Optimization of Bus Scheduling and Bus-Berth Matching at Curbside Stops under Connected Vehicle Environment

Shiqi Ou, Wanjing Ma, and Chunhui Yu

The Key Laboratory of Road and Traffic Engineering of the Ministry of Education, Tongji University, 4800 Cao’an Road, Shanghai 201804, China

Correspondence should be addressed to Chunhui Yu; hughyu90@tongji.edu.cn

Received 5 February 2022; Revised 1 May 2022; Accepted 12 May 2022; Published 28 May 2022

Copyright © 2022 Shiqi Ou 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.

It is commonly seen that buses are blocked by the ones in front serving passengers and have to queue outside a curbside bus stop although there are vacant berths at the stop. The resultant bus delays degrade the service level of urban public transportation. A potential solution is to reschedule the arrivals of the buses at the stop for full utilization of the berths with the aid of connected vehicle technologies. This study proposes a mixed-integer linear programming model to optimize the scheduling of bus arrivals and the bus-berth matching at a curbside stop under the connected vehicle environment. The objective is the minimization of the bus delays weighted by the number of passengers on the buses. Bus arrival times at the stop and the assignment of berths are optimized together with bus departure times from the stop. Bus punctuality is also taken into consideration. The proposed model could be applied dynamically to cater to time-varying traffic conditions. Numerical studies validate the advantages of the proposed model over the “first-come-first-service” strategy and the relaxed model without bus punctuality in terms of weighted bus delays and bus punctuality. Sensitivity analyses show that (1) the proposed model is robust to the fluctuation of bus service time and (2) a smaller number of berths may be preferred on the condition that the bus demand does not exceed the stop capacity.

1. Introduction

Urbanization has increased the population of human beings and vehicles in cities. This has resulted in road congestion by automobiles, a major proportion of which is private vehicles. The adoption of public transportation in urban areas can significantly decongest the roads. As an important component of public transportation, buses have higher capacities, lower costs, and larger energy savings than private vehicles. In some cities, buses are prioritized by applying dedicated bus lanes and signal priority to enhance bus service efficiency, which is crucial to urban transportation systems. Bus system management for the improvement of bus service efficiency is a widely studied topic [1–3].

Lots of studies have been dedicated to bus operations. One research direction is to enhance the reliability of bus systems. This category of studies mainly focuses on equalizing bus headway by bus holding at stops [4], stop skipping [5], and a combination of both [6]. Moreover, connected vehicle technologies were introduced by Bie et al. [7] to dynamically control bus headways with the development of wireless communication systems and global positioning systems [8]. Another research direction is to improve the efficiency of bus systems. Lots of studies have investigated signal timings at intersections along bus lines such as bus signal coordination [9–11]. Due to the connected vehicle environment, bus information (e.g., bus service time [12] and travel time [13]) can be collected in real time for signal timings and control strategies can be sent to buses for trajectory control [8]. Bus speed advisory in dedicated bus lanes [14], signal priority at intersections [15, 16], and their combination [17] have been widely studied. To improve the utilization of dedicated bus lanes, the management and the performance evaluation of intermittent bus lanes have also been explored under the connected vehicle environment [15, 18].

Besides intersections, bus operation at stops is another key to the enhancement of the service level of bus systems. To improve bus operation at stops, relocation of bus stops has been investigated, for example, by comparing the
advantages and disadvantages of near- and far-side bus stops [19] and optimizing the distance between bus stops [20]. Further, signal timings at the downstream intersection are taken into consideration to avoid stops of buses at the stop line after leaving the upstream bus stop [21]. Another category of studies focuses on the design of bus stops. Bus-stop designs (curbside or lay-by), berth numbers, and berth lengths are optimized [1, 22, 23]. Typically, there are four types of berths, namely, drive-through berths, angle berths, sawtooth berths, and linear berths [24]. The first three types of berths allow independent bus movements but require large space. In contrast, linear berths require much less space and are most commonly used at curbside stops.

However, at a curbside stop with linear berths, it happens that the adjacent lane of bus berths is occupied by other cars and the interberth spaces are small, which makes it difficult for buses to conduct overtaking. Further, overtaking may be prohibited at a curbside stop in some cities, e.g., in Shanghai, P.R.C [1]. As a result, it is commonly seen that buses are blocked from entering or leaving a curbside stop by other buses serving passengers. This phenomenon is called mutual blockage effects [1, 25] as shown in Figure 1. In Figure 1(a), the light grey bus is blocked from entering the stop by the dark grey one occupying the uppermost berth although the downstream berth is vacant. In Figure 1(b), the light grey bus is blocked from leaving the stop by the dark grey one occupying the downstream berth. For curbside stops, capacity approximation, bus queueing models, and overtaking maneuver effects have been studied and analyzed [26–29], all of which find that the mutual blockage phenomenon is one of the causes of the reduction in bus-stop operation efficiency. Shen et al. [26] developed a capacity approximation method that could be used by practitioners to decide the location and berth number of a new bus stop. The model considered the queued buses entering a curbside stop as convoys of the size of the berth number, which took first-come-first-service as dwelling strategies. Gu et al. [27] and Gu and Cassidy [28] proposed analytical models predicting maximum bus flows at a bus stop under prohibiting overtaking and only permitting overtaking out situations, respectively. Both of them found stop capacities diminished due to the mutual blockage effects, especially when bus service time variability was high. Bian et al. [29] furnished queueing models of four overtaking rules (free overtaking, only overtaking out, only overtaking in, and overtaking prohibited) as well as the stop capacities and waiting time. Comparisons between different overtaking rules indicated the significant deterioration of bus operation due to mutual blockage effects when overtaking was prohibited at curbside stops. The mutual blockage also accounts for the effect that increasing the number of berths at a linear bus stop has an ever-decreasing effect on capacity as the number of berths increases [25–27]. These studies found that one of the causes for bus-stop discharge flow deduction is mutual blockage effects between successive buses. Unfortunately, few studies so far consider potential optimization control methods for bus-dwelling processes to alleviate mutual blockage effects.

When locations of bus stops and the numbers of berths are not able to change, it is realized that speed advisory and berth assignment for buses provide a prospective solution to the alleviation of bus mutual blockage at curbside stops under the connected vehicle environment. But limited studies have been reported. Gibson [30] suggested that buses must dwell at the lowermost berth to enhance bus-stop capacity. But this principle-based method cannot guarantee the global optima, especially when the mutual blockage is severe. Zhou [31] established a system structure for real-time queue guiding at a multitherth bus stop. Speed advisory was integrated, but no specific algorithms were presented in detail. Li [32] and Huang et al. [33] proposed rule-based methods for the dynamic assignment of berths to incoming buses based on the bus arrival order and berth-occupancy conditions. Li [32] only considered the first two incoming buses within 200 m upstream of the stop. Huang et al. [33] predicted the berth allocation only to release bus-berth information at the stop, but buses might not dwell at the predicted berth. And the nature of the rule-based approach cannot guarantee the global optima. In addition, berth assignment has also been explored at a bus stop with multiple bus lines [34]. However, these studies focused on the assignment of berths to bus lines rather than to individual buses. That is, buses in the same line always dwelled at the same berth, which might cause huge delays due to bus bunching.

Notwithstanding the above studies, the integrated optimization of bus arrivals and berth assignment to individual buses at a curbside bus stop is missing. It could further alleviate the mutual blockage, especially when the service time of buses at the stop varies to a great extent. To fulfill the research gap, this study proposes a mixed-integer linear programming (MILP) model to optimize the scheduling of bus arrivals and bus-berth matching at a curbside bus stop together with bus departure times from the stop under the connected vehicle environment. The minimization of the total weighted bus delay at the bus stop is adopted as the objective function. The delay of each bus is weighted by the number of its accommodated passengers. Bus arrival and departure processes at the stop are modeled with the

Figure 1: Mutual blockage situations: (a) in front of the stop and (b) inside the stop.
consideration of safety, bus mutual blockage, time of serving boarding and alighting passengers, and bus punctuality.

The remainder of this paper is organized as follows. Section 2 describes the problem in detail. Section 3 formulates the MILP model. In Section 4, numerical studies are conducted to evaluate the performance of the proposed model in both static and dynamic ways. Additionally, the sensitivity analysis is conducted in terms of service time, bus volumes, and berth numbers. Section 5 concludes the study.

2. Problem Description

Figure 2 shows a typical curbside bus stop (the target stop) with the linear berth design, which is most commonly seen in the real world. The target stop is an intermediate stop because there is no need to optimize the dwelling order at the first stop (with schedules) or at the terminal stop (without boarding passengers). The berths are indexed from 1 to $P$. The most downstream and upstream berths are indexed as Berth 1 and Berth $P$, respectively. When buses have arrived at the stop, no overtaking is allowed, but overtaking can occur on the road. The target stop may serve multiple bus lines with multiple upstream stops. Under the connected vehicle environment, with the control framework shown in Figure 3, each coming bus from its upstream stop could send its arrival time window (i.e., the earliest and latest arrival times) to the control center when it enters the control area. The real-time bus arrival windows are taken as the inputs to the optimization model embedded in the control center. The task is to optimize the scheduling of bus arrivals at the target stop and the bus-berth matching for the buses in the control area together with bus departure times from the stop. Specifically, the optimized bus arrival time refers to the time when the bus can directly start serving passengers without any waiting. If the berths are all occupied, none of the coming buses will arrive at the stop until all of the earlier arriving buses finish their service at the stop. Bus punctuality is taken into consideration with the prescheduled bus departure times from the target stop, which are given in advance. And then, the optimized arrival/departure times and the berth assignment are sent back to the buses from the control center for trajectory planning to avoid blockage and stops outside the target stop. The roadside infrastructure, including bus stops and detectors, receives real-time information from buses (e.g., arrival time and passenger numbers) and overall instructions from the control center (e.g., weather and bus schedules). The control center is a cloud center, and on-board unit (OBU) and road side unit (RSU) devices are furnished on each bus and at each bus stop, respectively. With this control framework, the control center can run the proposed model with real-time bus information and sends the optimal bus scheduling and berth matching to each bus to realize the optimization. For simplicity, the following assumptions are made:

1. Dedicated curbside bus lanes are used and the impacts of car traffic are not considered. Dedicated bus lanes could improve the service level of public transportation and are commonly seen where public
transportation plays an important role (e.g., in Shanghai, China) [35].

(2) Each bus could accurately predict its arrival time window based on its state, travel distance, speed limits on roads, vehicle dynamics, etc., which is not the focus of this study. There are lots of existing studies on vehicle travel time prediction [13].

(3) Each bus could plan its trajectory to arrive at the target stop at the scheduled arrival time received from the controller, which is not the focus of this study. There are lots of existing studies on vehicle trajectory planning [36, 37].

3. Problem Formulation

In this section, the MILP model is developed to optimize the arrival/departure times at/from the target stop for the buses in the control area as well as the berth assignment. The constraints and the objective function are presented in the following subsections. Table 1 lists the notations used in this study.

3.1. Constraints. The constraint structure is illustrated in Figure 4. The constraints dealing with decision variable domains, bus arrival rules, and bus departure rules are established in this section.

3.1.1. Decision Variable Domains. The arrival time of bus \( n \) at the target stop is constrained to fall into the arrival time window.

\[
\tilde{t}_{n}^{\text{in}} \leq t_{n}^{\text{in}} \leq \tilde{t}_{n}^{\text{in}}, \quad \forall n \in \Omega,
\]

where \( \tilde{t}_{n}^{\text{in}} \) and \( \tilde{t}_{n}^{\text{in}} \) denote the earliest and latest arrival times of bus \( n \) at the target stop, respectively, and \( t_{n}^{\text{in}} \) denotes the earliest arrival time of bus \( n \) at the target stop.

Table 1: Notations.

<table>
<thead>
<tr>
<th>Notations</th>
<th>Variable description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \Omega )</td>
<td>Set of buses heading toward the target stop, ( \Omega \in {1, 2, \ldots, N} ), where ( N ) denotes the number of the incoming buses in the control area. Each incoming bus is denoted as bus ( n \in \Omega )</td>
</tr>
<tr>
<td>( \omega )</td>
<td>( \omega ) Bus whose berth index is the largest at the target stop when the optimization is conducted</td>
</tr>
<tr>
<td>( a_n )</td>
<td>Number of passengers alighting from bus ( n ) at the target stop</td>
</tr>
<tr>
<td>( b_n )</td>
<td>Number of passengers boarding bus ( n ) at the target stop</td>
</tr>
<tr>
<td>( g_n )</td>
<td>( g_n ) Weight of the delay of bus ( n )</td>
</tr>
<tr>
<td>( M )</td>
<td>A sufficiently large number</td>
</tr>
<tr>
<td>( \epsilon )</td>
<td>( \epsilon ) A sufficiently small positive value</td>
</tr>
<tr>
<td>( P )</td>
<td>( P ) Number of berths at the target stop</td>
</tr>
<tr>
<td>( S_n )</td>
<td>Service time for bus ( n ), ( s )</td>
</tr>
<tr>
<td>( T_{\text{punc}} )</td>
<td>( T_{\text{punc}} ) Acceptable schedule delay, ( s )</td>
</tr>
<tr>
<td>( T_{\text{safe}} )</td>
<td>Safety headway between two successive buses, ( s )</td>
</tr>
<tr>
<td>( t_{\text{acc}} )</td>
<td>Time taken by a bus to decelerate to or accelerate from a berth, ( s )</td>
</tr>
<tr>
<td>( t_b )</td>
<td>Time taken by a bus to traverse a berth, ( s )</td>
</tr>
<tr>
<td>( t_{n}^{\text{in}}, t_{n}^{\text{out}} )</td>
<td>Earliest/latest arrival time of the arrival time window of bus ( n ) at the target stop, ( s )</td>
</tr>
<tr>
<td>( p_n )</td>
<td>Berth index assigned to bus ( n )</td>
</tr>
</tbody>
</table>

### Decision variables

- \( t_{n}^{\text{in}} \): Arrival time of bus \( n \) at the target stop, \( s \)
- \( t_{n}^{\text{out}} \): Departure time of bus \( n \) from the target stop, \( s \)
- \( p_n \): Berth index assigned to bus \( n \)

### Auxiliary variables

- \( b_{m,n}^{\text{in}} \): 1, if bus \( m \) departs later than bus \( n \) arrives; otherwise, 0 (\( m \neq n \))
- \( p_{m,n} \): 1, if bus \( m \) arrives earlier than bus \( n \); 0, otherwise (\( m \neq n \))
- \( \tau_n \): Unacceptable schedule delay of bus \( n \) to be penalized

### Figure 4: Constraint structure.
arrival time of bus \( n \) to be scheduled. Note that, as described in Section 2, a bus arriving at \( t^m_{in} \) could enter the stop without queuing outside the stop. \( t^m_{in} \) is also the entry time into the target stop. Although the bus departure times are also decision variables, they are constrained by arrival times and departure rules, which are described in Section 3.1.3.

One of the \( P \) berths is assigned to bus \( n \) to serve alighting and boarding passengers at the target stop.

\[
1 \leq p^n \leq P, \quad \forall n \in \Omega, \tag{2}
\]

where \( p^n \) denotes the index of the berth assigned to bus \( n \) and \( P \) denotes the number of berths at the target stop.

### 3.1.2. Arrival Rules.

An auxiliary variable \( l^{m,n} \) is introduced to indicate the arrival sequence of two buses.

\[
-Ml^{m,n} + T_{safe} \leq t^m_{in} - t^n_{in} \leq M(1 - l^{m,n}) - T_{safe}, \quad \forall m, n \in \Omega; m \neq n, \tag{3}
\]

\[
t^{\omega}_m + T_{safe} \leq t^n_{in}, \quad \forall n \in \Omega, \tag{4}
\]

\[
l^{m,n} = 1, \quad \forall n \in \Omega, \tag{5}
\]

where \( M \) is a sufficiently large number, \( T_{safe} \) denotes the safety headway between two successive buses, and bus \( \omega \) is the bus whose berth index is the largest at the target stop when the optimization is conducted as shown in Figure 5. Equation (3) guarantees that \( l^{m,n} = 1 \) if bus \( m \) arrives at the target stop earlier than bus \( n \) \( (t^m_{in} + T_{safe} \leq t^n_{in}) \); otherwise \( (t^m_{in} + T_{safe} \leq t^n_{in}) \), \( l^{m,n} = 0 \). Note that \( t^{\omega}_m \) is the arrival time of bus \( \omega \), which is a constant. Equations (4) and (5) are effective only when bus \( \omega \) exists.

An auxiliary variable \( b^{m,n} \) is introduced to model the relationship between the departure time of bus \( m \) and the arrival time of bus \( n \).

\[
-(1 - b^{m,n})M + \epsilon \leq t^m_{out} - t^n_{in} \leq b^{m,n}M, \quad \forall m, n \in \Omega; m \neq n, \tag{6}
\]

\[
-(1 - b^{\omega,n})M + \epsilon \leq t^{\omega}_m - t^n_{in} \leq b^{\omega,n}M, \quad \forall n \in \Omega, \tag{7}
\]

where \( \epsilon \) is a sufficiently small positive value defined as time accuracy and \( t^m_{out} \) denotes the departure time of bus \( m \).

Equation (6) guarantees that \( b^{m,n} = 1 \) if bus \( m \) departs later than bus \( n \) \( (t^m_{out} > t^n_{in}) \); otherwise \( t^m_{out} \leq t^n_{in} \), \( b^{m,n} = 0 \). Note that \( t^{\omega}_m \) is the departure time of bus \( \omega \), which has been scheduled in the previous optimization and is a constant. Equation (7) is effective only when bus \( \omega \) exists.

If bus \( n \) arrives at the target stop later than bus \( m \) \( (l^{m,n} = 1) \) but earlier than bus \( m \) departs \( (b^{m,n} = 1) \), then bus \( n \) is blocked. That is, the berth assigned to bus \( n \) is constrained by the one assigned to bus \( m \).

\[
p^m + 1 \leq p^n + (1 - l^{m,n})M + (1 - b^{m,n})M, \quad \forall m, n \in \Omega; m \neq n, \tag{8}
\]

\[
p^\omega + 1 \leq p^n + (1 - l^{m,n})M + (1 - b^{m,n})M, \quad \forall n \in \Omega. \tag{9}
\]

### 3.1.3. Departure Rules.

The departure time of a bus is related to its arrival time, service time for boarding and alighting passengers, and the time for entering and leaving its assigned berth.

\[
t^m_{in} + ((P - p^n)t_b + t_{acc}) + S^n + (t_{acc} + p^n t_b) \leq t^m_{out}, \quad \forall n \in \Omega, \tag{10}
\]

where \( t_b \) denotes the time for traversing a berth, \( t_{acc} \) denotes the time for decelerating to or accelerating from a berth, and \( S^n \) denotes the service time of bus \( n \) for boarding and alighting passengers at the target stop. It is assumed that the service time of each bus could be accurately predicted, for example, by the compound Poisson service time estimation model in Bian et al. [12], which takes into consideration the interactions among incoming buses and the uncertainty of boarding and alighting passenger numbers. Besides, there are lots of other models for bus service time prediction [12, 24], which are beyond the scope of this study.

Due to the linear berth design, overtaking is prohibited at the target stop. If bus \( n \) follows bus \( m \) to depart from the stop \( (l^{m,n} = 1) \), the safety headway between their departure times should be guaranteed.

\[
t^m_{out} + T_{safe} \leq M(1 - l^{m,n} + t^n_{out}), \quad \forall m, n \in \Omega; m \neq n, \tag{11}
\]

\[
t^{\omega}_m + T_{safe} \leq M(1 - l^{m,n} + t^n_{out}), \quad \forall n \in \Omega, \tag{12}
\]

It is assumed that the prescheduled bus departure times are given in advance. For the sake of boarding passengers, buses are not allowed to depart earlier than the prescheduled departure times [24]. Additionally, the schedule delay tolerance of \( T_{punc} \) is acceptable in terms of punctuality [24]. And the punctuality constraints of the departure times are set as follows:

![Figure 5: Bus ω when the optimization is conducted.](image-url)
Equations (1)–(14) are linear. Therefore, the proposed model (M0) is applied.

where \( t_{out} \) denotes the prescheduled departure time of bus \( n \), \( T_{punc} \) denotes the acceptable schedule delay (e.g., 5 min according to [24]), and \( r^n \) is the unacceptable schedule delay of bus \( n \) to be penalized.

3.2. Objective Function. Total bus delay is used as the performance measure because it is tangible, commonly used, and calculable [38]. To take into consideration passenger benefits [39], weights that are related to the numbers of passengers on the buses are applied, and therefore, the bus delay term is called weighted bus delay hereafter. The objective function is formulated as follows:

\[
\min \sum_{n=1}^{N} (g^n \cdot (t^n_{out} - t^n_{in} - S^n) + M r^n). 
\]  

(15)

In Equation (15), the bus delay is calculated as the difference between the actual departure time \( t^n_{out} \) and the minimum departure time \( t^n_{in} + S^n \). The unacceptable delay \( r^n \) is penalized for punctuality. Equation (15) indicates that bus punctuality is the primary objective. \( g^n \) is the passenger number weight of bus \( n \) defined as follows:

\[
g^n = \frac{N G^n}{\sum_{n=1}^{N} G^n}, \quad \forall n \in \Omega, 
\]

(16)

\[
G^n = \frac{G^n \cdot (G^n - d^n + b^n)}{2}, \quad \forall n \in \Omega, 
\]

where \( G^n \) denotes the average number of passengers on bus \( n \) over the process of serving passengers, \( C^n \) denotes the number of passengers on bus \( n \) before arriving, \( d^n \) is the number of passengers alighting from bus \( n \), and \( b^n \) is the number of passengers boarding bus \( n \).

The objective function Equation (15) and the constraints Equations (1)–(14) are linear. Therefore, the proposed model is an MILP model, which can be solved efficiently by many existing solvers (e.g., Gurobi and Cplex). The numbers of continuous variables, integer variables, and constraints are \( O(N) \), \( O(N^2) \), and \( O(N^3) \). Note that the proposed model could be dynamically applied with updated inputs to cater to time-varying traffic conditions.

4. Numerical Studies

4.1. Experimental Settings. To explore the benefits of the proposed model, a curbside stop with two berths \( (P = 2) \) is used, and both static and dynamic application scenarios are tested. In the static application scenarios, static information of six incoming buses is applied \( (N = 6) \), which is a common situation in rush hours in big cities, e.g., in Shanghai, China. The optimization-free model without bus scheduling and berth assignment (M0, i.e., the “first-come-first-service” strategy), the relaxed model without bus punctuality (M1), and the proposed model (M2) are applied. M0 and M1 are the benchmark models. M1 is derived by relaxing constraints 14–15 and removing the penalty \( M r^n \) from the objective function of M2. Weighted bus delays and bus punctuality per the prescheduled timetable of departure times from the stop are used as the performance measures. When the optimization is conducted, no buses are dwelling at the stop. All the buses share the same timeline, in which the start of the timeline is the time when the optimization is conducted. Thus, a negative value of a bus prescheduled departure time means that the prescheduled departure time is ahead of the start time of the timeline. For example, the prescheduled departure time of bus 2 in scenario 1 is 261 s ahead of the time when the optimization is conducted as shown in Table 2. Model-related parameters are summarized in Table 2. For simplicity, the service time of each bus is assumed to be accurately predicted in advance, for example, by the service time prediction model in Bian et al. [12]. The standard deviation (SD) settings of service time are different in the three scenarios to explore the impacts of the fluctuation of service time because it contributes to the bus mutual blockage (Gu et al., 2011; Ryu et al., 2013). Note that any generated service time less than 3 s is abandoned because the minimal service time should be the dead time for opening and closing doors which is about 3 s in Bian et al. (2015). The arrival time window of each bus at the target stop is given based on the maximum/minimum speeds and the distance to the stop. The safety headway \( T_{safe} \) between two buses is 3 s. Both the time taken by a bus to traverse a berth \( t_b \) and to accelerate or decelerate \( t_{acc} \) is 5 s. The acceptable schedule delay \( T_{punc} \) is 5 min according to [24]. The sufficiently small positive value \( \epsilon \) defined as time accuracy is 0.1 s.

In the dynamic scenarios, buses arrive in fleets to simulate bus bunching. The probability of an incoming bus fleet with one bus, two buses, three buses, and four buses are 50%, 30%, 15%, and 5%, respectively. Successive buses in the same fleet remain the safety headway \( T_{safe} \). The model performance is evaluated under different bus volumes with average arrival intervals of bus fleets from 60 s to 180 s. For example, the average arrival interval of 60 s with the average fleet size of 1.75 buses indicates that the bus volume is \((360/60) \times 1.75 = 105 \) buses per hour. To cater to bus fleet arrival uncertainty, the difference between the generated bus fleet arrival time and the prescheduled bus fleet arrival time at the control area is assumed to follow normal distribution. To simulate the delayed bus arrival phenomena, the mean and the SD of the normal distribution are 72 s and 168 s, respectively [40]. Note that the prescheduled arrival time at the control area of each bus fleet is obtained by assuming uniform arrivals with the average arrival interval. The bus service time and the average number of passengers on each bus also follow normal distribution. The bus service time \( S \sim N(20, 15) \) [41] and the number of passengers \( G \sim N(20, 10) \). The incoming bus fleets are generated at the upstream stops of 340 m and 400 m with the same probability. With the maximal and minimal bus speeds, bus arrival time windows can be calculated. The prescheduled departure time of a bus is its prescheduled arrival time at the control area plus its minimal travel time and its predicted service time.
measure the bus punctuality \[42\], and it is defined as \( t_{n} \) and \( M_{1} \) relaxed in \( M_{2} \) dwellingsystem.

In the actual application of this model, the MILP model is solved by Gurobi in Python on a desktop with an Intel 1.8 GHz CPU with 8 GB of memory. The average computational time of solving the proposed model in different scenarios is shown in Table 3, which reveals the potential computational time of real-time applications of the proposed model.

4.2. Results and Discussion

4.2.1. Static Application Scenarios. Three static application scenarios with the settings in Table 2 are used. The weighted bus delay is applied as the performance measure as shown in Figure 6. The delay consists of two parts, namely, travel delay and dwelling delay. Travel delay occurs when a bus travels toward the stop at a speed less than the speed limit, and dwelling delay occurs when a bus keeps dwelling after serving boarding and alighting passengers, both of which may be caused by the bus mutual blockage. Figure 6 shows that the weighted bus delays of \( M_{2} \) are noticeably reduced compared with \( M_{0} \) by around 12%–33%. The reduction is the most significant (~33%) in scenario 3 with the largest SD of the service time. Moreover, the travel delays of \( M_{2} \) decrease more remarkably than its dwelling delays in all the scenarios. This is due to the alleviation of the mutual blockage in front of the stop which raises bus travel delay. Even though the dwelling delays in scenario 2 and scenario 3 increase, the total weighted bus delays still decrease significantly, which means that the model can reorganize berth resources to improve the overall efficiency of the bus-dwelling system.

As expected, the weighted bus delays of \( M_{1} \) are lower than those of \( M_{2} \) because the bus punctuality constraints are relaxed in \( M_{1} \). Table 4 shows the schedule delay of \( M_{0} \), \( M_{1} \), and \( M_{2} \) in scenario 3. The schedule delay is introduced to measure the bus punctuality [42], and it is defined as \( t_{out}^{n} - t_{in}^{n} \) different from the weighted bus delay in Equation (15). In \( M_{0} \), all the buses are designed to depart later than the prescheduled departure times. \( M_{1} \) reduces the weighted bus delays by ~40% but at the cost of bus punctuality. Only three buses are on time. In contrast, there are four buses on time in \( M_{2} \). And the reduction of the schedule delay of \( M_{2} \) is larger than that of \( M_{1} \) although \( M_{1} \) has less weight bus delay. As bus punctuality is regarded as the primary objective in this study, \( M_{2} \) outperforms \( M_{1} \).

The berth occupation diagrams of \( M_{0} \), \( M_{1} \), and \( M_{2} \) in scenario 3 shown in Figure 7 demonstrate how the proposed model alleviates mutual blockage by rescheduling the bus arrivals. In Figure 7, the left and right ends of the dark blue boxes are the start and end times of the berth occupation time of a bus. And the left and right ends of the light blue boxes are the start and end times of the service time of a bus, which represents the exact time for passenger boarding and alighting. Comparing Figure 7(b) with Figure 7(a), buses with similar service times (i.e., bus 2 and bus 4, and bus 1 and bus 3) are scheduled to dwell at the same time in \( M_{1} \) to reduce the service time difference of successive buses to relieve mutual blockage. Bus 5 and bus 6 are scheduled to dwell at the same time despite the different service times because no more incoming buses will be blocked by them. Moreover, for two buses dwelling at the same time, the bus with the shorter service time is scheduled to arrive at the stop earlier. For example, bus 2 with a service time of 8 s arrives earlier than bus 4 with a service time of 11 s. In this way, mutual blockage inside the stop is expected to be avoided, which can decrease bus-dwelling delays. A similar phenomenon can be seen in \( M_{2} \) as shown in Figure 7(c). However, different from the schedule in \( M_{1} \), bus 5 is scheduled to arrive earlier than bus 6 in \( M_{2} \) to guarantee the punctuality of both bus 5 and bus 6 (Table 4). The serving occupation rate, defined as the ratio of the total service time and the total berth occupation time, is introduced to measure the effective utilization of the spatiotemporal resources at the bus stop. And a higher rate is preferred. Figure 7 shows that the serving occupation rate is the highest

### Table 2: Parameter settings in three static scenarios.

<table>
<thead>
<tr>
<th>Scenarios</th>
<th>Scenario 1</th>
<th>Scenario 2</th>
<th>Scenario 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bus no. (n)</td>
<td>( S_{n} (s) )</td>
<td>( G_{n} )</td>
<td>( t_{in}^{n} / t_{out}^{n} (s) )</td>
</tr>
<tr>
<td>1</td>
<td>18</td>
<td>17</td>
<td>11/87</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>16</td>
<td>14/108</td>
</tr>
<tr>
<td>3</td>
<td>16</td>
<td>17</td>
<td>15/119</td>
</tr>
<tr>
<td>5</td>
<td>24</td>
<td>15</td>
<td>19/152</td>
</tr>
</tbody>
</table>

SD of \( S_{n} (s) \) | 6.83 | 9.94 | 13.37

Note. \( S_{n} \) is the service time, \( G_{n} \) is the average number of passengers, \( t_{in}^{n} / t_{out}^{n} \) is the earliest/latest arrival time of the arrival time window, and \( t_{out}^{n} \) is the prescheduled departure time.

### Table 3: Average computational time of solving the model.

<table>
<thead>
<tr>
<th>Numbers of incoming buses</th>
<th>Computational time (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>12.6</td>
</tr>
<tr>
<td>2</td>
<td>18.5</td>
</tr>
<tr>
<td>3</td>
<td>26.3</td>
</tr>
<tr>
<td>4</td>
<td>45.4</td>
</tr>
<tr>
<td>5</td>
<td>64.1</td>
</tr>
<tr>
<td>6</td>
<td>77.8</td>
</tr>
</tbody>
</table>
Figure 6: Weighted bus delays.

Table 4: Schedule delays (s).

<table>
<thead>
<tr>
<th>Schedule delays (s)</th>
<th>Bus 1</th>
<th>Bus 2</th>
<th>Bus 3</th>
<th>Bus 4</th>
<th>Bus 5</th>
<th>Bus 6</th>
<th>Average</th>
<th>Reduction (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>M0</td>
<td>310</td>
<td>301</td>
<td>301</td>
<td>401</td>
<td>313</td>
<td>307</td>
<td>22.17</td>
<td>—</td>
</tr>
<tr>
<td>M1</td>
<td>347</td>
<td>On time</td>
<td>On time</td>
<td>327</td>
<td>302</td>
<td>On time</td>
<td>12.67</td>
<td>42.9%</td>
</tr>
<tr>
<td>M2</td>
<td>347</td>
<td>On time</td>
<td>On time</td>
<td>327</td>
<td>On time</td>
<td>On time</td>
<td>12.33</td>
<td>44.4%</td>
</tr>
</tbody>
</table>

Berth 1
- Total serving time: 138 s
- Total berth occupation time: 314 s
- Serving occupation rate: 43.9%

Berth 2
- Bus 2
- Bus 4
- Bus 6

(a)

Figure 7: Continued.
in M1 (∼54%), followed by that in M2 (∼48%). Therefore, M2 can utilize the berth resources much more efficiently on the condition of prioritizing bus punctuality by optimizing bus arrival and departure times and bus-berth matching.

4.2.2. Dynamic Applications to Time-Varying Traffic Conditions. The performance of the proposed model is also evaluated in dynamic applications. Figure 8 shows the weighted bus delays (Figure 8(a)) and the reduced delay values (Figure 8(b)) with different arrival intervals of bus fleets. Figure 8(a) shows that the average weighted bus delays decrease from ∼200 s to ∼50 s with increasing arrival intervals of bus fleets. This is because the bus mutual blockage is more severe when berth resources are intense with short arrival intervals of bus fleets. Similar to the observations in the static application scenarios in Section 4.2.1, Figure 8(a) shows that the weighted bus delay of M0 is the largest, and M1 and M2 can significantly reduce the delay. Due to the relaxation of the bus punctuality constraints, M1 outperforms M2 in terms of the weighted bus delay.

Figure 8(b) shows the reduced values of the weighted bus delays with increasing arrival intervals of bus fleets. M1 outperforms M2 at all the demand levels due to the ignorance of bus punctuality. The reduced weighted bus delays descend when the arrival interval of bus fleets increases from 60 s to 180 s. When the arrival interval of bus fleets is short (e.g., 60 s), the heavy bus traffic indicates more buses are expected to dwell at the stop within a unit of time. As a result, the mutual blockage is more likely to occur and the potential improvement is large. When the arrival interval of bus fleets is large (e.g., 180 s), the low bus traffic indicates fewer buses are considered in each optimization and

![Figure 7: Berth-time occupation diagrams of scenario 3: (a) M0, (b) M1, and (c) M2.](image-url)
consequently less improvement can be achieved. Take the extreme case as an example in which there is only one bus in each optimization. Then, M0, M1, and M2 make no difference.

Figure 9 illustrates the average schedule delays and the numbers of delayed buses among M0, M1, and M2 with different arrival intervals of bus fleets. In all of the cases, M2 has the best performance in reducing average schedule delays and the number of delayed buses compared with M1. The average schedule delay is the highest in M1, which minimizes weighted bus delays at the cost of bus punctuality. Although the reduction of the average schedule delay of M2 is slight, M2 can reduce the number of delayed buses, and the avoidance of one delayed bus is regarded as a remarkable benefit in this study. Therefore, M2 outperforms M1 and M0 in terms of bus punctuality. Moreover, with the increase in the arrival intervals of bus fleets, the difference between the numbers of delayed buses of M2 and M1 decreases. The reason is the low bus traffic as explained in Figure 8(b).

4.3. Sensitivity Analysis

4.3.1. Bus Service Time. The bus service time is one critical parameter. It influences the bus-dwelling time and thus the mutual blockage between successive buses. The sensitivity
analysis of the bus service time is conducted in the dynamic application scenarios to explore its impacts on the performance of the proposed model (M2). The average service time varies from 10 s to 30 s, and the SD varies from 3 s to 18 s following Jiang and Yang [41]. The average arrival interval of bus fleets is 90 s. Other parameters are the same as those in Section 4.1.

Figure 10(a) illustrates the reduced weighted bus delay of M2 compared with M0 with increasing SD of service time. The reduced weighted bus delays rise when the SD of service time increases. Therefore, M2 has the capability of handling the fluctuations of bus service time. Figure 10(b) shows the reduced weighted bus delays with increasing average service time. Generally, the reduced delays show an upward trend with fluctuations. When the average service time is 30 s, M2 gains a much better performance in reducing weighted bus delays. Every bus has to occupy the stop for around 50 s (average service time 30 s + the time for acceleration and deceleration 10 s + the time for traversing two berths 10 s), which is comparable to the average arrival interval of one bus (~51.4 s (the arrival interval of bus fleets 90 s divided by the average fleet size of 1.75 buses)). As a result, there is a

Figure 10: Impacts of service time on weighted bus delay: (a) SD of service time and (b) average service time.

Figure 11: Impacts of service time on schedule delay: (a) SD of service time and (b) average service time.
Figure 12: Impacts of service time on numbers of delayed buses: (a) SD of service time and (b) average service time.

Figure 13: Continued.
high possibility of mutual blockage with the uncertainty of bus fleet arrivals in \textbf{M0}. And the potential advantage of \textbf{M2} in relieving mutual blockage is huge.

Figures 11 and 12 show the reduced schedule delay of \textbf{M2} compared with \textbf{M1} from the perspective of bus punctuality. No noticeable patterns are observed with increasing SD and average of service time. The reduced schedule delay values vary between 1 s and 10 s. Figure 12 shows the reduced numbers of delayed buses of \textbf{M2} compared with \textbf{M1}. Similar to the reduced schedule delay, the reduced numbers of delayed buses are not sensitive to the SD and the average of service time. But \~6 delayed buses can be avoided on average, which indicates remarkable benefits. Figures 9, 11, and 12 indicate bus punctuality is mainly affected by the uncertainty of arrival times of bus fleets rather than bus service time.

**4.3.2. Bus Volumes and Berth Numbers.** The bus volume and the number of berths at the stop have combined effects on the scarcity of berth resources and consequently influence the bus travel delay due to queuing outside the stop. More berths indicate larger stop capacity [24] and are likely to reduce bus travel delays. However, the mutual blockage may occur when the bus volume is high. The sensitivity analysis of the bus volume and the berth number is conducted in the dynamic application scenarios to explore their impacts on the performance of the proposed model (\textbf{M2}). The berth number \( P \) varies from one to four, and the average arrival intervals of bus fleets vary from 60 s to 180 s. Other parameters are the same as those in Section 4.1.

Figure 13 illustrates the reduced weighted bus delays, the reduced schedule delay, and the reduced numbers of delayed buses with different arrival intervals of bus fleets and berth numbers. Note that the model is infeasible with an arrival interval of 60 s and one berth because the corresponding bus demand (105 buses per hour) exceeds the capacity of the one-berth stop, which is \( 3600/S + t_b + 2 \times t_{acc} = 3600/20 + 5 + 2 \times 5 = 102.86 \) buses per hour. As a result, buses cannot arrive at the stop within the arrival time window (constraints Equation (1)). In Figure 13(a), the weighted bus delay reduction shows a decreasing trend with increasing arrival intervals, which is consistent with the results in Figure 8(b). One berth leads to the largest reduction of the weighted bus delays with the arrival intervals larger than 60 s because the benefits of the alleviated mutual blockage inside the stop by reducing berths dominate the negative impacts of the reduced bus-stop capacity. Therefore, a smaller number of berths may be preferred on the condition that the bus demand does not exceed the stop capacity. Figure 13(b) shows the reduction of the average schedule delays does not have noticeable patterns with different arrival intervals of bus fleets. Similar to the observance in Figure 13(a), one berth has the best performance with arrival intervals larger than 60 s. Figure 13(c) shows the reduced numbers of delayed buses have a decreasing trend. When the arrival intervals of bus fleets are large, the low bus traffic indicates the low possibility of mutual blockage both inside and outside the stop. Further, fewer berths help reduce delayed buses with high bus demand, for example, when the arrival interval of bus fleets is 90 s.

**5. Conclusions**

This study proposes an MILP model to optimize bus scheduling of bus arrivals and the bus-berth matching at a curbside bus stop to improve the efficiency of bus services under the connected vehicle environment, in which buses can share real-time location information through communications technology and receive speed advisory from the control center. The objective function of the proposed model
is to minimize bus delays weighted by the number of passengers on buses. Bus arrival times at the stop and the assignment of berths are optimized together with bus departure times from the stop. The prioritizing of bus punctuality is also taken into consideration. The proposed model could be applied dynamically to cater to time-varying traffic conditions. In the numerical studies, the advantages of the proposed model are validated considering both weighted bus delay and bus punctuality in both static and dynamic scenarios. The sensitivity analyses show that (1) the proposed model is robust to the fluctuation of bus service time and (2) a smaller number of berths may be preferred on the condition that the bus demand does not exceed the stop capacity.

No overtaking is allowed due to the linear berth design at the stop in this study. The investigation of curbside stops with drive-through berth designs, angle berth designs, and sawtooth berth designs is planned with the consideration of overtaking. The extension of the proposed model to the coordination between bus scheduling and bus-berth matching at multiple stops with multiple bus lines is another research direction. Accurate prediction of bus arrival time windows and service time is assumed in this study. It is a great challenge to take uncertainty into consideration, especially when there are no bus dedicated lanes and the bus operation is affected by car traffic. In addition, it is a worthy exploration to integrate bus scheduling and bus-berth matching at stops, bus trajectory planning on road segments, and signal timings at intersections.

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

Conflicts of Interest
The views presented in this paper are those of the authors alone. The authors declare that there are no conflicts of interest regarding the publication of this article. A preprint has previously been published [43].

Acknowledgments
This research was supported by the National Natural Science Foundation of China (Nos. 52131204 and 61903276) and Shanghai Science and Technology Innovation Action Plan Project (Nos. 19DZ1209004 and 20DZ1202805).

References


