Multiobjective Optimal Formulations for Bus Fleet Size of Public Transit under Headway-Based Holding Control

Shidong Liang,1 Minghui Ma,2 and Shengxue He1

1Institute: Business School, University of Shanghai for Science and Technology, China
2Institute: College of Automobile Engineering, Shanghai University of Engineering Science, China

Correspondence should be addressed to Minghui Ma; maminghui989@hotmail.com

Received 9 October 2018; Revised 11 December 2018; Accepted 31 December 2018; Published 10 January 2019

Academic Editor: Dongjoo Park

Copyright © 2019 Shidong Liang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In recent years, with the development of advanced technologies for data collection, real-time bus control strategies have been implemented to improve the daily operation of transit systems, especially headway-based holding control which is a proven strategy to reduce bus bunching and improve service reliability for high-frequency bus routes, with the concept of regulating headways between successive buses. This hot topic has inspired the reconsideration of the traditional issue of fleet size optimization and the integrated bus holding control strategy. The traditional headway-based control method only focused on the regulation of bus headways, without considering the number of buses on the route. The number of buses is usually assumed as a given in advance and the task of the control method is to regulate the headways between successive buses. They did not consider the bus fleet size problem integrated with headway-based holding control method. Therefore, this work has presented a set of optimal control formulations to minimize the costs for the passengers and the bus company through calculating the optimal number of buses and the dynamic holding time, taking into account the randomness of passenger arrivals. A set of equations were formulated to obtain the operation of the buses with headway-based holding control or the schedule-based control method. The objective was to minimize the total cost for the passengers and the bus company through calculating the optimal number of buses and the dynamic holding time. The effects of this optimization method were tested under different operational settings. It was found that the model was capable of reducing the costs of the bus company and passengers through utilizing headway-based bus holding control combined with optimization of the bus fleet size. The proposed optimization model could minimize the number of buses on the route for a guaranteed service level, alleviating the problem of redundant bus fleet sizes caused by bus bunching in the traditional schedule-based control method.

1. Introduction

In a stochastic traffic environment, deviations from the schedules of the bus system are unavoidable. The disturbances in bus travel times and passenger arrivals at the stops can lead to failure in schedule adherence and headway uniformity. A delayed or slowed bus usually encounters a larger number of passengers at each stop than expected, and the bus has to dwell longer, creating further delays. At the same time, the following bus has fewer passengers to board and will move faster. Therefore, large headways tend to become larger and small headways smaller. Eventually, some successive buses have little or even no time between them, and “bus bunching” occurs. Bus bunching results in longer waiting and travel times for the passengers. In addition, the unbalanced load of passengers on successive buses wastes bus capacity, because the leading bus will be quite crowded while the trailing bus will be relatively empty.

Under the schedule-based control method, if the bus cannot travel back to the terminal bus stop when it should depart from the bus stop according to the timetable, it means that the schedule has failed. In general, a bus company will increase the number of buses on the bus route to avoid this issue. However, the redundancy in the bus fleet will
bring further chaotic operation and become a burden on urban transportation. Therefore, scholars have focused on real-time headway-based control methods to eliminate bus bunching.

This issue of bus bunching was first presented by Newell and Potts in 1964. Transit agencies have usually handled this issue by introducing slack time into the scheme, giving a delayed bus the ability to recover at control points. Recently, real-time dynamic control strategies have greatly benefited from the advances in technology and the development of monitoring tools such as AVL (Automatic Vehicle Location), APC (Automatic Passenger Counter), and GPS (Global Positioning System). Using these tools, decision makers are able to know the actual behavior of the transit network and implement real-time control strategies. Ibarra-Rojas et al. [1] provided an excellent summary of the previous work on real-time headway control strategies. In terms of spatial configuration, different control strategies can be classified into two categories: station control, including holding strategies [2–4] and stop-skipping strategies [5–7], and interstation control, including bus speed regulation strategies [8] and traffic signal priority strategies [9–11]. In addition, two control methods can be integrated to regulate bus headways or optimize operations of public transport systems [12–16].

Daganzo [17] proposed a method to achieve the target headways. This method attempts to make the headways pre-settled static values. Daganzo and Pilachowski [8] proposed a control strategy that continuously adjusts the bus's cruising speed on the route based on a cooperative two-way system to achieve proper spacing between the buses. As an extension of this idea, Xuan et al. [18] proposed a holding strategy to regularize headways while maximizing the commercial speeds, as well as considering both the forward and the backward headways. Bartholdi and Eisenstein [19] adjusted the trip regularity to achieve a unique and static headway, leading to the satisfactory performance of the system. To integrate the advantages of the “two-way-looking control” with the “self-adaptive equalizing bus headway,” recently, Liang et al. [20] proposed a self-equalizing control strategy based on the two-way-looking control method (the headways between the bus at the control point and both its leading and following buses) with zero slack. More recently, Zhang and Lo [21] proposed two-way-looking self-equalizing headway control that considered multifarious variables, enriching the headway-based bus holding control method system.

In general, the bus fleet size problem is considered a subproblem of Transit Network Timetabling or Vehicle Scheduling Problem instead of independent. Transit Network Timetabling is to define arrival and departure times of buses at all stops along the transit network in order to achieve different goals such as the following: meet a given frequency, satisfy specific demand patterns, maximize the number of well-timed passenger transfers, and minimize waiting times. Vehicle Scheduling Problem is to determine the trips-vehicles assignment to cover all the planned trips such that operational costs based on vehicle usage are minimized.

In terms of Transit Network Timetabling, Mauttone and Urquhart [22] propose a multiobjective optimization approach for minimizing users' costs (based on in-vehicle time, waiting time, and transfer time) and the fleet size. The authors propose a GRASP algorithm for the multiobjective optimization problem. Medina et al. [23] formulate an optimization problem to simultaneously determine the stop density in a bidirectional corridor and the lines' frequency for several periods. The objective is to minimize the user costs and operating costs based on waiting time, in-vehicle travel time, fleet operating costs, and stops installation. Later, Nikolić and Teodorović [24] extend their previous approach outlined by Nikolić and Teodorović [19] considering elastic demand and the minimization of the weighted sum of the total number of unsatisfied passengers, the total travel time of all passengers, and the fleet size. Amiriour et al. [25] determine a set of lines to be implemented for an entire year considering seasonal demand patterns with a probability of occurrence. The authors propose a mathematical formulation to minimize the expected value of the weighed sum of passengers' total waiting time, unused seat capacity, unsatisfied demand (passengers with more than a specific number of trip legs), and the fleet size.

In terms of Vehicle Scheduling Problem, Baita et al. [26] present several approaches to solve the Vehicle Scheduling Problem considering the following practical elements: it is possible to perform deadheading; it is possible to fuel a vehicle after finishing a trip; and it is possible to hold a bus at the end of the trip to wait for the starting time of the following assigned trip. Moreover, the authors considered criteria such as minimizing the fleet size, minimizing the number of lines that a bus is assigned to, minimizing the deadhead costs, and minimizing the idle times of buses waiting for the next trip. Liu and Shen [27] integrate the Transit Network Timetabling presented by Liu et al. [28] and a multiple depot vehicle scheduling problem minimizing the number of buses and deadhead costs. To solve the problem, the authors develop a Bi-level Nesting Tabu Search which is implemented in a small example. Guillaume and Hao [29] propose a mathematical formulation for the integration of Transit Network Timetabling and Vehicle Scheduling Problem minimizing a weighted objective function based on the following elements: (i) number and quantity of transfers; (ii) headway evenness; (iii) fleet size; and (iv) deadhead costs. The model considers a limited deviation from an initial timetable which allows the authors to define feasible shifting procedures. Kéri and Haase [30] and Kéri and Haase [31] address this problem, minimizing the number of buses and crew costs subject to constraints for task covering, bus-driver coupling, and a flexible timetable. This flexibility is defined based on flexible groups, that is, sets of trips that could be shifted together in order to keep the level of service in terms of waiting times. The proper division of an operational day into smaller planning periods based on demand behavior is an important aspect in the implementation of deterministic optimization approaches. In this matter, Zhoucong et al. [32] propose a clustering-based method to generate short planning periods with low variability of travel times. Ceder [33] addresses an interactive heuristic approach to find timetables with
even loads and even headways which take advantage of different vehicle types. The approach is a combination of previously defined tools [34] to compute timetables with even loads, timetables with even headways, and vehicle schedules with minimum operational costs. The weight parameters in weighted objective functions must represent the planner’s preferences for the different objectives. The latter is an issue if two or more objectives are in conflict. Then, Ibarra-Rojas et al. [35] propose a biobjective optimization problem to jointly solve problem considering time windows for departure times and assuming constant demand. The objectives are maximizing the number of passengers benefited by well-timed transfers and minimizing the fleet size. The authors implement a constraint algorithm to obtain Pareto optimal solutions; thus, they are able to measure the “cost” of a vehicle in terms of passengers transfers and vice versa.

The method that integrated “two-way-looking control” with “self-adaptive equalizing bus headway” has been proven to perform well in reducing bus bunching and improving the service level for passengers. However, this kind of headway-based control method only focused on the regulation of bus headways without considering the number of buses on the route. The number of buses is usually assumed in advance and the task of the control method is to regulate the headways between successive buses. The control method may fail to regulate the bus headways due to quite enough fleet size on the route, even though the control method itself is effective. In addition, the benefits brought about from the advanced control method cannot be achieved directly by the bus company, such as reducing the cost of buying buses or operating costs. Therefore, the primary objective in this paper has been to optimize the bus system using the bus holding control method considering the bus fleet size. Furthermore, the fleet size problem was usually considered a subproblem of schedule-based bus operation process. However, with the development of real-time headway-based holding control, the headway between two successive buses is paid more attention, instead of the time the bus arrived at the control point. However, if the headway-based holding control replaces the schedule-based method, the fleet size should be considered under the novel operation process, which has not been addressed.

Therefore, the main contribution of this paper has been to integrate the self-adaptive control method with the optimal bus fleet size in order to optimize the bus system, which can increase the performance of the bus holding control method and bring benefits created by the advanced holding control method to the bus company.

The remainder of this paper has been organized as follows. In Section 2, the problem description and notations have been given first. The buses’ operating process has been described by formulas under the scheme-based bus holding control in Section 3, followed by the basic control scheme that has been used in this paper. The enhanced self-adaptive control method has also been proposed, which can dynamically decide which bus should dwell at the control point and its holding time. In Section 4, a cost objective function has been formulated including buying buses, operation, and cost for passengers waiting at bus stops. Finally, a set of tests have been conducted in Section 5, using the proposed scheme-based control method and the enhanced self-adaptive control method, respectively.

2. Problems Statement and Notations

In practice holding control is usually applied to regulate the bus headways on the bus route and resist bus bunching. Compared with the no control scene or schedule-based control method, the headway-based control method can make the bus operation process, maintaining stronger stability. There are three major advantages for the public transit system.

First, the more balanced distribution of bus headways can obviously reduce the waiting time for passengers at the bus stops, because the large bus headways are vanished. Therefore, with the same bus fleet size, the headway-based control method can provide better service level for the passengers. In other words, in order to provide the same service level for passengers, the bus headway-based control method may use less number of buses.

Second, the unbalanced load of passengers on successive buses wastes bus capacity. This is because the leading bus is quite crowded, while the trailing bus is relatively empty. Therefore, the chaos bus operation process cannot make maximum utilization of bus fleet size. If the full bus capacity of each bus can be used properly, the bus fleet size, to some extent, can be cut down.

Third, as shown in the previous research works (e.g., [17, 20]), the proper headway-based control method can improve the cruising speed of buses on the bus route. Because the holding time or slack time can be greatly cut down, the bus can run back to the terminal bus stop with less time with the headway-based control method. If the length of bus route and the bus headway are fixed, the bus fleet size can be saved (the number of buses is assumed to be obtained by the bus travel time of one cycle divide the bus headway).

Therefore, although the bus headway-based control method was used to regulate bus headways and resist bus bunching, it can further cut down the redundancy buses on the bus route to save cost for bus company, instead of saving the time cost for passengers only. The redundancy buses will waste the money for company of buying buses, bus operation, and pay for drivers, for example, and the redundancy buses will increase burden for urban traffic. This is the insight for the relationship between the headway-based control method and the bus fleet size, and the motivation of optimizing the bus fleet size under holding control.

In this paper, the fleet size optimization problem is integrated with the headway-based control method. In fact, the headway-based control method is an optimization problem as well to regulate the bus headways and maintain the stable bus operation. If the bilevel optimization model,
including fleet size optimization problem and headway-based optimization control method, is formulated to obtain the optimal fleet size, the model will be complex to be solved. Therefore, we select an analytical version of headway-based control method proposed by Liang et al. [20] and Zhang and Lo [21], which can obtain the optimal real-time holding time by calculation model according to the location between the bus at control point and its leading and trailing bus instead of solving the complex objective function. Therefore, the bilevel optimization model can become down to one single objective function of cost, including the cost of passengers’ total travel time and the cost of company for buying buses and operation. The “lower level optimization” can be replaced by the analytical headway-based control.

The fleet size optimization is a static problem, because for a long time (several years) the total fleet size cannot be changed. However, the headway-based control method is a real-time problem; the dynamic holding time is determined by the current state of bus system. Therefore, the problem should be solved integrated by the Monte Carlo simulation based method. According to enough simulation tests, we can obtain a relatively reliably optimal fleet size which can make the minimum cost of bus system. The theoretical optimal fleet size can be obtained.

In addition, in this paper, the schedule-based bus holding control means the fixed departure interval of buses from the control point. The bus control model assumes that the positions of all the buses are known, as well as the number of passengers in each bus and the number of passengers that are waiting at each stop.

The bus system can be completely defined by the following state variables:

(i) $S$: number of bus stops on the bus route
(ii) $N$: number of buses on the bus route
(iii) $s$: index of bus stops on the bus route ($s \leq S$)
(iv) $n$: index of buses on the bus route ($n \leq N$)
(v) $\tau_{n,s}$: bus n’s arrival time at bus stop s
(vi) $\theta_{n,s}$: bus n’s departure time at bus stop s
(vii) $k_{n,s}$: dwell time required for bus n at stop s to provide service to passengers
(viii) $\eta_{s-1,s}$: bus travel time between two successive bus stops, s-1 and s
(ix) $H$: target bus fixed departure interval for schedule-based bus holding
(x) $C$: capacity of buses on the bus route
(xi) $r$: average arrival rate of passengers at the bus stop s
(xii) $X$: passengers that have arrived in the period $\Delta t$, $X \sim P(r)$
(xiii) $\alpha$: passenger boarding time (min/passenger)
(xiv) $\beta$: passenger alighting time (min/passenger)
(xv) $m_{n,s}$: number of passengers that board bus n at stop s
(xvi) $m_{s,n}$: number of passengers that board bus n at stop s
(xvii) $m_{n,s}$: number of passengers that alight bus n at stop s
(xviii) $m_{s,n}$: number of passengers waiting to board and alight, which can be written as
(xix) $m_{s,n}$: number of passengers at stop s because of bus n capacity limitation
(xx) $L$: length of the bus route
(xxi) SC: the control point.

### 3. Bus Operation Process Description

In this section, the bus operation processes have been described under the schedule-based control method and the headways-based control method, respectively. The schedule-based control method means the bus leaves the control point with a fixed headway. The headway-based control method refers to the enhanced driven self-adaptive control method described by Liang et al. [20]

#### 3.1. Bus Operation Process with Schedule-Based Bus Holding

Each bus run can be regarded as a series of ‘events’ of arrivals and departures which specify the arrival and departure times of each bus n at each stop s along the service route. The evolution of $\theta_{n,s}$ is subject to the boundary condition at the stop s, shown in

$$\theta_{n,s} = \tau_{n,s} + k_{n,s}$$  \hspace{1cm} (1)

The evolution of bus n from bus stop s-1 to s can be written as Eq. (2). Bus travel time $\eta_{s-1,s}$ between two successive bus stops s-1 and s can be obtained or estimated from the detectors [36, 37]:

$$\tau_{n,s} = \theta_{n,s} - \eta_{s-1,s}$$  \hspace{1cm} (2)

In Eq. (1) and Eq. (2), the dwell time of bus n at stop s is an unknown variable. In the traditional method, the dwell time can be determined by the maximum service time for passengers to board and alight, which can be written as

$$k_{n,s} = \max(\alpha \cdot m_{b,s,n}, \beta \cdot m_{a,s,n})$$  \hspace{1cm} (3)

We assume that the passengers’ arrival process obeys the Poisson distribution. Therefore, in a short period $\Delta t$, $x_i$ passengers have arrived at the bus stop. The passengers that arrived during $\theta_{n,s} - \theta_{n-1,s}$ can be expressed by

$$\sum_{i=1}^{\theta_{n,s} - \theta_{n-1,s}}/x_i$$

With respect to the boarding passengers, the passengers waiting at the bus stop s $m_{b,s,n}$ can be expressed as

$$\sum_{i=1}^{\theta_{n,s} - \theta_{n-1,s}}/x_i + m_{a,s,n}$$

Some of these passengers may not board because of the bus’s capacity limitation, and the number of passengers waiting to board should be discussed. If the bus’s capacity is less than the total number of passengers on the bus and also waiting at the bus stop, some passengers are not allowed to board because of capacity limitation. Therefore, the number of boarding passenger can be expressed by
expressed by passengers at bus stops because of capacity limitation can be written as. The departure time from control point can be written as:

\[ m_{b,s,n} = \begin{cases} 
\frac{(\theta_{b,s} - \theta_{b,s-1})}{\Delta t} & \text{if } \sum_{i=1}^{\lambda} x_i + m_{i,s,n-1} - m_{a,s,n} - C \\
\sum_{i=1}^{\lambda} x_i + m_{i,s,n-1} - m_{a,s,n} + m_{b,s,n} - m_{a,s,n} - C & \text{if } \sum_{i=1}^{\lambda} x_i + m_{i,s,n-1} + m_{b,s,n} - m_{a,s,n} \leq C \\
C - m_{o,s-1,n} + m_{a,s,n} & \text{if } \sum_{i=1}^{\lambda} x_i + m_{i,s,n-1} + m_{b,s,n} - m_{a,s,n} > C
\end{cases} \tag{4} \]

When the bus \( n \) departs from bus stop \( s \), the rest of the passengers at bus stop \( s \) because of capacity limitation can be expressed by

\[ m_{l,s,n} = \begin{cases} 
0 & \text{if } \sum_{i=1}^{\lambda} x_i + m_{i,s,n-1} + m_{o,s-1,n} - m_{a,s,n} \leq C \\
\frac{(\theta_{a,s} - \theta_{a,s-1})}{\Delta t} & \text{if } \sum_{i=1}^{\lambda} x_i + m_{i,s,n-1} + m_{a,s,n} - m_{a,s,n} > C
\end{cases} \tag{5} \]

The unknown variable \( m_{o,s,n} \) can be regarded as the function of the number of passengers at the upstream bus stop in bus \( n \), and the number of alighting and boarding passengers at bus stop \( s \), which can be written as

\[ m_{o,s,n} = m_{o,s-1,n} + m_{b,s,n} - m_{a,s,n} \tag{6} \]

As mentioned in the beginning of this section, under the schedule-based bus operation, the buses are asked to depart the first bus stop with a fixed headway for its leading bus. If the headway between the bus at bus stop \( SC \) (control point) and its leading bus is less than \( H \), the bus at bus stop \( SC \) should be delayed for a certain time. However, in this research, the bus fleet may not be large and, when the bus arrives at bus stop \( SC \), if the headway with its leading bus is larger than \( H \), the bus that just arrived at bus stop \( SC \) should depart at once after picking up the waiting passengers. The departure time from control point can be written as

\[ \theta_{n-1,SC} = \begin{cases} 
\theta_{n,SC} + H & \text{if } \tau_{n-1,SC} + k_{n,SC} - \theta_{n,SC} \leq H \\
\tau_{n-1,SC} + k_{n-1,SC} & \text{if } \tau_{n-1,SC} + k_{n-1,SC} - \theta_{n,SC} > H
\end{cases} \tag{7} \]

3.2. Bus Operation Process with Headway-Based Bus Holding

In this section, the headway-based control method described by Liang et al. [20] has been presented first. In the original control model, the random variables were not considered. Therefore, in order to fix these gaps, the enhanced control method that has been proposed in this section makes the self-adaptive control method more flexible in handling the variables in the public transit system, such as the stochastic arrival process of the passengers.

3.2.1. Enhanced Version of Headway-Based Control Procedure

As shown in the Appendix, the bus headways will iterate to convergence by continually equalizing the headways between the bus at the control point with its forward and backward buses (\( \lambda = 0.5 \)). In practice, the headway is defined as the time between two successive buses leaving the same bus stop. Therefore, only three buses have been focused on at once, which are the bus at the control point and its following and leading buses. To obtain the bus headways among the three buses, it is necessary to predict the time of the last bus among the three buses leaving the control point, and the related research is relatively mature [38, 39]. Therefore, according to this control concept, a procedure has been designed to select the bus that should be controlled and find its corresponding holding time. The advantage of this method is that extending the self-equalizing bus headway control method makes it more flexible in handling the stochastic variables. The procedure of the proposed control method can be described in the flowchart shown in Figure 1.

\[ \begin{align*}
\text{Step 1.} & \text{ If the two buses' headways among the three buses are equalized without any control, the bus arriving at the control point can provide service for the passengers to alight and board at that control point. Then the bus can leave the control point directly.} \\
\text{Step 2.} & \text{ If the two headways are not equal, the bus holding control means should be considered to be used for the bus at the control point.} \\
\text{Step 3.} & \text{ Based on the headways evolution prediction, the hold time is an unknown quantity, and the predicted two bus headways are involved with the hold time parameter. To balance the successive headways, the two predicted headways should be equal. By solving this equation, the hold time can be obtained.} \\
\text{Step 4.} & \text{ The hold time calculated here may be either positive or negative. A positive hold time indicates that two successive bus headways can be balanced via the bus holding control}
\end{align*} \]
means to the bus at the control point with the proper hold time. In contrast, if the hold time is negative, it means that the bus holding control means is not suitable for balancing the headways at time $t$, and the bus at the control point should leave the control point directly after providing service to the passengers waiting at the control point.

In conclusion, according to the principle of balancing successive bus headways at the control point, the control means is dynamically selected by comparing the results of the headways evolution using the control means. The procedure of the proposed control method for the bus system has been shown in Figure 1. According to the procedure, we can obtain the holding time calculation model, as described in Section 3.2.2.

3.2.2. Dwell Time Calculation Model. The bus headways can converge to a common value, as long as the two bus headways at the control point are continuously balanced. As the analysis mentioned in Section 3.2.1, only three buses have been focused on when the bus arrives at the control point, which are the bus at the control point and its forward and backward buses. As shown in Step 3 of the control method procedure, in order to calculate the bus holding time, it is necessary to predict the time of the following bus leaving the control point. Let the two bus headways, among the three successive buses, be equal. The hold time can be obtained and written

$$
\tau_{n,s} + k_{n,s} + T - \theta_{n-1,s} = \theta_{n+1,s} - (\tau_{n,s} + k_{n,s} + T)
$$

$$
\Rightarrow T = \frac{1}{2}(\theta_{n+1,s} + \theta_{n-1,s}) - \tau_{n,s} - k_{n,s}
$$

The hold time calculated here may be positive or negative. As shown in Figure 1, if the hold time is nonnegative, it means that the two successive bus headways can be balanced by the bus holding control means for the bus at the control point with the proper extra hold time $T$. On the contrary, if the hold time is negative, it means that the bus holding control means is not suitable to balance the headways and the bus should leave the control point immediately after providing service to the passengers. Therefore, the extra hold time at the control points can be expressed by

$$
T = \max\left(\frac{1}{2}(\theta_{n+1,s} + \theta_{n-1,s}) - \tau_{n,s} - k_{n,s}, 0\right)
$$
The departure time can be expressed by the total dwell time (including providing service for passengers and the extra hold time) plus arrival time, as shown in

$$\theta_{n,s} = \max(\alpha \cdot m_{b,s,n} \cdot \beta \cdot m_{n,s,n} + T + \tau_{n,s})$$

(10)

The evolution of the buses' operation under the self-adaptive control method can be described by Equations (1)-(2), (4)-(6), and (9)-(10).

4. Optimization Method for Bus Fleet Size with Bus Holding

The objective is to minimize the total waiting times of passengers at bus stops and the total cost for the bus company including the cost of operating and buying buses. The passenger arrival process can be regarded as a Poisson distribution.

The parameter passenger arrival process in the objective function

$$J = \gamma_1 \cdot I_1 + \gamma_2 \cdot I_2 + \gamma_3 \cdot I_3$$

(11)

where $I_1$ is the cost for passengers waiting at bus stops, $I_2$ refers to the operating cost of the bus company, and $I_3$ is the cost to the bus company of buying buses; each of the three terms is multiplied by different weighting factors, $\gamma_1$, $\gamma_2$, and $\gamma_3$.

The first term in Eq. (11) refers to the at-stop waiting time experienced by passengers and the extra waiting time of passengers who are prevented from boarding the bus because it is at capacity:

$$I_1 = \sum_{c} \sum_{n=1}^{N} \sum_{s=1}^{S} \left( \varphi_1 \cdot \left( \frac{\theta_{n,s} - \theta_{n-1,s}}{\Delta t} - \frac{m_{n,s,n-1}}{\Delta t} \cdot (\frac{\theta_{n,s} - \theta_{n-1,s}}{\Delta t} - i) \right) + m_{n,s,n-1} \cdot (\theta_{n,s} - \theta_{n-1,s}) \right)$$

(12)

$$I_3 = \frac{\varphi_3 \cdot N}{8 \cdot 365}$$

(14)

5. Numerical Tests

A set of numerical tests were conducted to implement the two control methods, considering the bus fleet size and comparing the performance in reducing the cost to the company and the passengers. First, the optimal bus fleet sizes for the two control methods were obtained. Then, using the optimal bus fleet sizes, the total cost of the bus system was calculated. Due to there being a stochastic parameter passenger arrival process in the objective function Eq. (11), it is necessary to borrow the Monte Carlo method to solve the problem. It should be pointed out that the convergence of the Monte Carlo simulation based method has already been proved in the literature when the sample size approaches infinity. Mak et al. [40] gave the statistical analysis results between the quality of the approximated solution and sample size. Yan et al. [41] used the theory in implementation for bus operation. These results can provide a guide to take an appropriate sample size for a given instance.

There were fifteen bus stops on the bus route and the distance between two successive bus stops was 800m, so the total length of the bus route was 12,000m, and the bus's travel speed between two successive bus stops was 5m/s. The bus capacity was 80 pax/veh. Two control points were selected on the bus route. To explore the influence of passenger demand on the bus systems total cost, ten group arrival rates were set, as shown in Table 1. To further describe the passenger arrival process at each bus stop, it was assumed that the average arrival percentage at the bus stops and alighting proportions obeyed the distribution shown in Figure 2.
There were two parts in the numerical tests. In the first section, the optimal bus fleet sizes were calculated by solving the objective function under the two control methods, which were the schedule-based bus holding method and the headway-based bus holding method. For the schedule-based control test, the value range of the fixed headway of departure at control point 1 was from 120 s to 600 s, and the step length was 30 s.

The value of the travel time per passenger and the unit price to buy one bus and the running unit distance were set to be 7 dollars/h, 100,000 dollars/veh, and 2.86 dollars/km. The set values of the weight coefficients have two cases. When $\gamma_1$ was 0.5 and $\gamma_2\gamma_3$ were 0.5, it meant that the costs of the passengers and the bus company were balanced. If $\gamma_1$ was 0.75, it meant more attention was paid to the passengers and less attention was paid to the cost paid by the bus company.

Each test was conducted 100 times for one group shown in Table 1, and in one test the experiment time was 10 hours. In each test, an optimization bus fleet size was obtained by solving the objective function shown in Eq. (11). According to the optimization results, the final optimization bus fleet size for each group could be obtained by calculating the average value of the optimal bus fleet size in the tests repeated 100 times. By comparing the results, the performance of the two control methods in saving the cost of buying buses can be presented.

To further illustrate the performance of the two control methods in saving the costs of the passengers and the bus company, second parts of the tests were conducted. According to the optimization results obtained in Step 1, each test was conducted 100 times for one group shown in Table 1, and in one test the experiment time was 10 hours. The average cost to the passengers and the bus company could be calculated. By comparing the results, the performance of the two control methods in reducing the costs of the passengers and the bus company could be presented. The test results and comparison have been shown in Sections 5.1 and 5.2, respectively.

5.1. Optimal Bus Fleet Size. In the self-adaptive equalizing bus headway control method, only the optimal number of buses on the bus route should be obtained and the headways could be equalized automatically. In the schedule-based control method, the number of buses on the bus route and the fixed headway should both be obtained.

As shown in Table 2, the optimal bus fleet sizes in different groups have been presented. The first row of the table represents different weight coefficient values, and $N_1$ means the optimal number of buses under headway-based control, while $N_2$ and $H_2$ refer to the number of buses on the bus route and the fixed headway under the schedule-based control method. Along with the increasing passenger demand, the numbers of buses for the two control methods both increased, and the fixed bus headway under the schedule-based control method decreased. When the weight coefficient values were 0.5, it can be seen that the costs for buying buses under the two control methods were almost the same. When the weight coefficient values of $\gamma_1$ were 0.75 and $\gamma_2\gamma_3$ were 0.25, the cost of buying buses under headway-based control was less than that under schedule-based control.

To further explore the performance of the two control methods in reducing the total cost of the bus system, the bus fleet size shown in Table 2 can be used and the total cost can be calculated using Eq. (11). The test results have been shown in Section 5.2.

5.2. Performance of Headway-Based Holding Control. In this section, the comparison of the performance of the methods for reducing the cost of the bus system has been presented. First, the values of $J_1$, $J_2$, and $J_3$ have been shown in Table 3,
followed by the relative difference between the two control methods for the cost for both the passengers and the bus company, as shown in Table 4.

The self-adaptive control method was named control 1, while the schedule-based control method was named control 2. As shown in Table 3, $J_1$ refers to the cost for the passengers, $J_2$ refers to the cost of buying buses, and the $J_3$ is the cost of the buses’ operation. By comparison of the cost under different weight coefficients, it could be found that the cost for the passengers, when $\gamma_1$ was 0.5, was larger than when $\gamma_1$ was 0.75, because the bus system pays more attention to the feelings of passengers compared with the previous case. It can be seen from the table that the values of the cost increased along with the growing of passenger demand. More specifically, according to the comparison of the cost values under headway-based control and under schedule-based control, it has illustrated that the headway-based control method performed better in reducing the cost for the passengers, while the cost to the bus company was larger under the headway-based control method. To further explore the difference in the cost under the different situations, the relative difference between the two control methods for the cost for the passengers and the bus company was calculated, as shown in Table 4.

As presented in the first four columns of Table 4, when $\gamma_1$ was 0.5, the total cost for passengers under the headway-based control method was smaller than in the schedule-based control method and the relative difference changes changed from 7.7% to 9.1%. The contributions mainly came from the improvement of the service level for passengers reducing waiting time, while the cost for the bus company made a negative contribution. As presented in the last three columns, when $\gamma_1$ was 0.75, the total cost for the passengers under the headway-based control method was smaller than in the schedule-based control method and the relative difference changes changed from 9.2% to 13.5%. The contributions came from both the passengers’ costs and the bus company costs. From this table, the conclusion can be drawn that headway-based control could improve the service level for passengers with less cost to the company, because the schedule-based method increased the number of buses and had less of a benefit as shown in the last three columns in Table 4.

The stability of the control method is an important index to evaluate the performance of the control method. According to the description mentioned above, each test was conducted 100 times for one input condition. Therefore the deviation of cost, $J_1$, was calculated to reflect the performance of the control method in regard to stability. If the value of the deviation is small, it illustrates that the control method can maintain a stable performance for a certain service level. On the contrary, if the deviation is relatively large, it means the control method is not strong enough even if the average cost of the bus system is low. The deviations of the results of the cost $J_1$ have been shown in Figure 3. The full line denotes the deviation of cost $J_1$ under the weight coefficient $\gamma_1$ 0.5, while the dotted line denotes the deviation of cost $J_1$ under the weight coefficient $\gamma_1$ 0.75.

As presented in Figure 3, the deviation of the cost $J_1$ increased along with the increase of passenger demand. In addition, the deviation under the headway-based control method was less than under schedule-based control, which means the success rate of the headway-based control method was higher for regulating the bus’s operation. The dotted lines are below the full lines, which illustrates that the service level for passengers was better when the weight coefficient $\gamma_1$ was 0.75.

### 6. Conclusion

This paper has proposed an optimization method for a bus system using the bus holding control method, considering bus fleet size. First, the traditional self-equalizing bus headway control method was extended to handle more general situations, which enhanced the flexibility of the control method. Then the objective function was formulated including the cost of buying the bus, the operation, and the waiting time for passengers. Finally, a set of numerical tests were conducted to implement the proposed optimization method for the bus system.

The results suggested that the optimal number of buses, under the self-equalizing bus headway control method, changed gradually along with the increase in passenger demand. However, the optimal number of buses under the schedule-based control method changed greatly along
Table 3: The cost values of $J_1$, $J_2$, and $J_3$.

<table>
<thead>
<tr>
<th>Group</th>
<th>$\gamma_1=\gamma_2=\gamma_3=0.5$</th>
<th>$\gamma_1=0.75,\gamma_2=\gamma_3=0.25$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>control 1</td>
<td>control 2</td>
</tr>
<tr>
<td>1</td>
<td>$J_1$</td>
<td>$J_2$</td>
</tr>
<tr>
<td>2</td>
<td>5183</td>
<td>2192</td>
</tr>
<tr>
<td>3</td>
<td>5620</td>
<td>2192</td>
</tr>
<tr>
<td>4</td>
<td>5255</td>
<td>2466</td>
</tr>
<tr>
<td>5</td>
<td>5726</td>
<td>2466</td>
</tr>
<tr>
<td>6</td>
<td>5380</td>
<td>2740</td>
</tr>
<tr>
<td>7</td>
<td>5812</td>
<td>2740</td>
</tr>
<tr>
<td>8</td>
<td>5469</td>
<td>3014</td>
</tr>
<tr>
<td>9</td>
<td>5918</td>
<td>3014</td>
</tr>
<tr>
<td>10</td>
<td>6082</td>
<td>3288</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
with the increase in passenger demand, especially when the weight coefficient value was relatively large for the cost of the passengers. In addition, the self-equalizing bus headway control method could perform best for the optimal number of buses.

**Appendix**

In this section, the control method mentioned in Section 3.2 is proved effective with the homogeneous Markov chain theory. Our idealized model treats the bus route as a dynamical system with \( n \) buses moving at a constant average velocity \( v \) around a circular route of length 1, with a single control point at 0 (equivalently, 1). At any point in time each bus \( i \) has a location \( l_i \in [0, 1) \) on this circuit.

Let those times at which a bus arrives at the control point be indexed by \( t = 1, 2, \ldots \). At each such time we reindex the buses so that the bus that has just arrived at the control point is the bus 1, the next in the direction of bus movement is the bus 2, and so on, until the last bus, which is the next bus to arrive at the control point, is the bus \( n \). For each time \( t \) let the vector \( \mathbf{l} = (l_1^t, l_2^t, \ldots, l_n^t) \) represent the locations of the buses. From arbitrary starting positions \( \mathbf{l}^0 \), the trajectory of bus positions \( \{l_1^t, l_2^t, \ldots\} \) may be thought of as a series of snapshots of the bus route at those times when a bus arrives at the control point.

As shown in Figure 4, let the vector \( h_i^t = (h_1^t, h_2^t, \ldots, h_n^t) \) denote the headways of the buses at time \( t \). In the absence of perturbations, \( h_i^t = (l_{i-1}^t - l_i^t)/v \) for all buses \( i \), except for the bus 1, which we require to pause at the control point for time \( T_i \).

The headway evolutions can be obtained easily based on the research by Bartholdi and Eisenstein [19] and Liang et al. [20], which are expressed by Eq. (A.1), if the forward headway is smaller than the backward headway:

\[
\begin{align*}
    h_1^{t+1} &= \frac{l_1^{t+1} - l_1^t}{v} = (1 - \lambda) h_1^t + \lambda h_2^t \\
    h_2^{t+1} &= \frac{l_2^{t+1} - l_2^t}{v} = \left(\frac{h_1^t + h_2^t}{v} - (h_1^t - T_1)ight) v \\
    &= (1 - \lambda) h_1^t + \lambda h_2^t 
\end{align*}
\]

(A.1)

The headway evolutions can be expressed by Eq. (A.2), if the forward headway is larger than the backward headway:

\[
\begin{align*}
    h_1^{t+1} &= h_1^t \\
    h_2^{t+1} &= \frac{l_2^{t+1} - l_2^t}{v} = \left(\frac{h_1^t + h_2^t}{v} - (h_1^t - T_1)ight) v = h_1^t \\
    h_i^{t+1} &= \frac{(l_i^t + h_i^t)v - (l_{i-1}^t + h_{i-1}^t)v}{v} = \frac{l_i^t - l_{i-1}^t}{v} = h_i^{t-1} \\
\end{align*}
\]

(A.2)

We can write the linear system as \( \mathbf{h}^{t+1} = \mathbf{h}^t A \). Because the headways evolution has two cases according to the successive
bus headways at the control point, the matrix $A$ has two forms. If the forward headway is smaller than the backward headway, the matrix $A_1$ can be expressed by

$$A_1 = \begin{bmatrix}
\lambda & 1-\lambda & 0 & 0 & 0 & \cdots & 0 \\
0 & 0 & 1 & 0 & 0 & \cdots & 0 \\
0 & 0 & 0 & 1 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & 0 & 0 & \cdots & 1 \\
1-\lambda & \lambda & 0 & 0 & 0 & \cdots & 0
\end{bmatrix}$$

(A.3)

If the forward headway is larger than the backward headway, the matrix $A_2$ can be expressed by

$$A_2 = \begin{bmatrix}
0 & 1 & 0 & 0 & 0 & \cdots & 0 \\
0 & 0 & 1 & 0 & 0 & \cdots & 0 \\
0 & 0 & 0 & 1 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & 0 & 0 & \cdots & 1 \\
1 & 0 & 0 & 0 & 0 & \cdots & 0
\end{bmatrix}$$

(A.4)

It is easy to know that $A_2$ has no influence on the distribution of bus headways; therefore, we focus on the discussion of the matrix $A_1$. The state space is $I = 0, 1, 2, \ldots, n-1$, and the state transfer graph is shown in Figure 5.

As shown in Figure 5, all the states in the fig are interlinked, so every state is a regular return state. In addition, because $A_1(0,0)$ is $\lambda$ larger than zero, this Markov chain is traversed. Therefore, the probability of stability meets the equations as follows:

$$\pi(0) = \lambda \cdot \pi(0) + (1 - \lambda) \cdot \pi(n - 1)$$

$$\pi(1) = (1 - \lambda) \cdot \pi(0) + \lambda \cdot \pi(n - 1)$$

$$\pi(2) = \pi(3)$$

$$\vdots$$

$$\pi(n - 2) = \pi(n - 1)$$

(A.5)

In addition, the $\pi(i)$ meets the equation as follows:

$$\sum_{i=0}^{n-1} \pi(i) = 1$$

(A.6)

Therefore, the probability of stability $\pi(i)$ can be obtained by integrating Eq. (A.5) and Eq. (A.6):

$$\pi(0) = \pi(1) = \cdots = \pi(n - 1) = \frac{1}{n}$$

(A.7)

According to Eq. (A.7), we can obtain the convergent headways, shown as

$$h_c = h_1^\infty A_1^\infty = \frac{h_1 + h_2 + \cdots + h_n}{n} [1, 1, \ldots, 1]$$

(A.8)

We can figure out from Eq. (A.8) that the convergent headways are equal to each other after several iterations.

**Data Availability**

The data used to support the findings of this study are available from the corresponding author upon request.
Conflicts of Interest

We confirm that there are no conflicts of interest regarding the publication of this manuscript.

Acknowledgments

The work was supported by the National Natural Science Foundation of China (Grant NO. 71801153), National Natural Science Foundation of China (Grant NO. 71801149), Natural Science Foundation of Shanghai (Grant NO. 18ZR1426200), and Youth Research Projects of Shanghai (Grant NO. ZZslgl8013). This article is an extended version of the article (https://www.icevirtuallibrary.com/doi/10.1680/jmuen.ZZslg18013). This article is an extended version of the article presented in the Transportation Research Board TRB 98th Annual Meeting. All data included in this study is available upon request by contacting the corresponding author.

References


