connecting vehicle in line l arrives after the feeder vehicle in line k

The waiting time and respective conditions for each case are shown in Table 3.

Table 3. Different cases for delayed connection

Scenarios	Cost notations	Waiting time	Conditions
Case A	C_f^1	$s_{mk} - d_{mk}$	$0 \leq d_{mk} \leq s_{mk}, 0 \leq d_{ml} \leq s_{ml}$
Case B	C_f^2	$d_{ml} - s_{ml} + s_{mk} - d_{mk}$	$0 \leq d_{mk} \leq s_{mk}, s_{ml} \leq d_{ml} \leq h_l$
Case C	C_f^3	$d_{ml} - s_{ml} + s_{mk} - d_{mk}$	$s_{mk} \leq d_{mk} \leq s_{mk} - s_{ml} + d_{ml}, s_{ml} \leq d_{ml} \leq h_l$

Then the total delayed connection cost of transfer passengers is the summation of all possible conditions, see detailed derivation in Appendix 3.

$$C_f = C_f^1 + C_f^2 + C_f^3$$

$$= \mu_w \sum_m \sum_l \sum_k q_l^k [s_{mk} + \frac{1}{\lambda_{mk}}(e^{-\lambda_{mk}s_{mk}} + e^{-\lambda_{ml}h_l} - e^{-\lambda_{ml}(s_{mk}-s_{ml}+h_l)-\lambda_{ml}h_l} - 1) - (h_l - s_{ml} + s_{mk})e^{-\lambda_{ml}h_l}$$

$$- \frac{1}{\lambda_{ml}}(e^{-\lambda_{ml}h_l} - e^{-\lambda_{ml}s_{ml}}) + \frac{1}{\lambda_{mk}+\lambda_{ml}}e^{-\lambda_{mk}s_{mk}}(e^{-(\lambda_{mk}+\lambda_{ml})h_l+\lambda_{mk}s_{ml}} - e^{-\lambda_{ml}s_{ml}})] \tag{18}$$

3.6 Cost formulations for DDC operations

DDC operations allow bus holding at the operation stage to enable connections, subject to a maximum holding time SCM. The length of SCM is denoted as H. The derivations of the cost components for DDC follow those for DPC, with additional consideration on both the positive and negative impacts of SCM.

3.6.1 Induced tardiness cost C_s

In this case, induced tardiness time cost should consider the impact of holding strategy compared to DPC. The induced tardiness cost component in DPC should be modified considering the expected holding time. In addition, extra control cost, extra waiting time cost for downstream passengers should also be included.

Modifying Eq. (12) gives the following tardiness cost for DDC as:

$$C_s = \sum_m \sum_k (\mu_w Q_{mk} + \frac{B}{h_k})(s_{mk} + E_{mk}(H)) + C_r + C_d \tag{19}$$

where $E_{mk}(H)$ is the expected holding time for line k at transfer node m with a given SCM value H; C_r and C_d denote the extra control cost and the extra waiting time cost for downstream passengers, respectively. The expression of $E_{mk}(H)$ is presented as follows, for detailed derivation see Appendix 4.

$$E_{mk}(H) = \min_{i,i\neq k}\{-\frac{\lambda_{mi}}{\sum_i \lambda_{mi}}[He^{-\sum_i \lambda_{mi}(H+s_{mi})} + \frac{e^{-\sum_i \lambda_{mi}s_{mi}}}{\sum_i \lambda_{mi}}(e^{-\sum_i \lambda_{mi}H}-1)]$$

$$-[He^{-\lambda_{mi}(H+s_{mi})} + \frac{e^{-\lambda_{mi}s_{mi}}}{\lambda_{mi}}(e^{-\lambda_{mi}H}-1)]\prod_{l,l\neq i}(1-e^{-\lambda_{ml}s_{ml}})\} \tag{20}$$

Note in the above formulas that lines i, l and k share the same transfer node m.

Now we derive the formulations of C_r and C_d below.

➤ Extra control cost C_r

If the schedule recovery is needed, this will result in an increased capital cost and additional labour cost for drivers. However, we can disregard possible infrastructure investment such as global-positioning systems and wireless communication technology, which are actually a very small fraction of the long-term transit operation costs (Qiu et al., 2014), and this technology is readily available in the context of APTS mentioned earlier. Thus, the operating costs can be reasonably considered to be the normal operation cost plus the addition control cost for drivers. The additional labour cost is calculated as the product of the occurrence probability of holding and unit financial incentive.

We assume that the holding time can be fully covered by drivers' schedule recovery effort under SCM as discussed previously. We consider fixed payment scheme for the extra control cost, in which drivers would receive the revenue as long as recovery efforts have been devoted, whatever the holding time is. The reasons are twofold: (1) it is more convenient for practical implementation; (2) although the holding time is varied, it can be expected that the average holding time and thus the revenue for drivers in the same line would be roughly identical in a period.

Therefore, the labour cost paid to the drivers for making the recovery effort can be calculated as follows:

$$C_r = \rho\sum_m\sum_k \frac{P_{mk}}{h_k} \tag{21}$$

where $\frac{1}{h_k}$ represents the departure frequency on line k; P_{mk} is the occurrence probability of holding on line

k at transfer node m; $\frac{P_{mk}}{h_k}$ can be interpreted as the holding occurrence frequency.

The condition of holding for line k at transfer station m is that there is at least one late incoming bus and its predicted arrival time is within the SCM range H. The probability of such holding deployment is

$$P_{mk} = 1 - \prod_{i,i\neq k} P_i(d_{mi} < s_{mi} \bigcup d_{mi} > s_{mi}+H) = 1 - \prod_{i,i\neq k}(1-e^{-\lambda_{mi}s_{mi}}+e^{-\lambda_{mi}(s_{mi}+H)})$$

$$= 1 - \prod_{i,i\neq k}[1+e^{-\lambda_{mi}s_{mi}}(e^{-\lambda_{mi}H}-1)] \tag{22}$$

From Eq. (22), the holding probability P_{mk} is shown to be a monotonically decreasing function of slack time s_{mi} and thus average slack time $\bar{s}_{mi}$. If no slack time is set at the planning level, i.e., $s_{mi} = 0$, then the holding probability becomes $P_{mk} = 1 - e^{-H \sum_{i,i \neq k} \lambda_{mi}}$.

➤ Extra waiting time cost for downstream passengers C_d

Compared to DPC, once holding is implemented, DDC will increase waiting time of passengers further down the route. Under no schedule recovery, the additional waiting time for individual passenger would be equal to the holding time (Ting and Schonfeld, 2007). However, with the effect of schedule recovery, the additional waiting time caused by holding will be gradually diminished along the downstream route and eventually eliminated before the next transfer point.

Since the holding time may be different at each individual control deployment, it is natural to model the additional waiting time in an aggregated way during a time period as the expected additional waiting time. Here, we adopt the continuum approach to approximate such cost component to investigate the impact of holding on downstream waiting cost. As noted by Daganzo (2004), the main purpose of this type of approach is to obtain reasonable solutions with as little information as possible. By using continuum approach, we can easily investigate the impact of the number of downstream waiting passengers on the optimization performance.
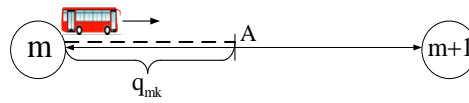

Fig.3 Virtual boarding demand along the link between consecutive nodes

An example is shown in Fig.3, a bus heading to transfer node $m+1$ is held at transfer node m. The resulting tardiness from the bus holding at m can be covered by the driver's schedule recovery effort by location A. For simplicity, we assume that the total number of impacted waiting passengers on specified link is q_{mk}, and they are distributed homogeneously along the line from the respective transfer node (m) to the downstream location A where the holding delay is recovered (shown as the dash line segment in Fig. 3). In a real bus network, q_{mk} can be estimated by the boarding demand of the intermediate bus stops that are affected by schedule recovery, which represents the adverse impact of DDC and is given exogenously but does not exist for DPC. Since we only model terminals and transfer stations in our bus network, this virtual demand between nodes is not explicitly represented by our modelled OD demand. The expression for the extra waiting time cost for downstream passengers is presented as follows (See detailed deviation in Appendix 5):

$$C_d = \frac{1}{2} \mu_w \sum_m \sum_k q_{mk} E_{mk}(H) \tag{23}$$

Eq. (23) is a closed-form approximation that is a representative of uniform distributed demand. The only information we need is the number of total impacted downstream passengers, and that is sufficient to capture the impact of holding policy.

Substituting Eqs. (21) and (23) into Eq. (19), we get the induced tardiness cost for DDC as:

$$C_s = \sum_m \sum_k [(\mu_w Q_{mk} + \frac{B}{h_k})(s_{mk} + E_{mk}(H)) + \rho \frac{P_{mk}}{h_k} + \frac{1}{2}\mu_w q_{mk} E_{mk}(H)] \tag{24}$$

3.6.2 Missed connection cost at transfer stations C_c

The formulations for missed connection cost with a given SCM follow those for DPC policy, and are also computed accounting for the following possible conditions:

a) The feeder vehicle on line k arrives late while the connecting vehicle on line l is not late

By a modification of Eq. (13) to consider a given SCM of length H, we have the following corresponding probability

$$\begin{aligned} p^a &= p(d_{mk} > s_{mk} + H \cap d_{ml} < s_{ml}) \\ &= p(d_{mk} > s_{mk} + H) p(d_{ml} < s_{ml}) \\ &= e^{-\lambda_{mk}(s_{mk}+H)}(1 - e^{-\lambda_{ml}s_{ml}}) \end{aligned} \tag{25}$$

b) Both vehicles are late, but the feeder vehicle in line k arrives after the predetermined maximum holding time H and the vehicle in line l not

The corresponding probability is:

$$\begin{aligned} p^b &= p(d_{mk} > s_{mk} + H \cap s_{ml} < d_{ml} < s_{ml} + H) \\ &= p(d_{mk} > s_{mk} + H) p(s_{ml} < d_{ml} < s_{ml} + H) \\ &= e^{-\lambda_{mk}(s_{mk}+H)-\lambda_{ml}s_{ml}}(1 - e^{-\lambda_{ml}H}) \end{aligned} \tag{26}$$

c) Both vehicles are late and beyond the predetermined maximum duration of time H, but the feeder vehicle in line k arrives after vehicle in line l, that is

$$\begin{aligned} p^c &= p(d_{mk} > s_{mk} + H \cap d_{ml} > s_{ml} + H \cap d_{mk} - s_{mk} > d_{ml} - s_{ml}) \\ &= p(d_{mk} > s_{mk} + H) p(d_{ml} > s_{ml} + H) p(d_{mk} - s_{mk} > d_{ml} - s_{ml}) \end{aligned} \tag{27}$$

where $p(d_{mk} - s_{mk} > d_{ml} - s_{ml})$ is given by Eq.(15).

Therefore, the total missed connection costs are also calculated with the following piecewise conditions:

$$\begin{aligned} C_c &= \mu_w \sum_m \sum_l \sum_k q_l^k h_l (p^a + p^b + p^c) \\ &= \mu_w \sum_m \sum_l \sum_k q_l^k h_l \cdot \begin{cases} [De^{-\lambda_{mk}(s_{mk}-s_{ml})}(1 - e^{-\lambda_{mk}h_k}) + E] & \text{for } s_{mk} \geq s_{ml} \\ [D(1 - e^{-\lambda_{ml}(s_{ml}-s_{mk})}) + E] & \text{for } s_{mk} < s_{ml} \end{cases} \end{aligned} \tag{28}$$

where $D = \dfrac{\lambda_{ml}}{\lambda_{mk} + \lambda_{ml}} e^{-\lambda_{mk}(s_{mk}+H)-\lambda_{ml}(s_{ml}+H)}$, and $E = e^{-\lambda_{mk}(s_{mk}+H)}(1 - e^{-\lambda_{ml}(s_{ml}+H)})$

3.6.3 Delayed connection cost at transfer stations C_f

Under DDC, the formulation of delayed connection cost is the same as the DPC case represented by Eq. (18). This is because the formulas are not related to SCM in all possible conditions as listed in those of under DPC.

4. RCSS-STT optimization formulations

4.1 A stochastic optimization model with given SCM

The main purpose in this study is to investigate the sensitivity of system performance under various SCM. Our proposed robust coordination schedule scheme with stochastic travel time (RCSS-STT) optimization problem is to find the best decision variables h_k (headways) and s_{mk} (slack times) that balance the trade-offs between the operating cost and user cost with a given SCM. For practical application convenience, we optimize headways as integer values for all operation modes. Although setting SCM to be multiples of 30 seconds (e.g., 0.5, 1, 1.5, 2 minute) may be favourable by operators since it is easier to handle, we test a series of discrete values of SCM with interval of 0.1 min to analyse the sensitivity. The number of variables varies with the size and topology of the bus network, which we will see in the numerical examples. As a result, the RCSS-STT is formulated as mixed-integer nonlinear programming (MINLP) models:

$$\min C^{DPC(DDC)}(h_k, s_{mk}) \tag{29}$$

subject to

$$h_k = h, \ \forall k \tag{30}$$

$$h_k^{min} \le h_k \le h_k^{max}, \ \forall k \tag{31}$$

$$h_k \in Z^+, \forall k \tag{32}$$

$$s_{mk} \ge 0, \ \forall k, m \tag{33}$$

$$H \in \{H_1, H_2, \ldots, H_i\}, \ i = 1, 2, \ldots, I \tag{34}$$

The objection function (29) minimizes the system total cost. Eq. (30) ensures that the headways of different lines are the same, where h is the common headway. Eq. (31) ensures that the headway should exceed a minimum acceptable headway and not exceed a maximum allowable headway. Eq. (32) ensures that headways are integer values. Eq. (33) ensures that the slack times should be positive values. Finally,

Eq. (34) assumes that the SCM takes one of the given discrete values. The model for DDC is equivalent to that for DPC operation when the SCM is set to be zero.

4.2 An enhanced optimization model for determining SCM

The above model can be enhanced by simultaneously optimizing the value of SCM, headways and slack times. To this end, H is treated as a decision variable in the objective function. The enhanced optimization model is presented as follows:

$$\min C^{DDC}(h_k, s_{mk}, H) \tag{35}$$

subject to conditions (30) to (33), plus (36) below

$$0 \leq H \leq H_{max} \tag{36}$$

Constraint (36) states that the SCM must be non-negative and may not exceed a maximum H_{max}. This upper bound H_{max} can be fine-tuned or estimated by rule-of-thumb from practical traffic safety requirement, as discussed in the problem description section (Section 3.1).

4.3 Solution method

The branch-and-bound algorithm is one of the most successful exact approaches to solve combinatorial optimization problems, both for linear and nonlinear programming models. It has been used to solve the dial-a-ride problem (Liu et al., 2015; Braekers et al., 2014), train timetabling problem (Zhou et al., 2007), and ship routing problem (Meng et al., 2015). For small- and medium-sized problem, we can use branch-and-bound method to obtain feasible solutions. For large-scale problem, however, branch-and-bound methods can be extended with a number of strategies added to improve their performance (e.g., Meng et al., 2015; Liu et al., 2015; Braekers et al., 2014). The procedures of branch-and-bound algorithm are briefly described below.

A problem with an integer variable is first being solved with the integer variable considered continuous (the first sub-problem). After this, the program generates sub-problems where the domain of the variable (still continuous) is being restricted. This is called branching. Then it solves these sub-problems. This process continues until the variable is fixed to an integer value. The advantage of this approach (when compared with explicit enumeration) lies in the fact that not all the sub-problems have to be solved (Zhou et al., 2007).

In this paper, we use BNB20[2] function to solve MINLP problem. The boundary conditions of optimized headways are chosen from [1, 60 min].

4.4 An additional performance indicator: Transfer failure rate

---

[2] BNB20 is a Matlab-based package developed by Koert Kuipers in University of Groningen
 (Source:http://uk.mathworks.com/matlabcentral/fileexchange/95-bnb)

As with costs, transfer failure rate is a key evaluation indicator on transfer reliability and service quality. Similar to Wu et al. (2015), we define transfer failure as the case where (at least) one passenger misses the planned connecting bus even if the passenger can transfer to the following bus. The generalized analytical formulation for transfer failure rate, represented as the missed connection probability, can be extracted from the missed connection cost component in Eq. (28), which is summarized as follows:

$$\text{TFR}_{mkl} = \begin{cases} [De^{-\lambda_{mk}(s_{mk}-s_{ml})}(1-e^{-\lambda_{mk}h_k})+E] & \text{for } s_{mk} \geq s_{ml} \\ [D(1-e^{-\lambda_{ml}(s_{ml}-s_{mk})})+E] & \text{for } s_{mk} < s_{ml} \end{cases} \tag{37}$$

where TFR_{mkl} denotes the transfer failure rate at transfer node m from line k to line l, $H=0$ and $H>0$ present DPC and DDC scenarios, respectively.

A network-wide average transfer failure rate can be obtained by the mean of transfer failure rates of transfer node m and line k in the bus network, i.e.,

$$\text{Avg TFR} = \frac{1}{\text{Num}} \sum_m \sum_k \sum_l \text{TFR}_{mkl} \tag{38}$$

where Num is the number of transfer route combinations.

5. Numerical Examples

In this section, we compare the performances of the three operation modes (i.e., uncoordinated operation, DPC and DDC) through numerical analysis. We conduct experiments on two hypothetical bus networks, as shown in Fig.4. The topology of the two networks is described as follows:

Network 1: A bus network with 2 transfer nodes and 3 bus lines.

Network 2: A bus network with 3 transfer nodes and 3 bus lines.

The two networks represent two distinct and classical bus network structures: a) a trunk-and-feeder (network 1 of Fig. 4a); and b) a looped structure (network 2 in Fig. 4b). In the hierarchical trunk-and-feeder network 1, the main 'trunk' route runs between nodes 3 and 4 and links to major destinations, while the two branches from nodes 5 and 6 act as feeder services to collect and re-route passengers in the network. Passengers may need to make more than one transfer to get to their destination, for example, between node 5 and node 6. There is no junction between the feeder lines (line 2 and line 3), thus the transfer volumes are zero, i.e. $q_2^3 = q_3^2 = 0$. On the other hand, in the triangular loop structure of network 2, all lines interact with each other providing direct access for passengers to their destinations with no more than one transfer. The two spatial designs represent two alternative generic structures with different connectivity or mobility featuring indirect transfer (network 1) and direct transfer (network 2) between lines respectively.
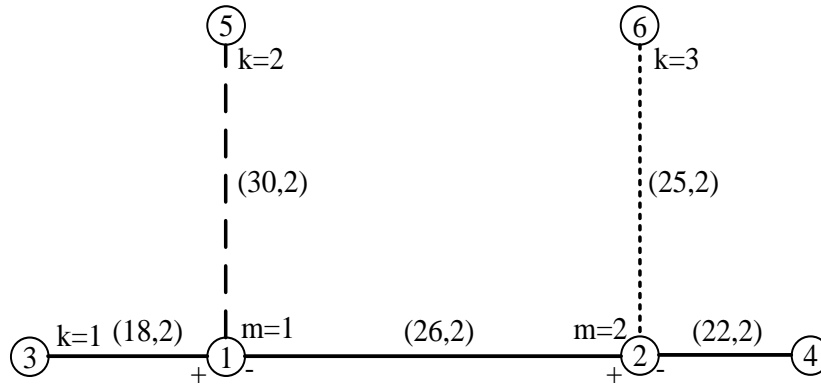
(a) Bus network 1
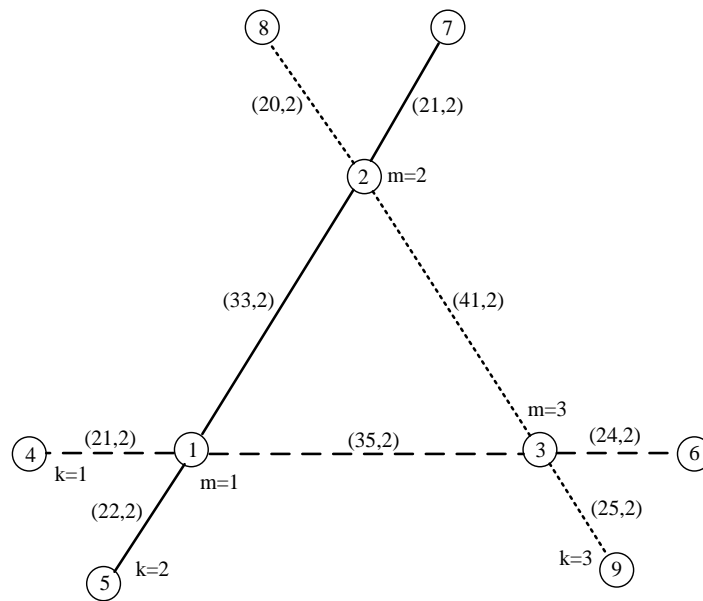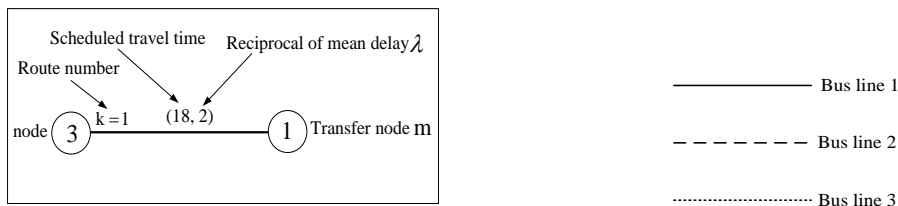
(b) Bus network 2

Fig.4 The test bus networks. The symbols are illustrated in the inserted box.

The OD demand matrix for network 1 is symmetrical (Table 4) and that for network 2 is asymmetrical (Table 5). The bus networks and OD demands are both adapted from Ting and Schonfeld (2005), while the route travel time and delay distribution parameters are re-defined here. The passenger route choice is assumed to be based on the shortest path. With this assumption, the boarding demand (Q_k), the through demand (Q_{mk}), and the transfer demand (q_k^l, $q_{l,}^k$ and R_k) can be uniquely derived. Note that the bidirectional headways of a line are identical in order to prevent internal disruption as mentioned previously.

The other default settings for both networks are:

(1) The values for the model parameters are fixed at: $B = 0.667\$/veh/min$, $\mu_w = 0.2\$/pax/min$, $\rho = \$1$, $t_w = 4$ min, $\alpha = 0.33$;

(2) The base delay parameter $\lambda_{mk} = 2 \, min^{-1}$ is fixed for all transfer stations and on all lines;

(3) The default SCM value is H=1min.

(4) The default demand for the two networks are as those listed in Tables 4 and 5; we introduce a demand multiplier n in the sensitivity analysis on the demand, where the base OD demand has a ratio $n = 1$.

(5) The downstream passengers who are affected by the holding are assumed to be in proportion to the demand multiplier as $q_{mk} = 3n \, pax/min$.

These default settings remain the same in all sensitivity tests, except where they are the subject of a test. For the uncoordinated scenario, the optimal headways as calculated according to Eq. (11) and taken the resulted integer values.

Table 4. A symmetric O/D demand matrix for Network 1 (pax/min)

Origin	Destination					
	1	2	3	4	5	6
1	0.0	0.8	0.8	0.6	0.2	0.8
2	0.8	0.0	0.6	0.7	0.4	0.5
3	0.8	0.6	0.0	0.4	0.6	0.0
4	0.6	0.7	0.4	0.0	0.0	0.8
5	0.2	0.4	0.6	0.0	0.0	0.0
6	0.8	0.5	0.0	0.8	0.0	0.0

Table 5. An asymmetric O/D demand matrix for Network 2 (pax/min)

Origin	Destination								
	1	2	3	4	5	6	7	8	9
1	0.00	0.50	0.30	0.30	0.50	0.20	0.40	0.30	0.20
2	0.50	0.00	0.40	0.40	0.00	0.30	0.30	0.30	0.00
3	0.30	0.25	0.00	0.20	0.20	0.30	0.15	0.15	0.25
4	0.30	0.20	0.20	0.00	0.20	0.10	0.10	0.00	0.10
5	0.50	0.40	0.40	0.40	0.00	0.30	0.30	0.30	0.00
6	0.20	0.20	0.30	0.10	0.10	0.00	0.00	0.10	0.20
7	0.40	0.50	0.30	0.30	0.30	0.00	0.00	0.40	0.30
8	0.15	0.25	0.15	0.00	0.05	0.05	0.15	0.00	0.05
9	0.15	0.15	0.25	0.05	0.00	0.15	0.05	0.05	0.00

The base RCSS-STT optimisation problem (29) are solved for the network 2 at the reduced demand (at a fraction n=0.1, i.e., 1/10 of the base demand), and the results are presented by Table 6. The results include optimized headways and slack times, and corresponding cost components and thus the total cost. One can see that the total transfer waiting time cost ($C_c + C_f$) of coordinated operation are significantly lower than that of uncoordinated operation (C_t), suggesting that the schedule coordination are effective in eliminating the user transfer costs. In this regard, DDC performs better than DPC, indicating the potential (further) improvement made by the SCM scheme. The DPC and DDC operations results in longer common headways than the averaged headway in the uncoordinated operation, leading to higher waiting time costs (C_w) under the coordinated operations than those without coordination.

Table 6. Optimized results of network 2 for reduced demand (n=0.1) scenario

Variables (min)		Uncoordinated	DPC	DDC (H=1)	Cost ($/min)	Uncoordinated	DPC	DDC (H=1)
h_k (note 1)		(31, 23, 49)	59	58	C	25.97	23.61	23.42
	s_{11}	—	(1.83, 1.66)	(1.30, 1.42)	C_o	9.09	5.57	5.47
	s_{12}	—	(1.93, 1.82)	(1.22, 1.24)	C_w	14.66	17.21	17.31
s_{mk} (note 2)	s_{21}	—	(1.76, 1.89)	(0.91, 1.05)	C_t	2.22	—	—
	s_{23}	—	(1.76, 1.50)	(1.24, 0.93)	C_s	—	0.47	0.43
	s_{32}	—	(1.56, 1.81)	(0.64, 0.96)	C_c	—	0.20	0.12
	s_{33}	—	(1.44, 1.75)	(0.55, 0.84)	C_f	—	0.16	0.09
TFR (%)		—	6.46	3.92		—		

Note 1: For uncoordinated operation, the three values in () are the optimal headways for the three lines of the network.

Note 2: The two values in () represent the slack times for the two different directions of the line.

The optimization model of (35) is constraint by an upper bound H_{max}. We examine the effectiveness of this model for different values of H_{max}. Interestingly, Fig. 5 shows that the optimal H is zero at low H_{max} bound, but jumps to the upper bound when the H_{max} values are high. The boundary where the jump happens varies with network and demand level. This suggests that for the optimal DDC system of (35) to be beneficial, it requires a certain minimum level of SCM; and that it is always better to set the maximum allowed SCM.
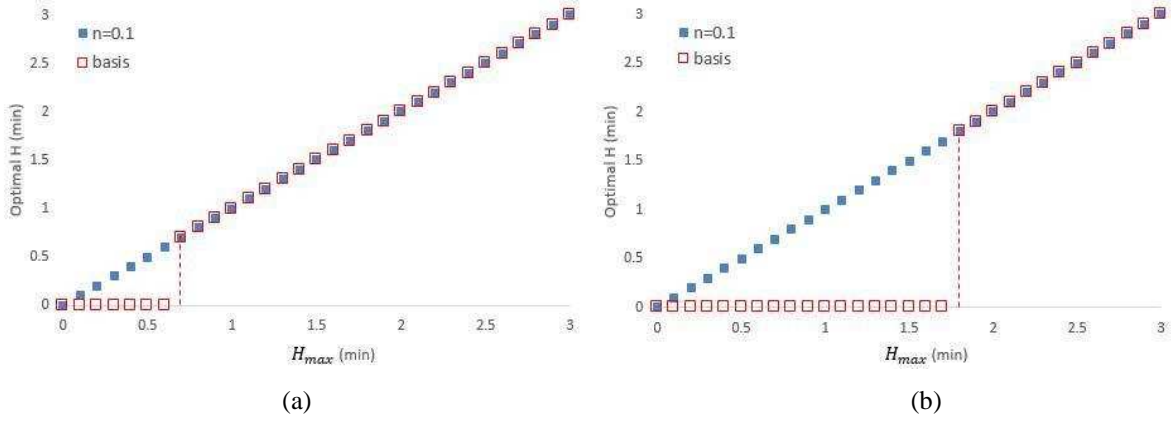
(a)   (b)

Fig.5 Optimal value of SCM with different demand and H_{max} for: (a) network 1; (b) network 2

In the following, we conduct sensitivity analysis to verify the effectiveness of the optimal design resulted from the proposed RCSS-STT method of (29) from the following aspects:

(a) Influence of SCM (the H value)
(b) Influence of total demand (n);
(c) Influence of arrival mean delay ($1/\lambda$).

The detailed output of the analytical model includes headways and slack times allocated to each transfer node. From this we can derive performance measures such as user cost and operating cost, and the average slack time that represents the schedule efficiency.

5.1 Influence of SCM

We examine the impact of different level of SCM on the resulted optimal slack time. We apply the RCSS-STT method to both networks at the base demand (as defined in Tables 4 and 5) and a reduced demand (n=0.1). Several different SCM levels (the H values) are tested, and the boundary is set as $H_{max} = 3$ min.

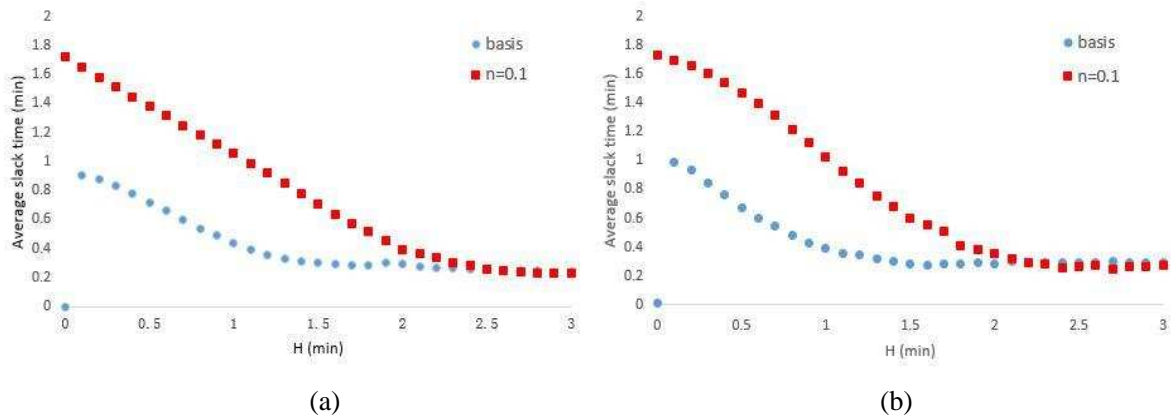(a)   (b)

Fig.6 Optimal average slack time with varying levels of SCM for: (a) network 1; (b) network 2.

Fig.6 shows the resulted average optimal slack times as the H value increases. We summarise the key observations from the results and discuss their implications below:

26

(a) Under the same H value, the average slack time for the reduced demand is generally larger than those for the base demand scenario. This is as expected, since at low demand, the service frequency is low and thus longer slack time is required to mitigate the travel time randomness and thus reduce transfer costs.

(b) With DPC operation (H=0), the optimal slack times at the base demand are zero for both networks. The result implies that it is uneconomical to impose slack time under DPC at the high demand levels. At low demand (n=0.1 case), however, a significant level of slack time is required.

(c) With DDC operations (H>0), there is a general trend in that the slack times decrease with increasing H values. This represents a trade-off between slack time (added at the planning stage) and a SCM (allowed at the operation stage). The larger H value, the smaller slack time is required.

(d) The slack times of different demand levels tend to converge as the value of SCM increases. When H is greater than 2 min, the slack times do not change much. The reasons are twofold. First, the delay arrival is exponentially distributed, underlying that the lower probability of vehicle arrival is desired for later time, hence marginal increase of SCM contributes less to compensate the cost reduction with smaller slack time. Second, the H value in DDC is only a maximum holding time and buses can depart earlier than H.

The above results, especially (b) – (d), show the complex relationships between slack time planning and operational SCM, and the inherent trade-off between different cost components in the system. For example, when SCM is small, the decrease in transfer waiting cost is not sufficient to overweight the additional economic incentives C_r (Eq. (21)) with a fixed payment scheme. Eq. (22) shows that the holding occurrence probability P_{mk} is a monotonically decreasing function of average slack time $\bar{s}_{mi}$. Accordingly, the greatest benefits of optimization is achieved when, due to the increased slack times, reduced holding probability P_{mk} and consequently the reduced additional economic incentives C_r ensues.

Following from the above discussion, we examine the effect of introducing SCM on total system cost and on individual cost components. Fig.7 shows the effect of SCM on the total costs C. The dashed lines point to a boundary in the H values below which the DDC costs are higher than that of DPC, while above this H threshold, it is cheaper to operation DDC. We postulate that with small SCM, the reduction in transfer waiting time cost is insufficient to compensate the increment of induced tardiness cost (Eq. (24)). As H increases, transfer costs reduces and eventually the benefits (of reduced transfer costs) outweigh the costs. For network 1, this threshold is at H=0.7min for the base demand levels, whilst for network 2, the value is 1.8min for the base demand.

We note that the above thresholds SCM correspond to the thresholds for the non-zero optimal H values shown in Fig. 5. This again reinforces the message that for DDC operations to be beneficial over DPC, a minimum level of SCM is required. The above result implies that, in practice, transit agencies can deal with random travel disruptions while maintaining lower cost by incorporating SCM without requiring substantial slack time (and therefore more efficient).
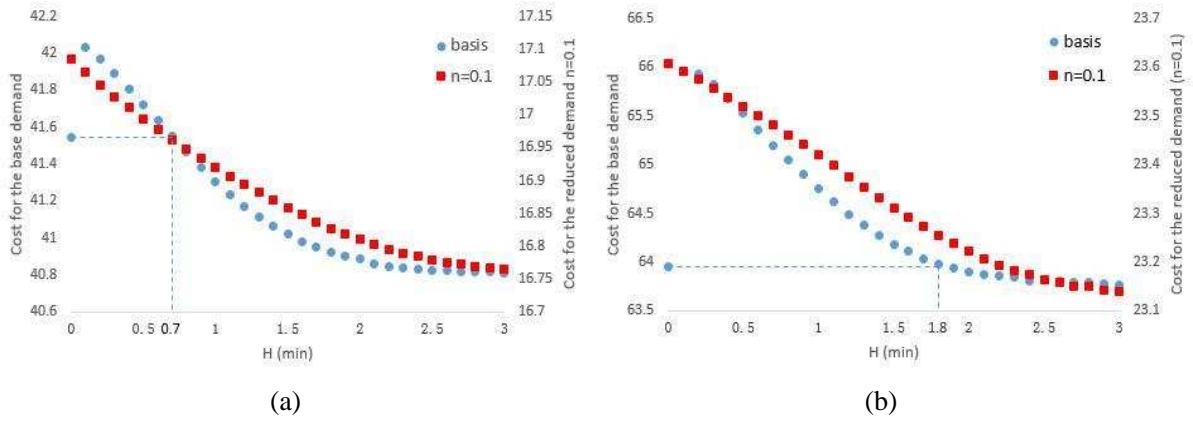
Fig.7 Optimal total cost C ($/min) vs different SCM levels: (a) network 1; (b) network 2.

It is worth noting that the total cost is composed of costs on the operation side (i.e., the operating cost C_o and tardiness cost C_s when introducing SCM), and costs to the passengers (including the transfer wait time costs C_c and C_f for common headway scenarios). We examine below the impact of SCM from the point of view of the transfer passengers.
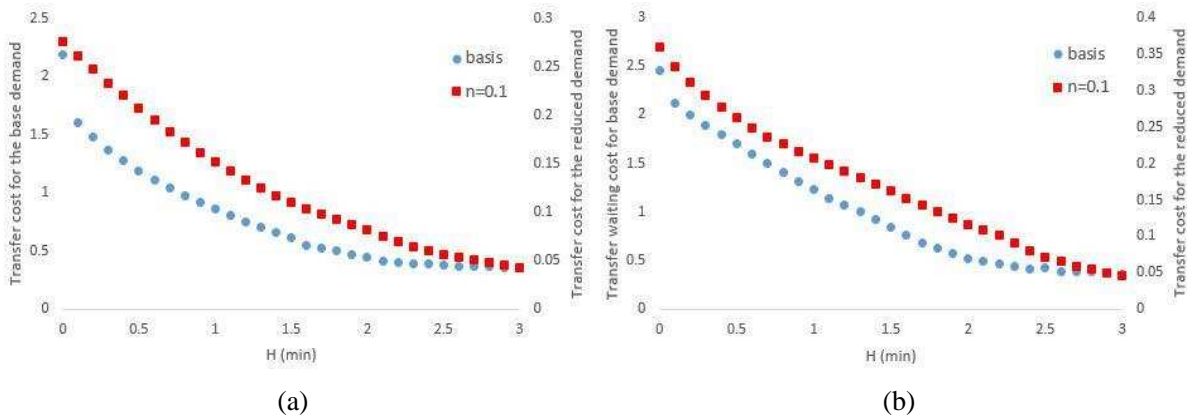

Fig.8 Transfer waiting time costs $C_F = C_c + C_f$ ($/min) of RCSS-STT performance for various SCM for both networks: (a) network 1; (b) network 2.

Fig.8 presents the results of one of the key performance indicators of the RCSC-STT results: the combined transfer waiting time costs $C_F = C_c + C_f$. It is clear that C_F decrease consistently as the SCM increases. This is as expected since such semi-flexible holding control strategy contributes to mitigating the travel time randomness and thus significantly reduce the possibility of miss connection.

Compared to the results on total costs (Fig.7), one can extrapolate that although adding slack times and SCM both reduce the transfer waiting time cost, the former contributes more to the adverse impact on some cost components (e.g. induced tardiness cost) at smaller H values. As shorter slack times are required when SCM gets larger (Fig.6), such negative effect will finally diminish as the slack time decreases and thus the total cost is reduced.
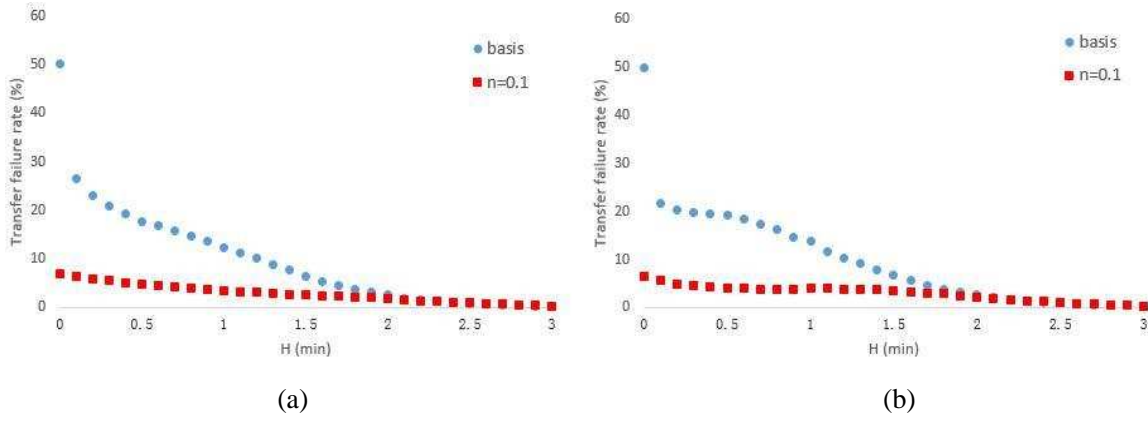
| (a) | (b) |

Fig.9 Average transfer failure rate vs different SCM: (a) network 1; (b) network 2

Fig.9 shows the effect of DPC and DDC operations on transfer failure rates. It can be seen that the transfer failure rate decreases with increasing H value, demonstrating again the benefit of DDC operations with large H range over the DPC operation, from the users' perspective. The significant reduction in transfer failure rate at the base demand in network 2 (Fig. 9b) between H=0 and H>0 is linked closely to the optimal slack times for the different scenarios (Fig. 6). As shown in Fig.6, it is uneconomical to insert slack time for DPC (H=0) for the base demand cases, while adding slack times will be needed with DDC (H>0), transfer failure rate will decrease considerably with imposed slack time.

Furthermore, the average transfer failure rates for the reduced demand scenario (n=0.1) are generally lower than those of for the base demand scenario, especially for DPC (H=0). Hence, in terms of transfer reliability, the potential benefit of schedule coordination is higher when the demand is relatively low for both DPC and DDC. However, as shown in Fig. 9, the gap between them is gradually narrowing as H increases. When H reaches a critical value (e.g., 1.8 min for both networks), the differences are negligible. This implies that the relative advantage of transfer failure rate based on low demand becomes less with increasing H under DDC policy, which also indicates that DDC is more robust to demand variation than DPC.

5.2 Influence of total demand

The results in Section 5.1 imply that the operation performances of the schedule coordination depend on the demand level, and that small SCM is undesirable in reducing total cost. We analyse here the effect of passenger demand on DPC operation vs that on DDC operation with a reasonable level of SCM, at H=1 min.

Fig.10 presents the optimal average slack times and headways with varying demand levels. It is as expected that the optimal common headways decrease with increasing demand. This is similar to that of uncoordinated operations (Eq. (11)). It is interesting to see that in all cases with demand ratio greater than 0.2, the optimized headways for both DPC and DDC reduce to 12 min or less. This suggests that passengers will switch from "schedule-dependent" to "schedule-blind" when the demand reaches a certain level.
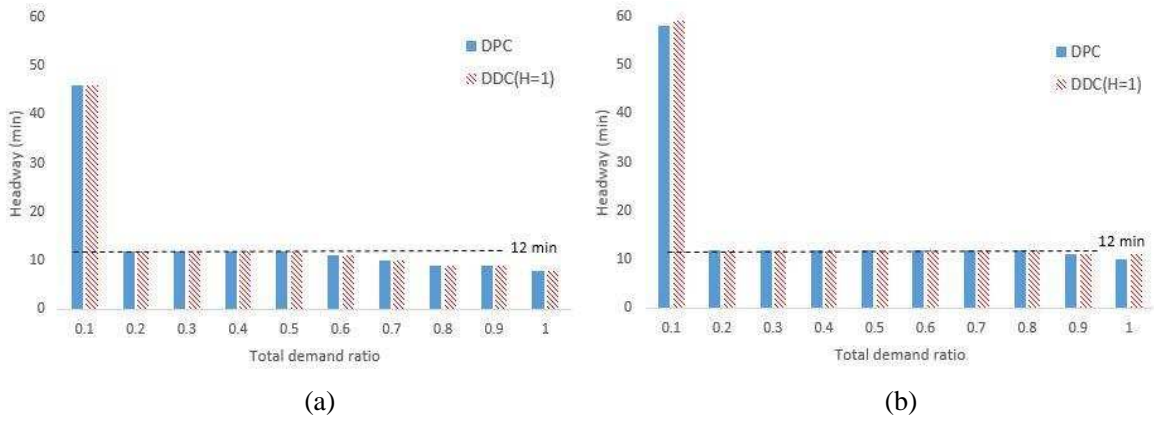
Fig.10 Optimal common headways vs different demand ratio for: (a) network 1; (b) network 2

To illustrate the trade-off of schedule coordination between the transfer and the boarding passengers, Fig.11 shows the waiting time cost of boarding passengers at various demand ratios for coordinated and uncoordinated operation. Due to the possible distinctive passenger behaviour under different demand, the effect of demand on waiting time costs for different operation modes are more complex. It can be seen that the waiting time costs of uncoordinated operation can be either larger or less than those of coordinated operation depending on the demand level: it is higher when the demand ratio is between 0.2 and 0.7 for both networks. This is because the passenger behaviour of uncoordinated and coordinated operations are different in this range of demand: schedule-independent for uncoordinated operation while schedule-dependent for coordinated operation. In other words, the passenger behaviour of the three operation modes are identical outside this range. This gives us an insight that for the same passengers' arrival pattern, the schedule coordination (with DPC and DDC) has a negative impact on boarding passengers, at the expense of reduced transfer costs for transfer passengers. The reason is that the uncoordinated operation optimize the headways independently, however, DPC and DDC focus only on the transfer nodes, and headways have to be synchronised to reduce transfer waiting time cost, at the expense of boarding passengers.
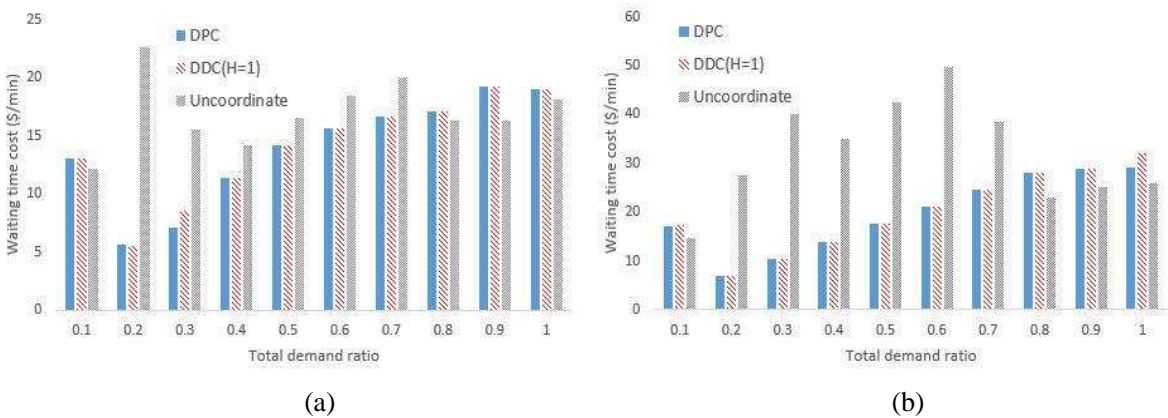

Fig.11 Waiting-time cost C_w vs different demand ratio: (a) network 1; (b) network 2

30

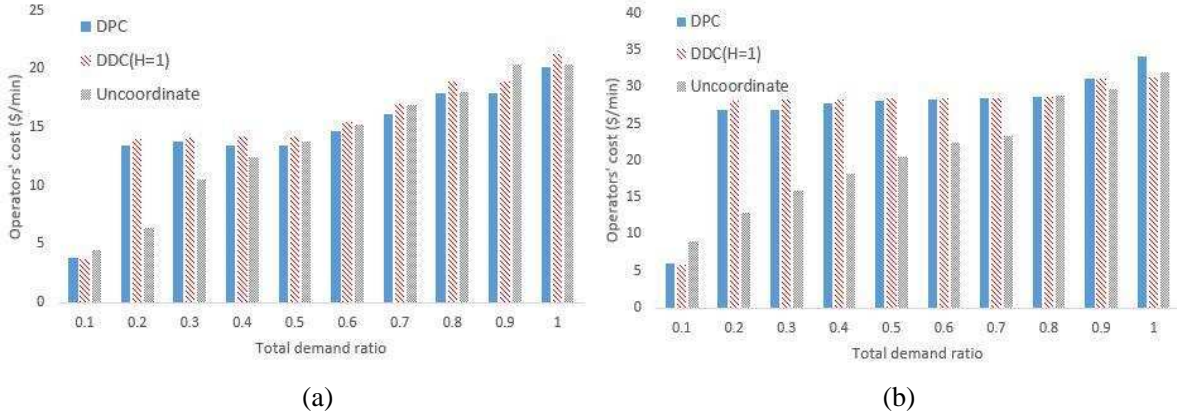(a)   (b)

Fig.12 Operators' cost $C_o (+C_s)$ at different demand levels for: (a) network 1; (b) network 2

We examine further the cost implications of schedule coordination from the point of view of the operators. Fig. 12 presents the operating cost C_o and induced tardiness cost C_s for various demand ratios. Since induced tardiness cost is only included in DPC and DDC and represents the extra cost for schedule coordination, we compare the extra total cost (i.e., $C_o + C_s$) of coordinated operation with the operating cost C_o of uncoordinated operation. While the total operators' cost for coordinated operations vs those for uncoordinated operations vary depending on the demand level, a common result for both networks is that the former is lower than that of latter one when the demand is relatively low, for instance when the demand ratio is 0.1. This suggests that coordinated operations may work better in low demand scenarios in that no increase of operators' cost is required.

5.3 Influence of arrival mean delay

In this section, we investigate the impact of the mean delay $(1/\lambda)$ on the slacks in the schedule. In practice, there are two slack times, one for the up and one for the down direction of the line at the transfer node. For simplicity and clarity, we take the average of the two, i.e., $s_{11} = (s_{11}^+ + s_{11}^-)/2$, since the delay distribution parameter are assumed to be the same for both the up and down directions of a line. The sensitivity results are presented in Fig. 13 for the reduced demand scenario (i.e., n=0.1) for both networks, for a range of delay values $1/\lambda = 0.1, 0.2, \ldots$ min. Notice that a $1/\lambda = 2$ min is used in all the other sensitivity analysis.

Fig.13 shows the effect of mean delay on the averaged (over the up- and down-directions) optimal slack time. One can observe that the slack time required initially increases with mean delay, but decreases when the mean delay is over certain threshold. With coordinated operations, over this delay threshold, the slack times drop to zero, which means that it is no longer beneficial to impose any slacks when the delay is beyond a critical value. Given our assumption that the delay arrival time is exponentially distributed, its variance $1/\lambda^2$ also increases with the mean delay $1/\lambda$. Therefore, the above results suggest that slack time is undesirable when arrival time uncertainties are high.

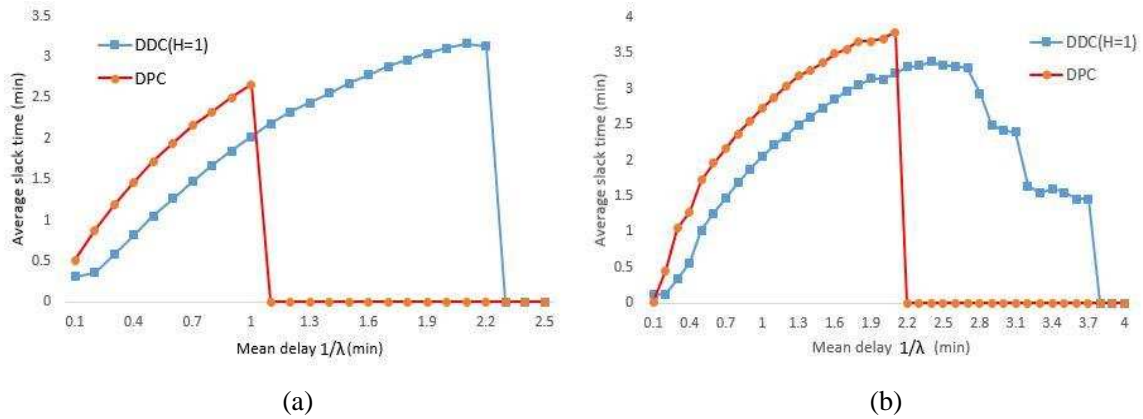(a)   (b)

Fig.13 The averaged optimal slack time for various mean delays: (a) network 1; (b) network 2

Note also from Fig. 13 that the critical delay values for DDC operations are much bigger than those for DPC operations. This suggests that DDC is more robust to delay uncertainty. Evidently, this result with DDC operations is an improvement over that under no holding policy in Ting and Schonfeld (2005), who concluded in their study that: "slack time is most feasible and desirable when arrival uncertainties are low". Therefore, our analysis suggests that relatively higher uncertain arrivals can be allowed for by introducing the SCM, which provides a possible way for delay management in public transport.

6. Concluding Remarks

This paper presents a novel schedule coordination model that couples planned slack time with real-time holding strategies, in which the variation in travel time due to stochastic disturbance is sufficiently compensated at both the planning and the operation stages. The findings are specific to two hypothetical transit networks that represent two alternative classical structures with different connectivity. The presented features with SCM in both networks have many commonalities with respect to the principal performance measures such as total cost, user cost, transfer reliability, and robustness to demand and delay, as well as the trends of slack times, verifying the effectiveness of SCM in schedule coordination. A summary of key findings is listed below:

1. Compared to DPC which relies only on adding a slack time in the planned timetable to account for historic delay, DDC can cope with random travel disruptions better and maintain lower cost from both user (see the results on passenger transfer waiting cost in Fig.8) and overall system costs (see the results on total cost in Fig.7) perspectives by suitably incorporating SCM. This is achieved without requiring substantial slack time, and thus obtaining a more cost-effective and efficient timetable.

2. From the user perspective, with DDC, both the passengers' transfer waiting time cost (Fig. 8) and transfer failure rate (Fig. 9) decreases considerably with the increase of SCM compared to DPC (Fig.8).

3. From the perspective of transfer failure rate and effective (non-zero) slack time, we demonstrate that the DDC is more robust to demand variation (Fig.9) and delay uncertainty (Fig.13) compared to DPC.

4. The passenger behaviour for both coordinated and uncoordinated operations may switch between 'schedule-independent' and 'schedule-dependent' with varied demand levels, while low demand scenarios are preferable for coordinated operation in that no increase of operators' cost is required (Fig.12).

This paper proposed an analytical framework to explore the potential of SCM setting in schedule coordination. Further research is required to explore more detailed design issues, such as jointly optimizing a set of non-common SCM for different transfer nodes in a more complex large-scale hybrid network, multiple vehicle types for different routes, and considering fleet size limitation. A limitation in this current paper is that the synchronization is only scheduled-based, i.e., with the same scheduled departure time (SDT) in a transfer node to ensure the maximum encounter probability. This is most suitable for the typical network whereby the sublines are not overlapped as illustrated in our example. However, this assumption might lead to bus bunching for the special networks in which many sublines share a common route segment as discussed in Ibarra-Rojas and Rios-Solis (2012). Investigating a new effective control solution to account for this situation maybe another future works. Furthermore, our current solution is based on the assumption that the forecasted delayed arrival time that beyond slacks is deterministic. A more generalized scenario considering uncertain forecasted delayed arrival should be explored in further work.

Acknowledgement

Appendix 1. Proof of pdf of the difference of two independent exponential random variables

Assuming that x_1, x_2 are independent continuous variables, their density probability functions are $f_1(x_1)$ and $f_2(x_2)$, respectively, according to the convolution theorem, then we have

$$f_{x_1-x_2}(x) = \int_{-\infty}^{+\infty} f_1(x+t) \cdot f_2(t)dt$$

Since $f(d_k) = \lambda_k e^{-\lambda_k d_k} (d_k \geq 0)$, $f(d_l) = \lambda_l e^{-\lambda_l d_l} (d_l \geq 0)$

Let $x = d_k - d_l$

Then we have

$$f(x) = \int_0^{+\infty} \lambda_k e^{-\lambda_k(x+t)} \cdot \lambda_l e^{-\lambda_l t}dt = \lambda_k \lambda_l e^{-\lambda_k x} \int_0^{+\infty} e^{-(\lambda_k+\lambda_l)t}dt$$

$$= \lambda_k \lambda_l e^{-\lambda_k x} \cdot \frac{1}{-(\lambda_k+\lambda_l)} \cdot e^{-(\lambda_k+\lambda_l)t} \Big|_0^{+\infty} = \frac{\lambda_k \lambda_l}{\lambda_k+\lambda_l} e^{-\lambda_k x} \qquad (x > 0)$$

Appendix 2. Derivation of equation (15)

I. when $s_{mk} \geq s_{ml}$

Let $x = d_{mk} - d_{ml}$

Since $f(d_{mk}) = \lambda_{mk} e^{-\lambda_{mk} d_{mk}} (d_{mk} \geq 0)$, $f(d_{ml}) = \lambda_{ml} e^{-\lambda_{ml} d_{ml}} (d_{ml} \geq 0)$, according to Appendix 1, we have

$$f(x) = \frac{\lambda_{mk} \lambda_{ml}}{\lambda_{mk} + \lambda_{ml}} e^{-\lambda_{mk} x} (x > 0)$$

Assuming that buses will seldom arrive late by more than one headway, then

$$
\begin{aligned}
& p(d_{mk} - s_{mk} > d_{ml} - s_{ml}) \\
&= p(0 < (d_{mk} - s_{mk}) - (d_{ml} - s_{ml}) < h_k) \\[6pt]
&= p(s_{mk} - s_{ml} < d_{mk} - d_{ml} < h_k + s_{mk} - s_{ml}) \\
&= p(s_{mk} - s_{ml} < x < h_k + s_{mk} - s_{ml}) = \int_{s_{mk} - s_{ml}}^{h_k + s_{mk} - s_{ml}} f(x) dx \\
&= \int_{s_{mk} - s_{ml}}^{h_k + s_{mk} - s_{ml}} \frac{\lambda_{mk} \lambda_{ml}}{\lambda_{mk} + \lambda_{ml}} e^{-\lambda_{mk} x} dx = \frac{\lambda_{ml}}{\lambda_{mk} + \lambda_{ml}} e^{-\lambda_{mk}(s_{mk} - s_{ml})} (1 - e^{-\lambda_{mk} h_k})
\end{aligned}
$$

II. when $s_{mk} < s_{ml}$

Let $x = d_{ml} - d_{mk}$, then

$$
\begin{aligned}
& p(d_{mk} - s_{mk} > d_{ml} - s_{ml}) \\
&= p((d_{ml} - s_{ml}) - (d_{mk} - s_{mk}) < 0) \\
&= p(0 < d_{ml} - d_{mk} < s_{ml} - s_{mk}) = \int_0^{s_{ml} - s_{mk}} f(x) dx \\
&= \int_0^{s_{ml} - s_{mk}} \frac{\lambda_{mk} \lambda_{ml}}{\lambda_{mk} + \lambda_{ml}} e^{-\lambda_1 x} dx = \frac{\lambda_{mk}}{\lambda_{mk} + \lambda_{ml}} (1 - e^{-\lambda_{ml}(s_{ml} - s_{mk})})
\end{aligned}
$$

Consequently, the probability $p(d_{mk} - s_{mk} > d_{ml} - s_{ml})$ is calculated with the following piecewise conditions:

$$
p(d_{mk} - s_{mk} > d_{ml} - s_{ml}) =
\begin{cases}
p(0 < (d_{mk} - s_{mk}) - (d_{ml} - s_{ml}) < h_k), & \text{for } s_{mk} \geq s_{ml} \\
p((d_{ml} - s_{ml}) - (d_{mk} - s_{mk}) < 0), & \text{for } s_{mk} < s_{ml}
\end{cases}
$$

$$
=
\begin{cases}
\dfrac{\lambda_{ml}}{\lambda_{mk} + \lambda_{ml}} e^{-\lambda_{mk}(s_{mk} - s_{ml})} (1 - e^{-\lambda_{mk} h_k}), & \text{for } s_{mk} \geq s_{ml} \\[12pt]
\dfrac{\lambda_{mk}}{\lambda_{mk} + \lambda_{ml}} (1 - e^{-\lambda_{ml}(s_{ml} - s_{mk})}), & \text{for } s_{mk} < s_{ml}
\end{cases}
$$

Note that the objective of using piecewise function is to make $x > 0$ holds (See the pdf in Appendix 1). This completes the derivation of Eq. (15).

Appendix 3 Derivation of Eq. (18)

a) Both vehicles are not late, no matter whether the connecting bus on line l arrives after or before the feeder bus on line k, the waiting time is $w = s_{mk} - d_{mk}$. Integration allows us to calculate the cost for waiting:

$$C_f^1 = \mu_w \sum_m \sum_l \sum_k q_l^k \int_0^{s_{ml}} \int_0^{s_{mk}} (s_{mk} - d_{mk}) f(d_{mk}) dd_{mk} f(d_{ml}) dd_{ml}$$

$$= \mu_w \sum_m \sum_l \sum_k q_l^k [s_{mk} + \frac{1}{\lambda_{mk}} (e^{-\lambda_{mk} s_{mk}} - 1)](1 - e^{-\lambda_{ml} s_{ml}})$$

b) The feeder vehicle on line k arrives early while the connecting vehicle on line l is late. The waiting time is $w = d_{ml} - s_{ml} + s_{mk} - d_{mk}$. Integration allows us to calculate the cost for waiting:

$$C_f^2 = \mu_w \sum_m \sum_l \sum_k q_l^k \int_0^{s_{mk}} \int_{s_{ml}}^{h_l} (d_{ml} - d_{mk} - s_{ml} + s_{mk}) \lambda_{mk} e^{-\lambda_{mk} d_{mk}} \lambda_{ml} e^{-\lambda_{ml} d_{ml}} dd_{mk} dd_{ml}$$

where

$$\int_0^{s_{mk}} \int_{s_{ml}}^{h_l} (d_{ml} - d_{mk} - s_{ml} + s_{mk}) \lambda_{mk} e^{-\lambda_{mk} d_{mk}} \lambda_{ml} e^{-\lambda_{ml} d_{ml}} dd_{mk} dd_{ml}$$

$$= [(h_l - s_{ml}) e^{-\lambda_{ml} h_l} + \frac{1}{\lambda_{ml}} e^{-\lambda_{ml} h_l} - \frac{1}{\lambda_{ml}} e^{-\lambda_{ml} s_{ml}}] e^{-\lambda_{mk} s_{mk}}$$

$$-[(h_l - s_{ml} + s_{mk}) e^{-\lambda_{ml} h_l} - s_{mk} e^{-\lambda_{ml} s_{ml}} + \frac{1}{\lambda_{ml}} e^{-\lambda_{ml} h_l} - \frac{1}{\lambda_{ml}} e^{-\lambda_{ml} s_{ml}}] + \frac{1}{\lambda_{mk}} (e^{-\lambda_{ml} s_{ml}} - e^{-\lambda_{ml} h_l})(e^{-\lambda_{mk} s_{mk}} - 1)$$

c) Both vehicles are late, but the connecting vehicle in line l arrives after the feeder vehicle in line k, that is $d_{mk} - s_{mk} < d_{ml} - s_{ml}$. Then the waiting time is $w = d_{ml} - s_{ml} + s_{mk} - d_{mk}$. Integration allows us to calculate the cost for waiting:

$$C_f^3 = \mu_w \sum_m \sum_l \sum_k q_l^k \int_{s_{ml}}^{h_l} \int_{s_{mk}}^{s_{mk} - s_{ml} + d_{ml}} (d_{ml} - d_{mk} + s_{mk} - s_{ml}) \lambda_{mk} e^{-\lambda_{mk} d_{mk}} \lambda_{ml} e^{-\lambda_{ml} d_{ml}} dd_{mk} dd_{ml}$$

where

$$\int_{s_{ml}}^{h_l} \int_{s_{mk}}^{s_{mk} - s_{ml} + d_{ml}} (d_{ml} - d_{mk} + s_{mk} - s_{ml}) \lambda_{mk} e^{-\lambda_{mk} d_{mk}} \lambda_{ml} e^{-\lambda_{ml} d_{ml}} dd_{mk} dd_{ml}$$

$$= -[(h_l - s_{ml}) e^{-\lambda_{mk} s_{mk}} + \frac{1}{\lambda_{mk}} e^{-\lambda_{mk} (s_{mk} - s_{ml} + h_l)} - \frac{1}{\lambda_{mk}} e^{-\lambda_{mk} s_{mk}}] e^{-\lambda_{ml} h_l} - \frac{1}{\lambda_{ml}} e^{-\lambda_{mk} s_{mk}} (e^{-\lambda_{ml} h_l} - e^{-\lambda_{ml} s_{ml}})$$

$$+ \frac{1}{\lambda_{mk} + \lambda_{ml}} (e^{-(\lambda_{mk} + \lambda_{ml}) h_l - \lambda_{mk} (s_{mk} - s_{ml})} - e^{-\lambda_{ml} s_{ml} - \lambda_{mk} s_{mk}})$$

Then the total delayed connection cost of transfer passengers is the summation of all possible conditions:

$$C_f = C_f^1 + C_f^2 + C_f^3$$

$$= \mu_w \sum_m \sum_l \sum_k q_l^k \{ [s_{mk} + \frac{1}{\lambda_{mk}}(e^{-\lambda_{mk}s_{mk}} - 1)](1 - e^{-\lambda_{ml}s_{ml}}) + [(h_l - s_{ml})e^{-\lambda_{ml}h_l} + \frac{1}{\lambda_{ml}}e^{-\lambda_{ml}h_l} - \frac{1}{\lambda_{ml}}e^{-\lambda_{ml}s_{ml}}]e^{-\lambda_{mk}s_{mk}}$$

$$- [(h_l - s_{ml} + s_{mk})e^{-\lambda_{ml}h_l} - s_{mk}e^{-\lambda_{ml}s_{ml}} + \frac{1}{\lambda_{ml}}e^{-\lambda_{ml}h_l} - \frac{1}{\lambda_{ml}}e^{-\lambda_{ml}s_{ml}}] + \frac{1}{\lambda_{mk}}(e^{-\lambda_{ml}s_{ml}} - e^{-\lambda_{ml}h_l})(e^{-\lambda_{mk}s_{mk}} - 1)$$

$$- [(h_l - s_{ml})e^{-\lambda_{mk}s_{mk}} + \frac{1}{\lambda_{mk}}e^{-\lambda_{mk}(s_{mk} - s_{ml} + h_l)} - \frac{1}{\lambda_{mk}}e^{-\lambda_{mk}s_{mk}}]e^{-\lambda_{ml}h_l} - \frac{1}{\lambda_{ml}}e^{-\lambda_{mk}s_{mk}}(e^{-\lambda_{ml}h_l} - e^{-\lambda_{ml}s_{ml}})$$

$$= \mu_w \sum_m \sum_l \sum_k q_l^k [s_{mk} + \frac{1}{\lambda_{mk}}(e^{-\lambda_{mk}s_{mk}} + e^{-\lambda_{ml}h_l} - e^{-\lambda_{ml}(s_{mk} - s_{ml} + h_l) - \lambda_{ml}h_l} - 1) - (h_l - s_{ml} + s_{mk})e^{-\lambda_{ml}h_l}$$

$$- \frac{1}{\lambda_{ml}}(e^{-\lambda_{ml}h_l} - e^{-\lambda_{ml}s_{ml}}) + \frac{1}{\lambda_{mk} + \lambda_{ml}}e^{-\lambda_{mk}s_{mk}}(e^{-(\lambda_{mk} + \lambda_{ml})h_l + \lambda_{mk}s_{ml}} - e^{-\lambda_{ml}s_{ml}})]$$

This completes the derivation of Eq. (18).

Appendix 4 Derivation of Eq. (20)

By assumption (A3), the expected holding time for the bus on line k waiting for at least one delayed bus (e.g., a bus on line i at transfer node m) can be computed as follows:

$$E_{mki}(H) = \int_0^H tf_{mi}(t) \prod_{l,l \neq i} p(t_l > t)dt + \int_0^H tf_{mi}(t) \prod_{l,l \neq i} p(-s_{ml} < t_l < 0)dt$$

where $f_{mi}(t) \prod_{l,l \neq i} p(t_l > t)$ and $f_{mi}(t) \prod_{l,l \neq i} p(-s_{ml} < t_l < 0)$ maybe interpreted as the probability density function of the holding time for being connected by at least one bus conditional to waiting a bus on line i ($i \neq k$), the former corresponds to the case when all buses are delayed, while the latter corresponds to the case when only one bus are delayed.

where $f_{mi}(t)$ denotes a shifted exponential distribution as follows, which is equivalent to Eq.(2) by considering the slack time

$$f_{mi}(t) = \lambda_{mi}e^{-\lambda_{mi}(t + s_{mi})}, \ t > -s_{mi}$$

Then $p(t_l > t)$ and $p(-s_{ml} < t_l < 0)$ are given with the above distribution as follows:

$$p(t_l > t) = \int_t^{+\infty} \lambda_{ml}e^{-\lambda_{ml}(t_l + s_{ml})}dt_l = e^{-\lambda_{ml}(t + s_{ml})}$$

$$p(-s_{ml} < t_l < 0) = \int_{-s_{ml}}^0 \lambda_{ml}e^{-\lambda_{mi}(t_l + s_{ml})}dt_l = 1 - e^{-\lambda_{ml}s_{ml}}$$

The corresponding expected holding time with respective to the two cases are therefore

$$E_{mki}^1(H) = \int_0^H tf_{mi}(t) \prod_{l,l \neq i} p(t_l > t)dt = \int_0^H t\lambda_{mi}e^{-\lambda_{mi}(t+s_{mi})} \prod_{l,l \neq i} e^{-\lambda_{ml}(t+s_{ml})}dt$$

$$= \int_0^H t\lambda_{mi}e^{-\lambda_{mi}(t+s_{mi})}e^{-\sum_{l,l \neq i}\lambda_{ml}(t+s_{ml})}dt = \int_0^H t\lambda_{mi}e^{-\sum_i \lambda_{mi}(t+s_{mi})}dt = -\frac{\lambda_{mi}}{\sum_i \lambda_{mi}}\int_0^H tde^{-\sum_i \lambda_{mi}(t+s_{mi})}$$

$$= -\frac{\lambda_{mi}}{\sum_i \lambda_{mi}}[He^{-\sum_i \lambda_{mi}(H+s_{mi})} + \frac{e^{-\sum_i \lambda_{mi}s_{mi}}}{\sum_i \lambda_{mi}}(e^{-\sum_i \lambda_{mi}H}-1)]$$

$$E_{mki}^2(H) = \int_0^H tf_{mi}(t) \prod_{l,l \neq i} p(-s_{ml} < t_l < 0)dt = \prod_{l,l \neq i} p(-s_{ml} < t_l < 0)\int_0^H tf_{mi}(t)dt$$

$$= -[He^{-\lambda_{mi}(H+s_{mi})} + \frac{e^{-\lambda_{mi}s_{mi}}}{\lambda_{mi}}(e^{-\lambda_{mi}H}-1)]\prod_{l,l \neq i}(1-e^{-\lambda_{ml}s_{ml}})$$

Then we have

$$E_{mki}(H) = E_{mki}^1(H) + E_{mki}^2(H)$$

$$= -\frac{\lambda_{mi}}{\sum_i \lambda_{mi}}[He^{-\sum_i \lambda_{mi}(H+s_{mi})} + \frac{1}{\sum_i \lambda_{mi}}(e^{-\sum_i \lambda_{mi}(H+s_{mi})} - e^{-\sum_i \lambda_{mi}s_{mi}})]$$

$$-[He^{-\lambda_{mi}(H+s_{mi})} + \frac{e^{-\lambda_{mi}s_{mi}}}{\lambda_{mi}}(e^{-\lambda_{mi}H}-1)]\prod_{l,l \neq i}(1-e^{-\lambda_{ml}s_{ml}})$$

Minimize $E_{mki}(H)$ over all lines i ($i \neq k$) gives the expected holding time for the bus on line k waiting for at least one delayed bus

$$E_{mk}(H) = \min_{i,i \neq k} E_{mki}(H)$$

$$= \min_{i,i \neq k}\{-\frac{\lambda_{mi}}{\sum_i \lambda_{mi}}[He^{-\sum_i \lambda_{mi}(H+s_{mi})} + \frac{e^{-\sum_i \lambda_{mi}s_{mi}}}{\sum_i \lambda_{mi}}(e^{-\sum_i \lambda_{mi}H}-1)]$$

$$-[He^{-\lambda_{mi}(H+s_{mi})} + \frac{e^{-\lambda_{mi}s_{mi}}}{\lambda_{mi}}(e^{-\lambda_{mi}H}-1)]\prod_{l,l \neq i}(1-e^{-\lambda_{ml}s_{ml}})\}$$

This completes the derivation of Eq. (20).

Appendix 5 Derivation of Eq. (23)

Proof. Assuming that q_{mk} is the expected total impacted number of boarding passengers on specified link. The boarding demands are distributed homogeneously along the line from the respective transfer node to the downstream location that the holding delay just totally be recovered. Therefore, we can assume the demand of segment differential to be q_{mk}/n, and by assumption (A5), the recovery time in each segment differential to be δ. Then we have $E_{mk}(H) = n\delta$.

The additional waiting time for sequential segment differential is as follows:

$$E_{mk}(H), \ E_{mk}(H) - \delta, ..., \ E_{mk}(H) - (n-1)\delta$$

Then the total extra waiting time for downstream passengers is the summation of individual extra waiting time of sequential segment differential, i.e.

$$\frac{q_{mk}}{n}E_{mk}(H) + \frac{q_{mk}}{n}(E_{mk}(H) - \delta) + ... + \frac{q_{mk}}{n}[E_{mk}(H) - (n-1)\delta]$$

Thus, the respective cost is calculated as

$$
\begin{aligned}
C_d &= \mu_w \sum_m \sum_k \lim_{n \to \infty} \{ \frac{q_{mk}}{n}E_{mk}(H) + \frac{q_{mk}}{n}(E_{mk}(H) - \delta) + ... + \frac{q_{mk}}{n}[E_{mk}(H) - (n-1)\delta] \} \\
&= \mu_w \sum_m \sum_k \lim_{n \to \infty} \frac{q_{mk}}{n}[nE_{mk}(H) - \frac{(0+n-1)n}{2}\delta] \\
&= \mu_w \sum_m \sum_k \lim_{n \to \infty} \frac{q_{mk}}{n}[nE_{mk}(H) - \frac{(0+n-1)n}{2}\frac{E_{mk}(H)}{n}] \\
&= \mu_w \sum_m \sum_k \lim_{n \to \infty} \frac{q_{mk}}{n}\frac{n+1}{2}E_{mk}(H) = \frac{1}{2}\mu_w \sum_m \sum_k q_{mk}E_{mk}(H)
\end{aligned}
$$

This completes the derivation of Eq. (23).

Reference

Bookbinder, J.H., and Desilets, A., 1992. Transfer optimization in a transit network. Transportation Science, 26(2), 106-118.

Braekers, K., Caris, A., Janssens, G.K., 2014. Exact and meta-heuristic approach for a general heterogeneous dial-a-ride problem with multiple depots. Transportation Research Part B, 67, 166-186.

Berrebi, S., Watkins, K., Laval, J., 2015. A real-time bus dispatching policy to minimize passenger wait on a high frequency route. Transportation Research Part B, 81, 377-389.

Ceder, A., Wilson. N.H.M., 1986. Bus network design. Transportation Research Part B, 20(4), 331-344.

Ceder, A., 2007. Public Transit Planning and Operation: Theory, Modeling and Practice. Elsevier, Butterworth Heinemann.

Ceder, A., Golang, B., Tal, O., 2001. Creating bus timetables with maximal synchronization. Transportation Research Part A, 35(10), 913-928.

Chien, S., Schonfeld, P., 1998. Joint optimization of a rail transit line and its feeder bus system. Journal of Advanced Transportation, 32(3), 253-284.

Chen, M., Liu, X., Xia, J., 2005. Dynamic prediction method with schedule recovery impact for bus arrival time. Transportation Research Record, 1923, 208-217.

Chowdhury,M.S, Chien, S.I.J..2001. Dynamic vehicle dispatching at the intermodal transfer station. Transportation Research Record, 1753, 61-68.

Chowdhury, M.S., Chien, S.I.J. 2002. Intermodal transit system coordination. Journal of Transportation Planning and Technology, 25(4), 257-288.

Chowdhury, M.S., Chien, S.I.J., 2011. Joint optimization of bus size, headway, and slack time for efficient timed transfer. Transportation Research Record, 2218, 48-58.

Chung, E.H, Shalaby, A., 2007. Development of control strategy for intermodal connection protection of time-transfer transit routes. Transportation Research Record, 2006, 3-10.

Dessouky, M., Hall, R., Nowroozi A., Mourikas K., 1999. Bus dispatching at a timed transfer transit stations using bus tracking technology. Transportation Research Part C, 5(4), 187-208.

Dessouky, M., Hall R., Zhang, L., Singh, A., 2003. Real-time control of buses for schedule coordination at terminal. Transportation Research Part A, 37(2), 145-164.

Daganzo, C. F., 2009. A headway-based approach to eliminate bus bunching: Systematic analysis and comparisons. Transportation Research Part B, 43(10), 913-921.

Daganzo, C.F., 2004. Logistics Systems Analysis. Springer-Verlag, Heidelberg, Germany.

Dalgado, F., Murioz, J.C., Ciesen, R., 2012. How much can holding and/or limiting boarding improve transit performance? Transportation Research Part B, 45(1), 267-277.

Eberlein, X.J., Wilson, N.H.M., Bernstein, D., 2001. The holding problem with real-time information available. Transportation Science, 35(1), 1-18.

Eranki, A., 2004. A Model to Create Bus Timetables to Attain Maximum Synchronization Considering Waiting Times at Transfer Stops. Master's Thesis, University of South Florida.

Fonzone, A., Schmocker, J-D, Liu, R., 2015. A model of bus bunching under reliability based passenger arrival patterns. Transportation Research Part C, 59,164-182.

Guenthner, R.P, Hamat, K., 1983. Distribution of bus transit on-time performance. Transportation Research Record, 1202, 1.

Hadas, Ceder, A., 2010. Optimal coordination of public-transit vehicles using operational tactics examined by simulation. Transportation Research Part C, 18, 879-895.

Hernandez D., Munoz J.C., Giesen R., Delgazo. F., 2015. Analysis of real-time control strategies in a corridor with multiple bus services. Transportation Research Part B, 78, 83-105.

Hollander, Y., Liu, R., 2008. Estimation of the distribution of travel times by repeated simulation. Transportation Research Part C, 16(2), 212-231.

Hickman, M., 2001.An analytic stochastic model for the transit vehicle holding problem. Transportation Science, 35(3), 215-237.

Ibarra-Rojas, O.J, Rios-Solis, Y.A. 2012. Synchronization of bus timetabling. Transportation Research Part B, 46(5), 599-614.

Ibarra-Rojas, O.J., Loprz-Irarragorri F., Rios-Solis, Y.A., 2012.Multiperiod bus timetable. Transportation Science, in press, doi: 10.1287/trsc.2014.0578.

Ibarra-Rojas, O.J., Delgado, F., Giesen, R., Munoz, J.C., 2015. Planning, operation, and control of bus transport systems: A literature review. Transportation Research Part B, 77, 38-75.

Ji, Y., He, L., Zhang, H.M., 2014. Bus drivers' responses to real-time schedule adherence and the effects on transit reliability. Transportation Research Record, 2417, 1-9.

Lee, M., Schonfeld, P., 1991. Optimal slack time for timed transferred at transit terminal. Journal of Advanced Transportation, 25(3), 281-308.

Liu, M., Luo, Z., Lim, A., 2015. A branch-and-cut algorithm for a realistic dial-a-ride problem. Transportation Research Part B, in press. doi:10.1016/j.trb.2015.05.009.

Lin, W., Bertini, R.L. 2004. Modelling schedule recovery processes in transit operations for bus arrival time prediction. Journal of Advanced Transportation, 38(3), 347-365.

Lu, J.J, Park, J., Pernia, J.C. and Dissanayake, S., 2003. Criteria for setting speed limits in urban and suburban areas in Florida. (No. Final Technical Report)

Meng, Q., Wang, S., Lee, C.Y., 2015. A tailored branch-and-price approach for a joint tramp ship routing and bunkering problem. Transportation Research Part B, 72, 1-19.

Moccia, L., Laporte, G., 2016. Improved models for technology choice in a transit corridor with fixed demand. Transportation Research Part B, 83, 245-270.

Nesheli M.M, Ceder A. 2014. Optimal combinations of selected tactics for public-transport transfer synchronization. Transportation Research Part C, 48, 491-502.

Parbo, J., Nielsen, O.A., Parto, C.G., 2014. User perspectives in public transport timetable optimization. Transportation Research Part C, 48, 269-284.

Qiu, F., Li, W.Q., Zhang. J., 2014. A dynamic station strategy to improve the performance of flex-route transit services. Transportation Research Part C, 48, 229-240.

Osuna, E.E., Newell, G.F., 1972. Control strategies for an idealized public transportation system. Transportation Science, 6(1), 52-72.

Quarmy, D.A., 1967. Choice of travel mode for the journey to work. Journal of Transport Economics and Policy, 1(3), 273-314.

Randolph W.H., 2001. Vehicle scheduling at a transportation terminal with random delay enroute. Transportation Science, 19(3), 308-320.

Sun, Y., Schonfeld, P., 2016. Holding decisions for correlated vehicle arrivals at intermodal freight transfer terminals. Transportation Research Part B, 90, 218-240.

Shafahi, Y., Khani, A., 2010. A practical model for transfer optimization in a transit network: model formulation and solutions. Transportation Research Part A, 44(6), 377-389.

Strathman, J.G., Hopper, J.R., 1993. Empirical analysis of bus transit on-time performance. Transportation Research, 27A, 93.

Ting, C.J, Schonfeld, P., 2005. Schedule coordination in a multiple hub transit network. Journal of Urban Planning and Development, 131(2), 112-124.

Ting, C.J, Schonfeld, P., 2007. Dispatching control at transfer stations in multi-hub transit networks. Journal of Advanced Transportation, 41(3), 217-243.

Ting, C.J. 1997. Transfer Coordination in Transit Networks. PhD dissertation, Civil Engineering Department, University of Maryland, College Park, Md.

Tirachini, A., Hensher, D.A., Jara-Diaz, S.R., 2010. Restating modal investment priority with an improved model for public transport analysis. Transportation Research Part E, 46(6), 1148-1168.

Vansteenwegen, P., Qudheusden, D.V., 2007. Decreasing the passenger waiting time for an intercity rail network. Transportation Research Part B, 41(4), 478-492.

Wu, W., Jin, W., Shen, L., 2013. Mixed platoon flow dispersion model based on speed-truncated Gaussian mixture distribution. Journal of Applied Mathematics, Volume 2013, Article ID 480965, 9 pages.

Wu, W., Shen, L., Jin, W., Liu, R., 2015. Density-based mixed platoon dispersion modelling with truncated mixed Gaussian distribution of speed. Transportmetrica B, 3(2), 114-130.

Wu, Y., Tang, J., Yu, Y., Pan, Z., 2015. A stochastic optimization model for transit network timetable design to mitigate the randomness of traveling time by adding slack time. Transportation Research Part C, 52, 15-31.

Wong, R.C.W, Yuen, T.W.Y., Fung, K.W, Leung, J.M.Y. 2008. Optimizing Timetable Synchronization for Rail Mass Transit. Transportation Science 42(1), 57-69.

Xuan Y.G, Argote, J, Daganzo, C. F., 2011. Dynamic bus holding strategies for schedule reliability: Optimal linear control and performance analysis. Transportation Research Part B, 45(10), 1831-1845.

Yan, Y.D, Meng, Q., Wang, S.A., Guo, X.C., 2012. Robust optimization model of schedule design for a fixed bus route. Transportation Research Part C, 25, 113-121.

Yu, B., Wu, S., Yao, B., Yang, Z., Sun, J., 2012. Dynamic vehicle dispatching at a transfer station in public transportation system. Journal of Transportation Engineering, 138(2), 191-201.

Zhao, J., Dessouky, M., Bukkapatnam, S., 2006. Optimal slack time for schedule-based transit operations. Transportation Science, 40(4), 529-539.

Zhou, X., Zhong, M., 2007. Single-track train timetabling with guaranteed optimality: Branch-and-bound algorithms with enhanced lower bounds. Transportation Research Part B, 41(3), 320-341.