Optimal Bus Service Design with Limited Stop Service in a Travel Corridor

David Z.W. Wang[1]*, Ashish Nayan[1], W.Y. Szeto[2]

[1] School of Civil and Environmental Engineering, Nanyang Technological University, 50 Nanyang Avenue, 639798, Singapore

[2] Department of Civil Engineering, The University of Hong Kong, Pok Fu Lam, Hong Kong

*Corresponding Author

ABSTRACT

This paper seeks to answer questions from the combined bus operator's and users' perspective on how to design limited stop service operation strategies when they are offered along with the normal bus services. The passengers' service choice is determined by the common line calculation. The problem is formulated as a Mixed Integer Nonlinear Program (MINLP) with equilibrium constraints. Thereafter, a global optimal solution method applying various linearization and convexification techniques is proposed. Numerical studies are then performed to evaluate the model validity and solution efficiency followed by concluding remarks.

Keywords: Limited stop service, line setting, attractive lines, mixed-integer nonlinear programming, global optimization

1. Introduction

Public transit services are lifelines for daily commute in many major cities in the world. In order to increase the service quality, constant improvement in operation and design is of paramount importance. In the presence of increasing daily travel demand, public transit service operators now seek to improve their service quality to efficiently satisfy the travel demand while maintaining operation in a financially sustainable manner. In many cities, bus transit services have become more convenient with the inclusion of differential services such as normal, express, and limited stop services which are operated to cater to various demand patterns. While a normal service serves all the bus stops/nodes on a route, an express service travels end to end without or with very few intermediate stoppages. A limited stop service serving a selected subset of nodes in a corridor provides another alternative and helps transit operators in reducing overall passenger travel time. Hence, a limited stop service is of reasonable financial and social importance to bus transit operation and due academic attention needs to be given to developing methodologies for bus operators to design their operation strategies.

In the literature, transit corridor design problems have attracted much research attention. Ceder and Wilson (1986) discussed the bus route planning problem that minimizes total system operation cost while also addressing the scheduling problem. Since then, a vast body of literature on transit corridor design has emerged which involves optimal decisions of routing and scheduling, service frequency design, inter-node spacing, fleet size design, etc. Curtin and Biba (2011) proposed a mathematical model that maximizes the service value of a route, rather than minimizing its cost, and the cost (distance) is considered as a budget constraint on the extent of the route.

Wang and Lo (2008) presented a related work on a multi-fleet ferry routing and scheduling problem that considered ferry services with different operational characteristics. Cortés et al. (2011) presented a methodology to optimise costs while integrating two kinds of services in the transit network with deadheading and short turning services. Yadan et al. (2012) proposed a robust optimization model for the bus route schedule design problem by taking into account the bus travel time uncertainty and the bus drivers' schedule recovery efforts. A few studies discussed the bus dwelling time which is critical towards determining the total travel time of passengers (e.g., Meng and Qu, 2013; Sun et al., 2013). Bus transit generally operates under different market regimes and a few studies in the literature have also contributed towards this aspect (e.g., Li et al., 2010; Li et al., 2008). Liu and Meng (2012) modelled the network flow equilibrium problem on a multimodal transport network with a bus-based park-and-ride system and congestion pricing charges. Li et al. (2011) addressed the design problem of a rail transit line located in a linear urban transportation corridor where the service variables include a combination of rail line length, number and locations of stations, headway, and fare. In addition to the above mentioned studies, there exist many other published works on transit service design; but unfortunately, few studies focus on the methodological design of a limited stop service in bus transit.

Limited stop services have been operating in cities like Bogota, Chicago, Montreal, New York City, Santiago. Afanasiev and Liberman (1983) described a limited stop service as a service with stops at intervals of about 0.8 km. Silverman (1998) proposed a few important considerations while designing a limited stop service: wider roadways, not too close to rapid transit corridors, operationally more successful over long distances. Conlon et al. (2001) noted that implementing a limited stop service parallel to a normal bus service drew appreciation from users in Chicago where user satisfaction for both the services increased after the inclusion of the former. El-Geneidy and Surprenant-Legault (2010) observed that a limited stop service is the most preferred choice of passengers as they tend to overestimate their time savings while using this service. Tétreault and El-Geneidy (2010) proposed a stop selection methodology for limited stop services based on archived Automatic Vehicle Location (AVL) and Automatic Vehicle Classification (AVC) data obtained from a travel behaviour survey in Montreal, Canada. This included different scenarios wherein stops were selected based on passenger activity and transfers. As it can be concluded, studies mentioned above mainly focused on the operational aspect of limited stop services which is data-driven and descriptive while no analytical approach was proposed for the service design.

In designing a limited stop service, bus stop selection is the prime decision variable, i.e., to determine which stops the bus service should stop or skip in the transit corridor. In addition, other operation strategies like optimal fleet size, service frequency and bus capacity should be determined with consideration of the passengers' service choices. Some research studies have been conducted to develop methodological frameworks to prescribe guidelines for their operation in terms of optimal service design. Larrain et al. (2010) proposed the methodology to select optimal express services for a bus corridor with capacity constraints considering various demand criteria, whereas, Larrain et al. (2015) designed zonal bus services which skip all intermediate nodes over a segment of the transit route while serving all nodes in the initial and final segment. Ulusoy et al. (2010) presented a methodology to optimize the operation of integrated normal, short turn, and express services. Leiva et al. (2010) presented an optimization approach to design a limited stop service with capacity constraints. However, in this work, the selection of bus stops for the limited stop service is given in priori, and the service frequency of limited stop services lines is the only operation strategy determined by the model, despite the fact that they discard some of the services assigning zero frequency. Using only several given subsets of bus stops as the candidate service design plan for the limited stop service, one cannot obtain the truly "best" bus service design for limited stop services. Chiraphadhanakul and Barnhart (2013) proposed a design of the limited stop service by optimally reassigning certain bus trips rather than providing additional trips. However, this work does not consider transfers or multiple lines operating over common route corridors where passengers could make a choice. Besides, it allows only one limited stop service to be operated over the transit network and the frequency of the limited stop service is not taken into account for passenger assignment on the respective services. Recently, Larrain and Muñoz (2016) proposed a design algorithm for limited stop services in a corridor to optimise a number of services and then calibrated a regression model to estimate the benefits. Hart (2016) developed a methodology for transit agencies to evaluate the potential for limited stop service along existing bus routes where net benefits of travel time savings would outweigh the net costs as a result of implementation of limited stop service. Zhang et al. (2016) proposed a methodology to determine frequencies and skip-stop strategy; however, genetic algorithm was used to develop the model. A detailed comparison between some of the above mentioned studies on limited stop services which are closer to the contribution in this study is illustrated in Table 1.

In this paper, we present a mathematical model formulation to explicitly design a limited stop service with optimal decisions on bus line configurations (the set of bus stops served by the limited stop service) along with other operation strategies including operating frequencies and the optimal fleet size assignment. The model developed primarily considers the perspective of operators. Basically, given a fixed bus fleet size, the bus operators who decide to offer a limited stop service other than the normal bus service need to determine optimal operation strategies pertaining to service fleet size, line setting for limited stop services, and the service frequencies so as to minimize the total operation cost. At the same time, due consideration must be given to the service performance from the perspective of passengers as poor service performance may lead to a drop in demand or the possibility of losing the franchise of operating the routes altogether. Therefore, the objective function also incorporates passengers' travel and waiting time as important factors which are to be adjusted by appropriate weights. Although this study assumes that the bus services are operated in a monopolistic market with fixed total demand, it should be noted that the model framework (which focuses on how to model and solve the optimal operation strategies for limited stop services) can be easily extended to consider elastic demand or competition with existing alternative bus services.

Table 1 Comparison between a few published studies on limited stop service design and this work

Factors\Studies	**Leiva et al. (2010)**	**Ulusoy et al. (2010)**	**Chiraphadhanakul et al. (2013)**	**This work**
Assumptions				
Origin-destination (O-D) matrix	Fixed	Fixed	Fixed	Fixed
Transfers	Allowed	Allowed	Not Allowed	Allowed
Number of limited stop services allowed	Unlimited	Unlimited	1	Unlimited
Objective	Minimize social costs	Minimize social costs	Maximize social welfare	Minimize social costs
Decision variables				
Fleet size	✓	✓	✓	✓
Frequency	✓	✓	-----	✓
Explicit design of a new limited stop service	----- (Given a predefined set of candidate services)	----- (Given a predefined set of candidate services)	✓	✓
Constraints				
Choice behavior	DUE	SUE (stochastic	System Optimal	DUE (deterministic

	(deterministic user equilibrium)	user equilibrium)		user equilibrium)
Common line approach	✓	-----	-----	✓
Capacity	✓	✓	✓	✓
Fleet size	✓	✓	✓	✓
Assignment	Proportional to frequency of each attractive line	Logit based model considering wait, transfer, in-vehicle times	A linear function of frequency share and in-vehicle travel time savings	Proportional to frequency of each attractive line
Multiple services along route segments	✓	✓	-----	✓
Shorter running times of limited stop services	✓	✓	-----	✓
Incorporating limited stop service(s) frequency in demand model	✓	✓	-----	✓

In designing the transit corridor with the normal service and limited stop service, one intrinsic issue to be considered is the travellers' choice behaviour between different services. In this study, such passenger choices are described by the classical common line problem. Specifically, it is assumed that the travellers choose a subset of bus services that minimizes the expected total travel time which was defined as the common line problem in Chriqui and Robillard (1975). The common line problem has been investigated in many other research works in the literature such as Spiess and Florian (1989), De Cea and Fernandez (1993), Cepeda et al. (2006).

In this study, the model is formulated into a MINLP. One may consider this bus service design problem with limited stop services as a bi-level mathematical program; the lower level program describes the passenger assignment problem whereas the upper level program is the bus service design problem. Alternatively, in this study, we formulate the lower level passenger assignment problem as equivalent mathematical conditions, thus reducing the bi-level program into a mathematical program with equilibrium constraints (MPEC). The formulated MINLP is inherently non-convex and even if the integrality is solved, the problem still remains non-convex. Hence, we devise a solution algorithm applying various linearization and convexification techniques to find the global optimal solution. A global optimal solution guarantees the best possible operation strategy and, therefore, it is necessary for bus service operators to have such a solution if the objective of minimizing the total operation cost is to be achieved. It should be noted that most of the previous research works on transit service design did not guarantee a global optimal solution. One can further note that the constraints in the model formulations of some bus service design problems in the literature were simply removed or relaxed to facilitate obtaining the solution efficiently; but unfortunately, this does affect the solution validity to a certain extent.

In summary, this paper contributes to the literature in two major aspects. First, we develop a mathematical model to fully address the optimal design of limited stop bus services which explicitly determines bus line configurations (the set of bus stops served by the limited stop services), operation frequencies and fleet size assignment while considering passengers' service choices. Previous research assumed pre-determined limited stop bus services, which simplified the problem and circumvented complicated model formulation, however, compromised the model's capability of determining the truly best limited stop bus services. Second, a global optimal solution method is developed to solve the model formulation. Due to the inherent non-convexity of the formulated problem, previous studies in transit service design did not achieve a global optimal solution.

The paper is structured as follows: Section 2 presents the problem with the model formulation. Section 3 discusses solution methodology. Section 4 is the numerical study which illustrates the accuracy of the methodology and problem properties and finally, section 5 gives the concluding remarks.

2. Model formulation

In this section, we present a model formulation to optimally design the limited stop service. Before the model formulation; we define the general corridor setting and notation, make a few preliminary assumptions and state the general mathematical constraints for a limited stop service.

2.1. Corridor setting

The transit corridor considered is single corridor, linear and numbered such that the last node number is equal to the total number of nodes in the corridor. A bus cycle refers to one complete trip from the first node to the last node of the corridor. A loop service may be considered if the first and the last nodes are the same. For the purpose of illustration, we show an example transit corridor of 5 nodes numbered from 1 to 5 as in Figure 1.

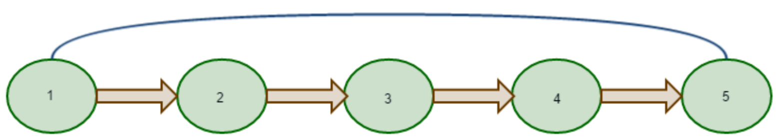

Fig. 1 Transit corridor

In this corridor, passengers can utilize the transit services to travel from any origin node *i* to any other destination node *j* (for all *j* that is greater than *i*). The normal services cater to every node of the corridor whereas the limited stop services serve only a certain subset of all the bus stops which are to be determined through optimization and are referred to as *special nodes*. However, all services to be considered in our model start at node 1 and end at the last node which can be referred to as node *b*, which is node 5 here in Fig. 1. The operator has a fixed fleet of buses which is to be split amongst the operating services.

In Fig.1, let us consider the O-D pair (1-5) connected by an arc as shown. For travelling between this O-D pair, passengers use multiple route sections (sets of transfer nodes) such that they can transfer at intermediate nodes. Hence, starting at node 1, passengers can travel on various route sections to reach to their destination node. One of the strategies could be as follows: take route section 1-3, transfer at node 3, take route section 3-5 and finally reach the destination. On each of these chosen route sections, the passengers can select any of the bus transit services. As already mentioned, the normal service serves all the nodes and, hence, every route section. However, the limited stop service would serve only selected route sections depending on the optimal line setting. Hence, depending on whether the limited stop service serves a particular route section or not, the passengers travelling on that route section would decide on which bus service/line to take. This is indeed the “common line” problem, as is further explained in the later part of this section.

2.2. Variable definitions

Sets and notations

N	Set of nodes in the transit corridor, $i \in N$
W	Set of O-D pairs $w \in W$
o,d	Indices for origin and destination of any O-D pair $w \in W$
S_{ij}	Set of node pairs
L	Set of all transit lines $l \in L$
L'	Subset of limited stop services, $L' \subset L$
ij	Route section joining nodes *(i,j)*
l_s	Limited stop service $l_s \in L'$
l_r	Normal service
l_{ij}	Any transit line serving route section *ij*

Parameters

W_c	Value of waiting time
T_c	Value of travel time
k	Parameter whose value depends on the distribution of bus arrival times at stops
T^h	Fixed dwelling time at node *h*
T_{ij}	Non-stop running time between node pair (i, j)
K_l	Bus ownership cost of any line *l*
X^w	Exogenous demand for O-D pair *w*
B	Available fleet size
Cap_l	Passenger capacity of a bus on any line *l*
λ_{trans}	Coefficient to convert the number of transfers to cost terms
b	Cardinality of the set *N* (number of nodes in the bus service corridor)
F^l	Operating cost per cycle of any transit line *l*
α, β	Time delay due to the alighting or boarding activity per passenger

Definitional variables

y^{i,l_s}	Binary variable, it equals one if the limited stop service l_s serves node i.
t_{ij}^{l}	Travel time between node pair *(i, j)* on line *l*
V_{ij}^{w}	Passenger flow over route section ij for an O-D pair *w*
x_{ij}^{l}	Line *l* over route section ij in the common line problem is attractive if it takes a value of one, binary variable
$v_{ij}^{w,l}$	Passenger flow on any line *l* over route section ij for an O-D pair *w*
$t^{l,cycle}$	Travel time for one bus cycle on any line *l*

Decision variables

n_l	Fleet size allocated to any line *l*
f^l	Operating frequency of any line *l*

$y_{ij}^{l_s}$ Binary variable, denoting the direct service between node pair *(i, j)* for a limited stop service l_s exists if it equals one

2.3. Assumptions

2.3.1. Passenger demand

Exogenous demand for each OD pair is assumed to be given. However, the demand for a particular bus service depends on its service quality and is determined by the choice behaviour of passengers.

2.3.2. In-vehicle travel time

It is assumed that the running time between two consecutive stops along the service is exogenously given; the standard bus dwelling time at each stop is given. For a given node pair in a given route, the total in-vehicle travel time is determined by the running time plus dwelling times at the stops located between these nodes. Therefore, the in-vehicle travel time for a limited-stop service is determined by its line setting, i.e., which stops to be served by the limited stop service; whereas for the normal service which serves all intermediate nodes, the in-vehicle travel time is fixed if the dwelling time at each bus stop is assumed to be fixed and given. In this study, effects of boarding and alighting on bus stop dwell time are also considered.

2.3.3. Choice behaviour

The model formulation captures choice behaviour through the attractive/common line approach (Chriqui and Robillard, 1975). It assumes that passengers consider only a subset of lines serving a pair of nodes and the first arriving bus among the subset of lines is chosen by the passengers. The factors primarily affecting bus service choice include service waiting time and in-vehicle travel time of the lines. The model formulation also takes bus service capacity into account to assign passengers to each of the operating services.

2.3.4. Total costs

As in Leiva et al. (2010), the total costs for the bus corridor comprise of *operator costs* and *user costs*. *Operator costs* in our model formulation are a combination of (*i*) *bus ownership costs* which is basically the cost of owning or renting a bus for operation on a particular service (the total ownership cost of a particular service is the product of the number of buses and the unit ownership cost of the type of bus given by $K_l n_l$) (*ii*) *bus operating costs* which accounts for variable operation costs such as employment costs, taxes, licenses and insurance (total operating cost for a service *l* can be given by $F^l f^l$). Hence, it is noted that the total operating cost term for a service *l* does not directly include cycle time and hence, is kept independent of the number of stops and the cycle time of any transit line *l*. However, if it is desired to include the effect of cycle time into the operating cost, the operating cost per cycle F^l can be easily expressed as a linear function of the cycle times, which are further defined in constraints (13) and (14) in the model formulation. As regards the *user costs*, they are a combination of *passenger waiting time costs* and *passenger travel time costs* over route sections for each O-D pair.

2.4. General mathematical constraint for a limited stop service

In this section, we model the service corridorof the limited stop service through a mathematical formulation. For the limited stop service l_s, the possible direct service ($y_{ij}^{l_s}$) between any node pair *(i, j)* can be shown as below in Fig. 2:

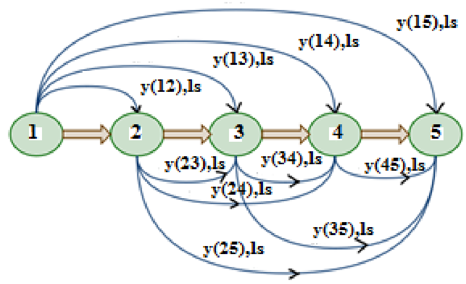

Fig. 2 Possible direct limited stop services between different node pairs in the transit corridor

The service corridor of the limited stop service can be modelled by the following conditions:

$$y^{i,l_s} = \sum_{j>i} y_{ij}^{l_s} \leq 1, \forall i \in N \setminus \{1,b\}, \forall l_s \in L'. \quad (1)$$

$$y^{1,l_s} = 1, y^{b,l_s} = 1, \forall l_s \in L'. \quad (2)$$

$$y_{ij}^{l_s} \in \{0,1\}, \forall l_s \in L'. \quad (3)$$

The binary variable $y_{ij}^{l_s}$ represents the direct limited stop service, which takes the value 1 if there is a direct limited stop service between node pair (*i, j*) with no intermediate stoppages in between and 0 otherwise. The variable y^{i,l_s} defined in constraint (1) describes whether there is a limited stop service from this particular node *i* towards any node *j* located further in the corridor. Eq. (2) ensures that the limited stop service starts from the first node 1 and ends its service at the last node *b*. Hence, at any intermediate node *i*, the maximum value of the flow variable y^{i,l_s} is equal to 1 for a certain limited stop service l_s, as defined in constraint (1). Constraint (3) defines that $y_{ij}^{l_s}$ is indeed a binary variable.

2.5. Model formulation for the design of limited stop services

2.5.1. Problem Description

In this section, a methodology is presented to design the limited stop service operating in conjugation with the normal service in terms of line setting, operating frequency and the fleet size with the attractive/common line approach. The problem description in this study is similar to that presented in Leiva et al. (2010).To describe the passengers' service choices, the concept of route sections (e.g., De Cea and Fernandez (1993)) is applied (each route section connecting nodes *i*,*j* denoted as *ij*). The normal service operates over all route sections throughout the corridor while a limited stop service operates over certain route sections depending on its line setting. At a particular node, a passenger intending to travel on a route section in the direction towards the destination node only boards a service that is attractive over that route section. As was defined in Chriqui and Robillard (1975), the attractiveness of a particular service over a route section is determined by the total expected travel time including the waiting time and in-vehicle travel time. Allowing for the transfers, the model assumes that passengers have the liberty to travel in stages over route sections towards their respective destinations. Starting from his/her origin node, a passenger boards the first bus from the set of attractive lines over a certain route section and transfers to the next route section until he/she reaches the destination node.

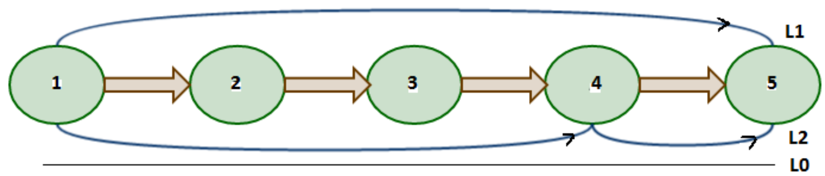

Fig. 3 Example transit corridor with line setting for different limited stop services

Consider the above example with two limited stop services l_1 and l_2, and the normal service l_0, as shown in Fig. 3. The line settings of l_1 and l_2 are 1->5 and 1->4->5 respectively. Table 2 shows the set of lines serving different route sections as per the given transit line settings. A passenger can choose one of these lines to travel over the route section depending on whether it is serving the route section and whether it is attractive. A particular bus line serving a route section is attractive when the passenger's travel time cost over this line is smaller than the combined travelling and waiting time cost of all other lines serving the same route section.

Table 2 Route sections and operating lines

Route section	1-2	1-3	1-4	1-5	2-3	2-4	2-5	3-4	3-5	4-5
Lines	l_0	l_0	l_0, l_2	l_0, l_1, l_2	l_0	l_0	l_0	l_0	l_0	l_0, l_2

De Cea and Fernandez (1993) presented a hyperbolic programming problem to find the set of common/attractive lines operating over a route section. If A_s constitutes the set of transit lines operating over a route section s and (t^l, f^l) defines the in-vehicle travel time and frequency of a bus line l operating over s, then the following optimization problem determines the set of common lines on route section s:

$$\min_{\{x^l\}} \frac{1+\sum_{l=1}^{k'} t^l f^l x^l}{\sum_{l=1}^{k'} f^l x^l} \tag{4}$$

subject to $x^l \in \{0,1\}, \forall l \in A_s$, where $k' = |A_s|$ is the number of lines operating over s and x^l equals one if line l is an attractive line to the passengers and zero otherwise. In formulation (4), the objective function to be minimized describes the expected total travel time if passengers board on the first arriving bus service within the set of common/attractive lines. Specifically, the first term (inverse of the frequency, i.e., $\frac{1}{\sum_{l=1}^{k'} f^l x^l}$) denotes the expected waiting time at bus stop; while the second term (i.e., $\frac{\sum_{l=1}^{k'} t^l f^l x^l}{\sum_{l=1}^{k'} f^l x^l}$) depicts the expected in-vehicle travel time. Hence, the set of common lines are determined such that the combined waiting and travelling time for passengers using these lines is minimized. More detailed information on this formulation can be referred to De Cea and Fernandez (1993).

While travelling over a particular route section, passengers might face a situation wherein they need to choose between the different lines/services serving the route section. The common line problem determines how passengers select the set of common lines such that passengers can board any of these lines, whichever comes first. It is interesting to note that, as congestion on the common lines increases, the lines which are not attractive also become attractive. This is because, the congestion on the common lines increases their dwell time at nodes and the expected waiting time at the downstream nodes, making the less attractive lines more attractive. It is noted that the bus service congestion is not considered in this study.

2.5.2. Model Formulation

The model formulation for the design of limited stop services is defined through an optimization problem as given below.

$$\min_{\{n_l, f^l, y_{ij}^{l_s}\}} Z = \sum_{l\in L} K_l n_l + \sum_{l\in L} F^l f^l + \sum_{w\in W}\sum_{ij\in S_{ij}} V_{ij}^{w} \frac{W_c k}{\sum_{l\in L} x_{ij}^{l} f^{l}} + \sum_{w\in W}\sum_{ij\in S_{ij}}\sum_{l\in L} T_c v_{ij}^{w,l} t_{ij}^{l} + \lambda_{trans} \sum_{w\in W} \left(\sum_{ij\in S_{ij}} V_{ij}^{w} - X^{w}\right) \tag{5}$$

Subject to:

$$\sum_{l\in L} n_l \le B. \tag{6}$$

$$t^{l,cycle} f^{l} \le n_l, \forall l \in L. \tag{7}$$

$$v_{ij}^{w,l} = V_{ij}^{w} \frac{x_{ij}^{l} f^{l}}{\sum_{l\in L} x_{ij}^{l} f^{l}}, \forall l \in L, \forall w \in W. \tag{8}$$

$$\sum_{ij\in S_i^{+}} V_{ij}^{w} - \sum_{ij\in S_i^{-}} V_{ij}^{w} = \begin{Bmatrix} X^{w}; i = o(w) \\ -X^{w}; i = d(w) \\ 0; otherwise \end{Bmatrix}, \forall w \in W. \tag{9}$$

$$\sum_{i\in N\setminus b}^{b-1} \sum_{j=i+1\in N}^{b} \sum_{w\in W} y^{i,l_s} y^{j,l_s} v_{ij}^{w,l_s} \le f^{l_s} Cap_{l_s}, \forall l_s \in L', \forall i \in N.$$

$$\sum_{i\in N\setminus b}^{b-1} \sum_{j=i+1\in N}^{b} \sum_{w\in W} v_{ij}^{w,l_r} \le f^{l_r} Cap_{l_r}, \forall i \in N. \tag{10-i,ii}$$

$$y^{i,l_s} = \sum_{j>i} y_{ij}^{l_s} \le 1, \forall i \in N \setminus \{1,b\}, \forall l_s \in L'. \tag{11}$$

$$y^{1,l_s} = 1, y^{b,l_s} = 1, \forall l_s \in L'. \tag{12}$$

$$t_{ij}^{l_s} = T_{ij} + \sum_{h=i+1}^{h=j-1} y^{h,l_s} T^{h}, \forall l_s \in L'. \tag{13}$$

$$t_{ij}^{l_r} = T_{ij} + \sum_{h=i+1}^{h=j-1} T^{h}. \tag{14}$$

$$x_{ij}^{l_{ij}} = 1 \Leftrightarrow T_c t_{ij}^{l_{ij}} \le \frac{kW_c + T_c \sum_{h_{ij}\in L\neq l_{ij}} t_{ij}^{h_{ij}} f^{k_{ij}}}{\sum_{h_{ij}\in L\neq l_{ij}} f^{h_{ij}}}, \forall l_{ij} \in L. \tag{15}$$

$$x_{ij}^{l_s} \le y^{i,l_s} y^{j,l_s}, \forall l_s \in L'. \quad (16)$$

$$f^l \ge 0, \forall l \in L; V_{ij}^{w} \ge 0, \ \forall w \in W. \quad (17)$$

$$x_{ij}^{l} \in \{0,1\}, \forall l \in L.$$
$$y_{ij}^{l_s} \in \{0,1\}, \forall l_s \in L'. \quad (18)$$

In the objective function (5), the first term is the total ownership cost of the fleet size allocated to all services; the second term is the total operating cost of all services, proportional to the line service frequency; the third term is the waiting time cost for all passengers, the fourth term is the passenger travel time cost and the fifth term is the penalty for transfers. Constraint (6) states that the total fleet size of all operating lines is lesser than or equal to the available fleet size. Constraint (7) ensures that the fleet size assigned to each service can fulfil the service frequency requirement. It should be noted that the total travel time for the full cycle of the bus service $t^{l,cycle}$ for limited-stop service is determined by the optimal design of the subset of stops to be served while for normal service, $t^{l,cycle}$ is fixed, as we assume that the in-vehicle traveling time between bus stops and bus dwelling time at bus stops are both fixed and exogenously given. Indeed, $t^{l,cycle}$ can be defined by equations (13) and (14) by letting *i* and *j* represent the first and last node respectively.

Equation (8) computes the passenger flow on individual lines over a route section depending on whether they are attractive or not. Passenger flow is assigned to a line only if it is attractive over the route section (i.e., when $x_{ij}^{l} = 1$). Constraint (9) defines passenger demand conservation at all nodes of the corridor. Constraint (10) states that the capacity of a line is greater than or equal to its total passenger demand (in this constraint, *(i,j)* refer to any node pair in the transit corridor). Constraints (11), (12) describe the binary variables for transit line setting of the limited stop services as in conditions (1)-(2). Constraints (13) and (14) compute the travel time over route sections for the limited stop service and the normal service respectively. Constraint (15) is used to determine the attractive lines over a route section as in the common line problem explained in section 2.5.1. It is assumed that the passengers choose a subset of attractive lines with the minimum expected total travel time cost including both waiting time and in-vehicle travel time to travel towards their destination. Constraint (15) implies that, if a transit line is attractive over a route section, then the travel time cost on that line must be less than or equal to the combined waiting and travel time cost of all other lines serving the same route section and vice versa. A similar formulation can be found in Leiva et al. (2010). The constraint does not consider the case when the line is not attractive as passenger flow over a route section is assigned to only those lines which are found attractive and hence, waiting time and travel time costs are computed with respect to only those lines. Constraint (16) states that if the bus stops *i* and *j* are not served by the limited stop service, it is certain that the limited stop service will not be attractive over route section ij as the limited stop service between the two stops does not even exist. Hence, constraints (15) and (16) entail that a limited stop service is attractive over a route section only if it serves the route section and its travel time cost is lesser than the combined waiting and travelling times of all other lines that operate on the same route section. Constraint (17) defines the non-negativity condition of the operating frequency of each line and the route section flow respectively. Constraint (18) defines the binary variables. One can notice that the prominent characteristic of this model formulation is that, all the possible line settings for limited stop services are clearly incorporated in the model, which avoids the tedious and inefficient enumeration. In addition, the common line problem that determines the attractive lines in route sections is depicted in logic condition as in (15), which will later be cast into an equivalent set of linear conditions.

Hence, it should be noted that this model formulation is indeed a bi-level problem wherein the upper level problem is to design an optimal limited stop service operation strategy and the associated lower level problem is to find the attractive set of lines on each route section which describes the passengers' service choice behaviour and also accounts for the trip assignment process. The lower-level problem is essentially a common-line problem assuming passengers choose a set of attractive lines so that the expected total travel time is minimized.

2.5.3. Incorporating the effects of boarding and alighting into bus dwell time

In the above formulation, we assume that bus dwell time at bus stops is fixed and the congestion level is low. This would mean that either the demand is low or that the marginal impact of an extra passenger on the dwell time is low. But since limited stop services are implemented considering high demand, hence the above assumption holds only if the marginal impact of passengers boarding and alighting is negligible. This is suitable when vehicles have multiple doors, level boarding and not highly crowded on-board or on the platforms. In this subsection, the effects of boarding and alighting on bus dwell time are explicitly considered by assuming that the bus dwell time spent at stops is linearly dependent on the boarding or alighting passenger flow. Basically, the boarding time or alighting time, whichever is bigger, will be treated as the bus dwell time at stops. Here, we re-define the travel time on the limited stop service and the normal service respectively by incorporating the effects of alighting and boarding.

$$t_{ij}^{l_s} = T_{ij} + \sum_{w \in W} \sum_{h=i+1, h \in N}^{h=j-1, h \in N} y^{h,l_s} \max\left[\alpha \sum_{k<h, k \in N} (v_{kh}^{w,l_s}), \beta \sum_{h<k, k \in N} (v_{hk}^{w,l_s}) \right], \alpha > 0, \beta > 0 \quad (19)$$

$$t_{ij}^{l_r} = T_{ij} + \sum_{w \in W} \sum_{h=i+1, h \in N}^{h=j-1, h \in N} \max\left[\alpha \sum_{k<h, k \in N} (v_{kh}^{w,l_r}), \beta \sum_{h<k, k \in N} (v_{hk}^{w,l_r}) \right], \alpha > 0, \beta > 0 \quad (20)$$

In the above equations (19) and (20), the travel time for the two services between a node pair *(i,j)* takes into account the effects of alighting and boarding. As passengers travel between various O-D pairs w and different route sections ij, at each node *h*, the greater of the boarding time and alighting time would determine the dwell time at *h*. The positive coefficients α and β describe the alighting time and boarding time per passenger respectively.

It should be noted that incorporating the boarding and alighting effects would lead to the common-line equilibrium problem (e.g., Cominetti and Correa, 2001; Larrain and Munoz, 2008), which will make the problem significantly more complicated. In this study, the equilibrium common lines are not considered by assuming low level of congestion. Capturing the boarding and alighting effects, as well as the congested common-line equilibrium, could be addressed in the future study.

3. Solution Method

The model formulation in the previous section is indeed an MINLP. Due to the inherent nonlinear and nonconvex property, it is very hard to solve the MINLP. In this study, we seek to obtain a global optimal solution of the problem rather than only a local optimal solution. To achieve so, we first transform the nonlinear terms into linear ones by applying various linearization techniques so that the original MINLP can be transformed into a mixed integer linear program (MILP). Then, many existing solution algorithms like the branch and bound method can be used to solve the MILP which can guarantee a global optimal solution.

One can notice that the nonlinearity of the model formulation arises from the objective function (5) and the nonlinear constraints (7), (8), (10)-(i), (15), (16), (19) and (20). Constraint (7) is converted into equivalent linear conditions by using a Reformulation Linearization Technique (RLT) as demonstrated in section 3.1, constraint (8) is linearized by using the multidimensional piecewise linearization method mentioned in section 3.3, constraint (10)-(i) is linearized using a combination of linear transformation for product of binary variables and RLT as shown in section 3.2, constraint (15) is replaced with an equivalent set of linear conditions by using a linear inequality constraint as mentioned in section (3.4), constraint (16) is handled in section 3.2 and finally, constraints (19)-(20) are also treated by using the linear inequality method shown in section 3.5. The objective function (5) is linearized and subsequently reformulated using a combination of the various linearization techniques used to linearize the constraints.

3.1. Linearization with RLT

Constraint (7) is linear when considering the normal service as the cycle time for the normal service is known, but it is nonlinear for the limited stop service due to the product of a continuous variable (service frequency) and the travel time which includes binary variable representing the limited stop service as shown in constraint (13). Hence, a RLT as introduced in Sherali and Alameddine (1992) is used to represent this bilinear term through an equivalent set of linear conditions. Denote u_a as binary and $\tilde{x}_a$ as continuous such that $\underline{x}_a \le \tilde{x}_a \le \overline{x}_a$, where $\underline{x}_a$ and $\overline{x}_a$ are a sufficiently small positive number and a sufficiently large upper bound on $\tilde{x}_a$, respectively. If $x_a = u_a \tilde{x}_a$, the equivalent linear transformation of the bilinear term can be expressed as:

$$\begin{cases} x_a - u_a \underline{x}_a \ge 0 \\ x_a - u_a \overline{x}_a \le 0 \\ x_a - \tilde{x}_a + \underline{x}_a - u_a \underline{x}_a \le 0 \\ x_a - \tilde{x}_a + \overline{x}_a - u_a \overline{x}_a \ge 0. \end{cases} \tag{21}$$

Substituting equation (13) into constraint (7) for the case of limited stop services, the following can be obtained:

$$t^{l_s,cycle} f^{l_s} \le n_{l_s} \Rightarrow (T_{ij} + \sum_{h=2}^{h=b-1} y^{h,l_s} T_s^h) f^{l_s} \le n_{l_s} \tag{22}$$

Here, the nonlinearity of (22) arises from the product of service frequency variable f^{l_s} and binary variable y^{h,l_s} for the limited stop service. Further substitution can be done as follows to represent this bilinear term $y^{h,l_s} f^{l_s}$:

$$g^{h,l_s} = y^{h,l_s} f^{l_s} \tag{23}$$

Using (23) in constraint (22), the resultant expression becomes:

$$T_{ij} f^{l_s} + \sum_{h=2}^{h=b-1} g^{h,l_s} T_s^h \le n_{l_s} \tag{24}$$

Here, the bilinear term g^{h,l_s} as the product of service frequency variable f^{l_s} and binary variable y^{h,l_s} can be transformed into equivalent linear conditions as in (21). That is to say, the nonlinear constraint (7) is now completely converted into equivalent linear constraints by applying the RLT method. Indeed, the same RLT method can be used to linearize the nonlinear travel time function as given in (19) if effects of boarding and alighting need to be considered.

If the operating cost per cycle F^l is defined as a linear function of the cycle time as in (13) and (14), then the total operating cost term in the objective function will be nonlinear for the limited stop service involving the same nonlinear term of $t^{l_s,cycle} f^{l_s}$ as in (22), which can be easily linearized by using the RLT method.

3.2. Linear transformation to linearize a product of binary variables

For nonlinearity in (16) involving the product of two binary variables, the following equivalent linear transformation can be applied for its linearization. Without loss of generality, if *A*=*cd* where *c* and *d* are binary, this bilinear term can be expressed equivalently by the following linear constraints which can be easily verified by enumerating all the possible cases as *c* and *d* are binary variables.

$$\begin{cases} A \le c \\ A \le d \\ A \ge c + d - 1. \end{cases} \tag{25}$$

The nonlinear capacity constraint (10)-(i) for the limited stop service can be linearized by first using (25) to linearize the product of the binary variables y^{i,l_s}, y^{j,l_s} and then using the RLT in (21) to linearize the product of binary variables and continuous variables.

3.3. Multidimensional piecewise linearization method

This section deals with the nonlinearity in (8) and the objective function. To handle the nonlinearity in constraint (8), a multidimensional piecewise linear approximation technique as proposed by Misener and Floudas (2009) is adopted. The basic idea of this piecewise linear approximation technique to first partition the feasible domain (for this study, the nonlinear term has two variables) into a number of small rectangles and then use the linear convex combination to approximate the two-variable function within the small rectangles. The solution of the optimization with the piecewise linearization constraints will find the rectangle within which the optimal solution lies. This rectangle is referred to as the active rectangle. In this subsection, the constraint (8) is nonlinear in two dimensions and once it is linearized by using this technique, the third term in the objective function $\sum_{w \in W} \sum_{ij \in S_{ij}} V_{ij}{}^{w} \frac{W_c k}{\sum_{l \in L} x_{ij}^{l} f^{l}}$ can be reformulated into a linear form and the travel time cost term $\sum_{w \in W} \sum_{ij \in S_{ij}} \sum_{l \in L} T_c v_{ij}{}^{w,l} t_{ij}{}^{l}$ in the objective function can be further linearized by using the RLT as shown in (21).

Considering the objective function :

$$\min Z = \sum_{l \in L} K_l n_l + \sum_{l \in L} F^l f^l + \sum_{w \in W} \sum_{ij \in S_{ij}} V_{ij}{}^{w} \frac{W_c k}{\sum_{l \in L} x_{ij}^{l} f^{l}} + \sum_{w \in W} \sum_{ij \in S_{ij}} \sum_{l \in L} T_c v_{ij}{}^{w,l} t_{ij}{}^{l} + \lambda_{trans} \sum_{w \in W} (\sum_{ij \in S_{ij}} V_{ij}^{w} - X^{w}).$$

This can be reformulated as:

$$\min Z = \sum_{l \in L} K_l n_l + \sum_{l \in L} F^l f^l + \sum_{w \in W} \sum_{ij \in S_{ij}} V_{ij}{}^{w} \frac{W_c k}{q_{ij}} + \sum_{w \in W} \sum_{ij \in S_{ij}} \sum_{l \in L} T_c v_{ij}{}^{w,l} t_{ij}{}^{l} + \lambda_{trans} \sum_{w \in W} (\sum_{ij \in S_{ij}} V_{ij}^{w} - X^{w}). \tag{26}$$

In (26), we introduce new terms $q_{ij}{}^{l}$ and q_{ij} to represent:

$$q_{ij}{}^{l} = x_{ij}^{l} f^{l}, \text{ and} \tag{27}$$

$$q_{ij} = \sum_{l \in L} q_{ij}{}^{l} = \sum_{l \in L} x_{ij}^{l} f^{l}. \tag{28}$$

Equation (27) and (28) can be linearized using the RLT approach as shown above in the previous sub-section. Based on (27) and (28), constraint (8) can be written as:

$$v_{ij}^{w,l} = V_{ij}^{w} \frac{q_{ij}^{l}}{q_{ij}}, \forall l \in L, \forall w \in W. \quad (29)$$

The constraint (29) is nonlinear as the right-hand side of this equation involves multiplication and division of multiple variables. Now, we will apply the multi-dimensional piecewise linearization method as proposed by Misener and Floudas (2009) to linearize the reformulated constraint (29). Consider the two variables, i.e., V_{ij}^{w}, passenger flow over route section ij for an O-D pair w, and q_{ij}^{l}, route section line frequency for line *l*, both of which fall into bounded intervals partitioned into M_1 and N_1 smaller segments. These intervals can be explicitly stated as below respectively:

$$V_{ij}^{w} \in [V_{ij}^{w,m-1}, V_{ij}^{w,m}], m = 1,\ldots,M_1; q_{ij}^{l} \in [q_{ij}^{l,n-1}, q_{ij}^{l,n}], n = 1,\ldots,N_1.$$

Also, the bounds of the feasible domain for the two variables are explicitly shown as follows:

$$0 \le V_{ij}^{w} \le V_{ij}^{w,M_1}, \forall w \in W; \quad 0 \le q_{ij}^{l} \le q_{ij}^{l,N_1}, l \in L. \quad (30)$$

The segments given by $[V_{ij}^{w,m-1}, V_{ij}^{w,m}], [q_{ij}^{l,n-1}, q_{ij}^{l,n}]$ are not necessarily equal in size. If M_1 and N_1 are sufficiently large such that the distance between any two consecutive points of each segment is very small, the true values of the following functions can be closely approximated by using piecewise linear functions. In this study, we can take the lower bounds of both the variables to be zero and the upper bound values to be equal to the total exogenous demand and maximum allowable route section frequency value, respectively. Now, consider the following two nonlinear functions:

$$C_{ij}^{w} = \frac{V_{ij}^{w}}{q_{ij}}, \forall w \in W, \quad (31)$$

$$\overline{C_{ij}^{l,w}} = \frac{V_{ij}^{w}}{q_{ij}} q_{ij}^{l}, \forall l \in L, \forall w \in W, \quad (32)$$

The feasible domains of functions (31) and (32) cover the bounded intervals of the variables and are divided into $M_1 \times N_1$ rectangles. Each corner point *(m,n)* of these rectangles is associated with a particular value of variables in (31) and (32) which is explicitly computed by (34) and (35) below. Consider two sets of SOS1 variables (special ordered set of type 1 of which at most one variable is strictly positive whereas all others are at zero) i.e., S1 and S2 proposed by Beale and Tomlin (1970), to determine the active rectangle where the optimal values of the two variables are located.

$$S1: \mu_{ij}^{m} \in [0,1], m = 1,\ldots,M_1, \quad S2: \nu_{ij}^{n} \in [0,1], n = 1,\ldots,N_1. \quad (33)$$

Each candidate rectangle has four corner points, denoted by a set of co-ordinates *(m,n)*. Then, the functions (31) and (32) can be expressed as :

$$C_{ij}^{w,(m,n)} = \frac{V_{ij}^{w,m}}{q_{ij}^{n}}, \quad (34)$$

$$\overline{C_{ij}^{l,w(m,n)}} = V_{ij}^{w,m} \frac{q_{ij}^{l,n}}{q_{ij}^{n}}. \quad (35)$$

A convex combination of these points is used to determine the value of the two functions within that rectangle. Denoting the coefficient of convex combination ranging between 0 and 1 by:

$$\rho_{ij}^{m,n} : \rho_{ij}^{m,n} \in [0,1], m = 0,\ldots,M_1, n = 0,\ldots,N_1; \sum_{m=0}^{M_1}\sum_{n=0}^{N_1} \rho_{ij}^{m,n} = 1. \quad (36)$$

Hence, with the above described, the following equations are used to conduct the two-dimensional piecewise linearization:

$$C_{ij}^{l,w} = \sum_{m=0}^{M_1}\sum_{n=0}^{N_1} \rho_{ij}^{\ m,n} C_{ij}^{l,w(m,n)}, \forall l \in L, \forall w \in W, \tag{37}$$

$$\overline{C_{ij}^{l,w}} = \sum_{m=0}^{M_1}\sum_{n=0}^{N_1} \rho_{ij}^{\ m,n} \overline{C_{ij}^{l,w(m,n)}}, \forall l \in L, \forall w \in W, \tag{38}$$

$$V_{ij}^{\ w} = \sum_{m=0}^{M_1}\sum_{n=0}^{N_1} \rho_{ij}^{\ m,n} V_{ij}^{\ w,m}, \forall w \in W, \tag{39}$$

$$q_{ij}^{\ l} = \sum_{m=0}^{M_1}\sum_{n=0}^{N_1} \rho_{ij}^{\ m,n} q_{ij}^{\ l,n}, \forall l \in L, \tag{40}$$

$$q_{ij}^{\ n} = \sum_{l \in L} q_{ij}^{\ l,n}, n = 0,..., N_1, \tag{41}$$

$$\sum_{m=0}^{M_1}\sum_{n=0}^{N_1} \rho_{ij}^{\ m,n} = 1, \tag{42}$$

$$\rho_{ij}^{\ m,n} \in [0,1], m = 1,..., M_1; n = 1,..., N_1, \tag{43}$$

$$\sum_{n=0}^{N_1} \rho_{ij}^{\ 0,n} \le \mu_{ij}^{\ 1}, \tag{44}$$

$$\sum_{n=0}^{N_1} \rho_{ij}^{\ m,n} \le \mu_{ij}^{\ m} + \mu_{ij}^{\ m+1}, m = 1,.., M_1 - 1, \tag{45}$$

$$\sum_{n=0}^{N_1} \rho_{ij}^{\ M_1,n} \le \mu_{ij}^{\ M_1}, \mu_{ij}^{\ m} \in [0,1], m = 1,.., M_1, \tag{46}$$

$$\sum_{m=0}^{M_1} \rho_{ij}^{\ m,0} \le \nu_{ij}^{\ 1}, \tag{47}$$

$$\sum_{m=0}^{M_1} \rho_{ij}^{\ m,n} \le \nu_{ij}^{\ n} + \nu_{ij}^{\ n+1}, n = 1,.., N_1 - 1, \tag{48}$$

$$\sum_{m=0}^{M_1} \rho_{ij}^{\ m,N_1} \le \nu_{ij}^{\ N_1}, \nu_{ij}^{\ n} \in [0,1], n = 1,.., N_1. \tag{49}$$

In equations (37)-(40), the value of each of the variables $C_{ij}^{l,w}$, $\overline{C_{ij}^{l,w}}$, $V_{ij}^{\ w}$, $q_{ij}^{\ l}$ is computed as the convex combination of its values at the corner points of each of the $M_1 \times N_1$ rectangles that the domain was originally partitioned into. Equation (41) represents the combined frequency of all attractive lines serving a particular route section. Equations (44)-(49) lay the conditions of interdependence between the SOS1 variables ($\mu_{ij}^{\ m}, \nu_{ij}^{\ n}$) and the coefficient for convex combination $\rho_{ij}^{\ m,n}$. Hence, the two nonlinear functions in (31) and (32) are linearized with this multidimensional piecewise linearization approach. In this way, the nonlinear constraint (8) is converted into linear constraints; the nonlinear terms in the reformulated objective function (26) are also linearized.

The fourth term of the reformulated objective function (26) involves the product of the service passenger flow and the travel time. Using (32), this term can be expressed as:

$$\chi_{ij}^{l,w} = \overline{C_{ij}^{l,w}} t_{ij}^{\ l} \tag{50}$$

Since the travel time for the normal service is exogenously known, the product of passenger flow over normal service and travel time is linear. We consider the product of the passenger flow and travel time of the limited stop service over a route section ij as below:

$$\chi_{ij}^{l_s,w} = \overline{C_{ij}^{l_s,w}} t_{ij}^{l_s} = \overline{C_{ij}^{l_s,w}} (T_{ij} + \sum_{h=i+1 \in N}^{h=j-1} y_p^{h,l_s} T_s^{\ h}). \tag{51}$$

In the above equation, the product of the passenger flow on line *l* and binary variable is nonlinear, but can be easily represented by equivalent linear conditions as is shown in (21).

3.4. Treatment for logic constraints

Further, constraint (15), describing the common line problem, is expressed in the form of logic constraints, which cannot be tackled directly in a mathematical programming problem. Constraint (15) can also be written as:

$$x_{ij}^l = 1 \Leftrightarrow T_c t_{ij}^{\ l} \sum_{k \in L, k \neq l} f^{\ k} - (kW_c + T_c \sum_{k \in L, k \neq l} t_{ij}^{\ k} f^{\ k}) \leq 0, \forall l \in L. \tag{52}$$

This logical condition can be transformed into the following inequality constraint:

$$Lx_{ij}^l \leq (T_c t_{ij}^{\ l} \sum_{k \in L \neq l} f^{\ k}) - (kW_c + T_c \sum_{k \in L \neq l} t_{ij}^{\ k} f^{\ k}) \leq U(1 - x_{ij}^l) \tag{53}$$

In constraint (53), *L* and *U* are sufficiently large negative and positive constants, respectively. As per this constraint, when the binary variable x_{ij}^l is equal to 1, the left-hand side inequality is always true (as *L* is a sufficiently large negative constant) while only the right-hand side inequality is effective, entailing the condition for attractiveness of line *l* as in (52). Conversely, when the right-hand side inequality holds, to ensure constraint (53) is true, the binary variable has to be equal to 1. In another case, i.e., when the binary variable x_{ij}^l is equal to 0, the right-hand side inequality is always true (as *U* is a sufficiently large positive constant) while only the left-hand side inequality is effective, indicating the line *l* is not attractive in accordance with constraint (52). Hence, one can observe that constraint (53) is the equivalent reformulated inequality for constraint (15). The middle term in (53) can be easily transformed into equivalent linear constraints by applying the RLT approach as shown in (21). By doing so, the logic constraint (15) can be cast into an equivalent set of linear inequality constraints.

3.5. Treatment for the travel time functions incorporating the effects of alighting and boarding

We now consider the travel time functions for the limited stop service and the normal service when incorporating the effects of alighting and boarding as given in (19) and (20). Let $\alpha \sum_{k<h, k \in N} (v_{kh}^{w,l_s}) = a$ and $\beta \sum_{h<k, k \in N} (v_{hk}^{w,l_r}) = b$ respectively. In the travel time functions (19) and (20), the "max" function implies $\max[a,b] = a$ if $a > b$ and $\max[a,b] = b$ if $a < b$. Let us consider a binary variable $\omega \in \{0,1\}$ such that $\omega = 1$ when $a > b$ and $\omega = 0$ when $a < b$. This can be represented by the following mathematical expressions:

$$\max[a,b] = \omega a + (1-\omega) b \tag{54}$$

$$U\omega + a(1-\omega) < b < a\omega + L(1-\omega) \tag{55}$$

where U and L are very large negative and positive integers respectively. Now, when ω takes the value of 1 in (55), then $U < b < a$ and the "max" function returns 'a' as the solution in (54). Similarly, when ω takes the value of zero, then $a < b < L$ and the "max" function returns 'b' in (54). The nonlinearity in (54) and (55) can be linearized using the RLT as shown in (21).

3.6. Reformulated problem

Hence, the original model formulation of this limited stop service design problem has been completely transformed into a MILP, in which the objective function and all the constraints are linear.

Specifically, the reformulated model can be expressed as:

$$\min Z = \sum_{l\in L} K_l n_l + \sum_{l\in L} F^l f^l + \sum_{w\in W}\sum_{ij\in S_{ij}} kW_c C_{ij}^w + \sum_{w\in W}\sum_{ij\in S_{ij}}\sum_{l\in L} T_c \chi_{ij}^{l,w} + \lambda_{trans}\sum_{w\in W}\left(\sum_{ij\in S_{ij}} V_{ij}^w - X^w\right) \tag{56}$$

which is subject to constraints (6), (7), (9)-(14), (17)-(18), (24), (27)-(55).

As a result, the original model formulation is transformed into a MILP which can be solved by using many efficient solution algorithms like the branch and bound method. Most importantly, the solution property of global optimality of the MILP is guaranteed.

3.7. Additional operational constraints

For the sake of operational efficiency, a few more constraints could be added to the presented model formulation. At times, the operator could face a requirement of optimally selecting up to a fixed number of 'special' nodes (say P) for each operating limited stop service. Also, in case of multiple limited stop services, the operator might decide that a maximum of one limited stop service serves a particular 'special' node such that the benefit of travelling over limited stop service is equitable over all the nodes in the transit corridor. Therefore, the following constraints could be added:

$$\sum_i y^{i,l_s} \le P, \forall l_s \in L' \tag{57}$$

$$\sum_{l_s} y^{i,l_s} \le 1, \forall i \in N \tag{58}$$

Constraint (57) states that for each limited stop service, there is a maximum limit of P number of 'special' nodes that need to be optimally selected. Constraint (58) states that a maximum of one limited stop service can serve any intermediate node in the transit corridor. It should be noted that adding constraints (57) and (58) limits the size of the problem by reducing the search zone and the global optimal solution is faster to achieve, however, as mentioned earlier, the consideration of these constraints is upon the discretion of the operator.

4. Numerical studies

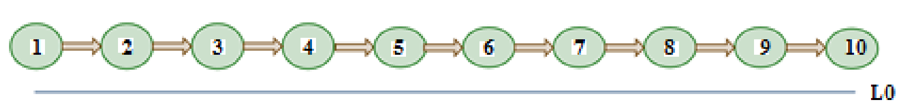

Fig. 4 Transit corridor for numerical study

Consider a corridor of 10 nodes as in Figure 4. From the example of the Singapore bus transit, two kinds of bus services are considered, i.e., normal bus service l_0 similar to service 179 (operating as a loop service from Boon Lay Interchange to NTU, Singapore campus and then terminating at the interchange) and limited stop services (l_1 and l_2) similar to service 179A(operating as limited stop service on the 179 route). We assume that the line setting of the limited stop service, i.e., which stops to be served is upon the discretion of the operator. As stated before, the problem is to determine the optimal line setting for limited stop services l_1 and l_2, as well as the operating frequencies and fleet size of all the transit lines while minimizing the total costs. We assume that nodes 6 and 10 are major attractors and hence, exogenous demand for these nodes is considered as given parameters. Also, in this study, a single corridor demand in the direction from node 1 towards node 10 is considered. .

4.1. General Parameters

1) Bus capacities (number/bus): l_0 = 60 passengers/bus, l_1 =60 passengers/bus, l_2 = 60 passengers/bus.

2) Intrinsic demand D at nodes for the destination node 10:

Table 3 Intrinsic demand D at each node for destination node 6 per hour

Node	1	2	3	4	5	6	7	8	9	10
Demand	75	65	40	25	15	0	0	0	0	0

Table 4 Intrinsic demand D at each node for destination node 10 per hour

Node	1	2	3	4	5	6	7	8	9	10
Demand	60	40	20	20	15	30	35	40	35	0

3) Standard running time (minutes): For $j>i$ where (i, j) represent a node pair, $T_{ij} = 2(j-i)$

(Assumed to vary linearly with distance and equal inter-node spacing)

4) Operating cost per operational hour: l_0 = 70\$/bus, l_1=50\$/bus, l_2 = 60\$/bus.

5) Ownership cost per operational hour: l_0 =40\$/bus, l_1 = 40\$/bus, l_2 = 40\$/bus.

6) Standard dwelling time at each node T^h (mins.): 1 min

7) Total available fleet size= 20 buses

8) Assuming Poisson's arrivals, $k=1$, $\lambda_{trans} = 5\$$, W_c = 0.25\$/min, T_c = 0.25\$/min.

9) Number of nodes in the transit corridor *b* = 10.

4.2. Optimization Results

The model was evaluated by using the solver Gurobi on the programming platform YALMIP (Löfberg 2004) interfaced with MATLAB on a Precision T1650 Dell PC, with a 3.20 GHz processor, 16 GB RAM, and a 64-bit operating system. In the numerical study, we first consider a transit corridor with one normal service and one limited stop service followed by another example with one normal service and two limited stop services. The results are then analysed and inferences are discussed. For the generic case, we consider the constraint (58) and do not include constraint (57) in the numerical study.

4.2.1. Numerical example with one limited stop service

In this numerical example, we consider only one single limited stop service l_1 operating in conjugation with the normal service l_0. Using the input parameters from Section 4.1., the optimization model is solved.

The optimal solution of this numerical example is listed as follows:

1) Optimal limited stop service pattern: 1->2->3->4->10

2) Service frequency: l_1 = 5 buses/hr, l_0 = 10 buses/hr

3) Fleet size: l_1 = 3 buses, l_0 = 6 buses

4) Total operating cost: 2992$

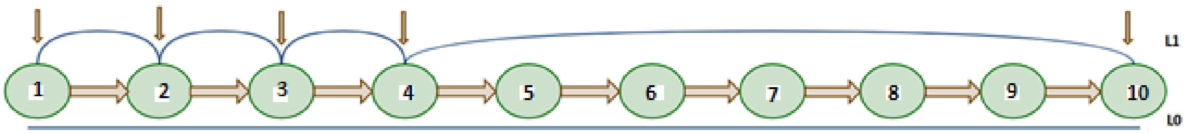

Fig. 5 Optimized line setting for the limited stop service l_1

4.2.2. Numerical example with two limited stop services

In this numerical example, we assume that exactly two limited stop bus services l_1 and l_2 are provided along with the normal service l_0. The detailed model solutions are listed as below:

1) Optimal flow pattern:
l_1: 1->7>8->10; l_2: 1->2->3->5->10.

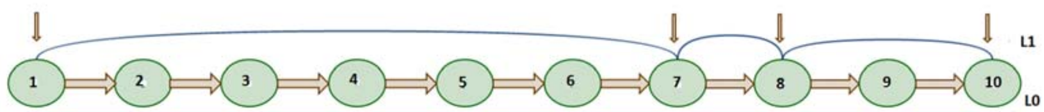

Fig.6 Optimal line setting for the limited stop service l_1

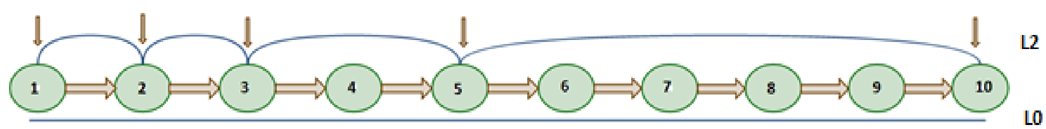

Fig. 7 Optimal line setting for the limited stop service l_2

2) Optimal service frequency for the transit lines in buses/hr: l_1 = 5 buses/hr, l_2 = 6 buses/hr, l_0 = 5 buses/hr.

3) Fleet size**:** l_1 = 3 buses, l_2 = 3 buses, l_0 = 3 buses.

4) Total cost: 3431$.

Using the model formulation and solution algorithm proposed in this study, one can further understand how the problem parameters like travel demand would affect the optimal design of limited stop services. Considering different levels of travel demand, we can obtain the following results as sensitivity analysis with respect to model parameters:

Table 5 Sensitivity analysis with respect to travel demand

Demand	Optimal total	Line settings for	Service	Fleet size for

w.r.t. original demand	operating costs	l_1, l_2	frequency for $l_1, l_2,\ l_0$	$l_1, l_2,\ l_0$
50%	$1845	1->7->8->10; 1->2->3->10	2 buses/hr, 3 buses/hr, 2 buses/hr	2 buses, 2 buses, 4 buses
80%	$2547	1->7->8->10; 1->2->3->5->10	4 buses/hr, 5 buses/hr, 4 buses/hr	3 buses, 3 buses, 4 buses
130%	$3596	1->6->7->8->10; 1->2->3->5->10	7 buses/hr, 8 buses/hr, 7 buses/hr	4 buses, 4 buses, 5 buses
150%	$4019	1->6->7->8->10; 1->2->3->5->10	8 buses/hr, 9 buses/hr, 8 buses/hr	5 buses, 5 buses, 5 buses

4.2.3. Numerical example with only one normal service and no limited stop service

In this section, we assume that only one normal service operates on the travel corridor. In this case, the transfers need not be considered and all passengers can reach their destination in the same normal service they boarded at the origin node. Since only one normal service would be used without any transfers, the waiting time term in the objective function would consider the overall service frequency for the aggregated demand and not for different route sections. Hence, as we can realise, there would be no choice available to passengers in this case as there exists only one operating service. Upon computation, we observed the following results:

Total operating cost: $3318; Fleet size: l_0 = 5 buses; Service frequency: l_0 = 9 buses/hr.

For illustration purpose, a table with the corresponding value of each term in the objective function for all the above three numerical cases has been given below. One can observe that, for the case when no limited stop service is provided, the travel time and waiting time cost are much higher due to the lack of choice of limited stop service, despite the transfer cost is removed.

Table 6 Case based analysis of objective function and corresponding computation time

	Numerical Case	Total($)	Term 1($)	Term 2($)	Term 3($)	Term 4($)	Term 5($)
1	4.2.1	2992	360	950	344	828	510
2	4.2.2	3431	360	960	635	756	720
3	4.2.3	3318	200	630	850	1638	-

4.3. Further discussions

4.3.1. Comparison with enumeration approach

The proposed model formulation and its solution algorithm employing convexification and linearization techniques seeks to find the global optimal solution of the problem of limited stop bus service design. In previous studies in the literature, without aid of the model formulation and the global optimal solution method, one has to resort to enumeration approach if global optimal solution of this problem is required. However, it is computationally prohibitive to enumerate all the possible line settings when problem size is large.

In this study, additional constraints (57) and (58) could be added into the model formulation for certain practical considerations. It should be noted, when such constraints are not included in the model formulation, the number of candidate line settings is huge and to find the global optimal solution using enumeration approach would be very computationally intensive.

For example, if we use the enumeration method, the number of possible transit line settings when only one limited stop service line is considered in the numerical example with 10 bus stop nodes would be around 300. Indeed, in our numerical test, we assume two limited stop service lines are to be constructed and constraint (58) is imposed. In this case, by enumeration approach, there will be around 7,000 possible combinations of bus service line settings. One can imagine that, when problem size goes further up, the computational load would be prohibitively huge if enumeration method is applied. Specifically, in our numerical experiments conducted in subsections 4.2.1, the computational time is 236 seconds. Meanwhile, we tried to use enumeration approach to solve this problem and the computational time is around 300 seconds.

4.3.2. Comment on solution quality

To further investigate the quality of the solution obtained by the proposed methodology, we conduct more numerical experiments by assuming that the candidate transit line settings are predetermined, as was done in previous research in the literature. In this case, in order to make the analysis tractable, we consider a subset of the possible combinations of transit line settings by invoking the constraints (57) and (58). Then, we compare the solutions with the global optimal solution obtained from our method. It should be noted that the number of possible transit line settings could be prohibitively large, and therefore considering only a small set of candidate lines might not be sufficient to obtain the global optimal solution, while enumerating all the possible candidate lines would be computationally tedious and inefficient. Let us consider the following two cases:

(i) For the case of single limited stop service operating with normal service: If the additional operational constraints (57) and (58) are considered with a total of 2 allowable stoppages, the number of possible transit line setting for the limited stop service l_1 is 28 (e.g., 1-2-3-10, 1-2-4-10, 1-2-5-10 and so on). One can expect that, if the additional constraint is not imposed (i.e., the number of allowable stoppages is not restricted), a much larger number of possible transit line settings would exist as shown in the previous sub-section.
(ii) For the case of two limited stop services operating with the normal service: If the additional operational constraints (57) and (58) are considered, the possible combinations of line settings for l_1 and l_2 are totally 420. (e.g., {(1-2-3-10),(1-4-5-10)},{(1-2-3-10),(1-4-6-10)}, and so on).

It should be noted that the number of candidate line settings is large even when we consider the additional operational constraints. This number increases rapidly if more limited stop transit lines are to be added and the additional constraints are completely relaxed as shown in the previous sub-section. Hence, one can envisage that, with just a few predetermined set of transit line settings, it is not possible to find the best solution as the number of such candidate settings could be prohibitively huge.

Now, we consider the numerical study in which a single limited stop service operates with the normal service. We also consider the additional constraint (57) with a maximum of two allowable stoppages for the line l_1; hence the total operating costs of all 28 candidate transit line settings are computed.

Optimal total operating cost Z= 2965$; Optimal transit line setting for l_1: 1->3->4->10.

Table 7 Total operating cost for corresponding candidate transit line setting of the limited stop service l_1

Solutions closest to optimal solution with corresponding transit line setting for l_1	Intermediate solutions with corresponding transit line setting for l_1	Solutions farthest from optimal solution with corresponding transit line setting for l_1
3123$; 1->2->4->10	6136$; 1->4->5->10	8269$; 1->6->8->10
3735$; 1->2->4->10	6431$; 1->3->9->10	7687$; 1->5->8->10
3943$; 1->4->6->10	6494$; 1->3->8->10	6513$; 1->4->8->10

In Table 7, we demonstrate the optimal objective solutions resulting from some given transit line settings. Three groups of representative results are shown: specifically, three line settings leading to objectives solutions that are closest to the global optimal solution obtained from our method, as well as other groups of settings incurring medium and worst objective solutions as compared to the global optimal solution are listed. One can find that the gap between the global optimal solution and the second best one is 5.32% ($=\frac{3123-2965}{2965}$) and that with the third best one is 25.91%. However, the worst-case result (with objective solution of 8269) is much worse than the global optimal solution. This underlines our rationale that only considering the selected set of candidate line settings will not guarantee the global solution; on the contrary, it is possible that the design of transit service is much worse than the global optimal solution if a wrong set of candidate line settings is predetermined.

5. Concluding remarks

This study formulates a bus service design problem in which limited stop services are provided in conjugation with the normal service. From the transit operators' perspective, this model formulation proposes a methodology to determine optimal operation strategies in terms of service fleet size assignment, service frequencies, and the line setting for the limited services, i.e., which stops are served by the limited stop services. The common line approach is applied to determine the passenger assignment on the transit corridor. Mathematically, the model presented is a mixed integer nonlinear program. A solution method is developed to firstly transform the MINLP into an MILP with various linearization techniques and then solve the approximated mixed integer linear problem to attain a global optimal solution. Unlike most of the other published works which select the best transit line setting for the limited stop service from a predetermined set of candidate line settings, this study presents a methodology for an explicit design of the limited stop service while considering various operational constraints.

However, one limitation of the model lies in its incapability to handle very large size transit corridors which could be addressed in future studies. This research work contributes to the literature by building up an analytical framework and developing a global optimal solution method to solve the problem of optimal bus service design with limited stop service. In many major cities, operating differential services is believed to be a viable strategy to improve public transit service quality and reduce the car dependency. It is imperative for the public transit service operators to answer the practical question on how to determine the optimal operation strategies when differential services are offered. Rather than manually predetermining certain bus service line settings as is done in the current practice of the bus industry, this study provides a mathematical modelling approach to assist in optimal decision making for transit operation, aiming to achieve financially and socially sustainable bus transit services. The model formulation in this study avoids the tedious and inefficient enumeration of all possible bus line settings. The solution method for the model formulation as proposed in this study guarantees the global optimality of the solution quality, which ensures the best solution for the transit operation strategies. Indeed, the proposed methodology in this study can be readily applied to real-life transit operations. Besides, the methodology presented in this study can also be extended to solve similar complex system design problems in maritime studies, air-transportation, and various other engineering domains.

References:

Afanasiev, L.L., Liberman, S.Y., 1983. Principles for organizing express bus services. Transportation Research Part A 17(5), 343-346.

Beale, E.M.L., Tomlin, J.A., 1970. Special facilities in a general mathematical programming system for non-convex problems using ordered sets of variables. Proceedings of the Fifth International Conference on Operation Research, 447-454.

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

Cepeda, M., Cominetti, R., Florian, C., 2006. A frequency-based assignment model for congested transit networks with strict capacity constraints: Characterization and computation of equilibria. Transportation Research Part B 40(6), 437-459.

Chien, S.I., Qin, Z., 2004. Optimization of bus stop locations for improving transit accessibility. Transportation Planning and Technology 27(3), 211-227.

Chiraphadanakul, V., Barnhart, C., 2013. Incremental bus service design: combining limited-stop and local bus services. Public Transport 5, 53-78.

Chriqui, C., Robillard, P., 1975. Common bus lines. Transportation Science 9(2), 115-121.

Cominetti, R., Correa J., 2001. Common-Lines and Passenger Assignment in Congested Transit Networks. Transportation Science 35(3), 250-267.

Conlon, M., Foote, P., O' Malley, K., Stuart, D., 2001. Successful arterial street limited-stop express bus service in Chicago. Journal of the Transportation Research Board 1760, 74-80.

Cortés, C.E., Jara-Díaz, S., Tirachini, A., 2011. Integrating short turning and deadheading in the optimization of transit services. Transportation Research Part A 45(5), 419-434.

Curtin, K.M., Biba, S., 2011. The Transit Route Arc-Node Service Maximization problem. European Journal of Operational Research 208(1), 46-56.

De Cea, J., Fernandez, E., 1993. Transit assignment for congested public transport systems. Transportation Science 27(2), 133-147.

El-Geneidy, A., Surprenant-Legault, J., 2010. Limited-stop bus service: an evaluation of an implementation strategy. Public Transport 2(4), 291-306.

Hart, N., 2016. Methodology for evaluating potential for limited-stop bus service along existing local bus corridors. Transportation Research Record, 2543, 91-100.

Kikuchi, S., 1985. Relationship between the number of stops and headway for a fixed-route transit system. Transportation Research Part A 19(1), 65-71.

Larrain, H., Giesen R., Muñoz, J.C., 2010. Choosing the right express services for bus corridor with capacity restrictions.Transportation Research Record: Journal of the Transportation Research Board 2197, 63-70.

Larrain, H., Muñoz, J.C., 2008. Public transit corridor assignment assuming congestion due to passenger boarding and alighting. Networks and Spatial Economics 8(2), 241-256.

Larrain, H., Muñoz, J.C., 2016. When and where are limited-stop bus services justified? Transportmetrica A: Transport Science, 12(9), 811-831.

Larrain, H., Muñoz, J.C., Giesen, R., 2015. Generation and design heuristics for zonal express services. Transportation Research Part E: Logistics and Transportation Review 79, 201–212.

Leiva, C., Muñoz, J.C., Giesen, R., Larrain, H., 2010. Design of limited-stop services for an urban bus corridor with capacity constraints. Transportation Research Part B 44(10), 1186-1201.

Li, Z.C., Lam, W.H.K., Wong, S.C., Sumalee, A., 2011. Design of a rail transit line for profit maximization in a linear transportation corridor. Procedia - Social and Behavioral Sciences 17, 82-112.

Li, Z.C., Lam, W.H.K., Wong, S.C., 2010. Optimization of number of operators and allocation of new lines in an oligopolistic transit market. Networks and Spatial Economics 12(1), 1-20.

Li, Z.C., Lam, W.H.K., Sumalee, A., 2008. Modeling impact of transit operator fleet size under various market regimes with uncertainty in network. Transportation Research Record: Journal of the Transportation Research Board 2063, 18-27.

Liu, Z., Meng, Q., 2014. Bus-based park-and-ride system: a stochastic model on multimodal network with congestion pricing schemes. International Journal of Systems Science 45(5), 994-1006.

Löfberg, J., 2004. YALMIP: A toolbox for modelling and optimization in MATLAB. Proceedings of the CACSD Conference.

Meng, Q., Qu, X., 2013. Bus dwell time estimation at bus bays: A probabilistic approach. Transportation Research Part C: Emerging Technologies 36, 61-71.

Misener, R., Floudas, C.A., 2009. Piecewise-linear approximations of multidimensional functions. Journal of Optimization Theory and Applications 145(1), 120-147.

Sherali, H., Alameddine, A., 1992. A new reformulation-linearization technique for bilinear programming problems. Journal of Global Optimization 2(4), 379-410.

Silverman, N.C., 1998. Limited-stop bus service at New York City transit. Journal of Transportation Engineering 124(6), 503-509.

Spiess, H., Florian, M., 1989. Optimal strategies: A new assignment model for transit networks. Transportation Research Part B (2), 83-102.

Sun, L., Meng, Q., Liu, Z., 2013. Transit assignment model incorporating bus dwell time. Transportation Research Record: Journal of the Transportation Research Board 2352, 76-83.

Tétreault, P.R., El-Geneidy, A.M., 2010. Estimating bus run times for new limited-stop service using archived AVL and APC data. Transportation Research Part A 44(6), 390-402.

Ulusoy, Y., Chien, S., Wei, C., 2010. Optimal all-stop, short-turn, and express transit services under heterogeneous demand. Transportation Research Record: Journal of Transportation Research Board 2197, 8-18.

Wang, D.Z.W., Lo, H.K., 2008. Multi-fleet ferry service network design with passenger preferences for differential services. Transportation Research Part B 42(9), 798-822.

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

Zhang, H., Zhao, S., Liu, H., Liang, S., 2016. Design of integrated limited-stop and short-turn services for a bus route. Mathematical Problems in Engineering, 2016.