IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received June 7, 2020, accepted June 21, 2020, date of publication June 25, 2020, date of current version July 22, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3004856

A Prediction Model for Bus Arrival Time at Bus Stop Considering Signal Control and Surrounding Traffic Flow

HU ZHANG[1], SHIDONG LIANG[1], YIN HAN[1], MINGHUI MA[2], AND RONGMENG LENG[1]
[1]Business School, University of Shanghai for Science and Technology, Shanghai 200093, China
[2]School of Mechanical and Automotive Engineering, Shanghai University of Engineering Science, Shanghai 201620, China

Corresponding author: Shidong Liang (sdliang@hotmail.com)

This work was supported in part by the Project funded by China Postdoctoral Science Foundation under Grant 2020M671173, in part by the National Natural Science Foundation of China under Grant 71801149 and Grant 71801153, and in part by the Shanghai Innovation Training Program under Grant XJ2020124 and Grant SH2020069.

ABSTRACT The arrival time of bus at the stop is critical data in design of bus operational strategies. Especially for real-time control strategies (i.e., signal priority system), bus arrival time is usually predicted to assess the next bus operation condition in the future (i.e., bus bunching and reliability of transit service) and then can be used as a decision basis of current control actions. The signalized intersection and surrounding traffic are the key factors in bus travel time prediction, but most previous approaches focus on the impact of signal control on bus delay only. This paper proposes a prediction model for arrival time at bus stop under the influence of both upstream signalized intersection and surrounding traffic flow. Considering the affected range of signalized intersection and the dynamic variation of bus speed, bus running processes are evaluated separately (including processes from a given detection point to stop line, through intersection, and from intersection to bus stop). In the proposed models, bus speed is deduced according to the change of traffic density at different locations to reflect the micro-impact of surrounding traffic flow on bus operation. The observed bus travel time is collected from actual investigations at two bus stops incorporating signalized intersection in Jinan City and compared with the predicted travel time in the proposed model. The results show that the proposed model has a low mean relative error. In addition, through the analysis of the maximum relative error, it also can be seen that vehicle queuing with random arrival of vehicles at stop line makes a gap between the prediction and the actual situation, which will be the focus of further research.

INDEX TERMS Public transit, bus arrival time, signalized intersection, bus speed, traffic density.

I. INTRODUCTION

Numerous intelligent transportation technologies have been widely applied in public transportation system recently. This enables an access to the real-time acquisition of information, such as vehicle location, bus speed, and number of passengers, which facilitates transit management. Meanwhile, because of the randomness of traffic flow and the influence of signal control at intersection, the travel time of buses fluctuates randomly, which reduces the stability of public transport. In order to further improve the level of bus service, based on these real-time information, a series of operational strategies is proposed for increasing the accuracy of timetable operation and the equilibrium of bus headway [1]–[3]. Then passengers can perceive the arrival time of buses well and the stability of passenger waiting time can be guaranteed. Therefore, as an important part of real-time control strategy, it is necessary to obtain accurate real-time bus arrival time. Based on the existing researches, prediction methods of bus travel time or arrival time can be classified as machine learning approaches, regression models and surrounding traffic based approaches.

The machine learning approaches are effective for predicting travel time of bus, in which there is a non-linear relationship between the bus travel time and the surrounding information which are acquired from inductive loops, GPS, etc. [4]–[8]. Amita *et al.* use Artificial Neural Network (ANN) to predict the bus travel time based on GPS data [9]. The number of boarding and alighting passengers and number of dwells are taken into account. Ma *et al.* propose a novel segment-based approach and predict the bus travel time and dwelling time using Support Vector

The associate editor coordinating the review of this manuscript and approving it for publication was Bohui Wang.

Machine (SVM), in which the required traffic data are derived from the real-time taxi and bus datasets, taking time period and route re-segmentation into account [10]. Yu *et al.* develop a bus arrival time prediction model for buses on multiple routes and divide bus travel time into several parts that are predicted using SVM, ANN, k-NN and LR, respectively [11]. Meanwhile, the effects of these methods are compared.

The regression models usually use statistical data to fit the bus travel time [12]. Bian *et al.* propose a compound Poisson distribution for estimating bus dwelling time at a curbside bus stop, considering the relationship between the number of boarding passengers and dwelling time [13]. Guenthner and Hamat analyze the arrival rule of multi-line vehicles in the bus network and find that the arrival time at the terminal obeys the Gamma distribution [14]. Zhang *et al.* and Liu *et al.* use a normal distribution to simulate travel time between bus stops to reflect the randomness of bus travel time [15], [16].

Both the machine learning approaches and regression models focus on the macroscopic travel time between stops, which were hardly related to microcosmic impact of a signalized intersection or surrounding traffic flow on bus travel time. Thus, the surrounding traffic based approaches are proposed.

The surrounding traffic based approaches develop relationships between the bus spatio-temporal operation and signalized intersection or surrounding traffic flow, then the travel time is derived. Chow *et al.* propose a kinematic wave model to display the space-time trajectory of buses within general traffic and calculate the bus delay under the adjustment of red light at intersection [17]. Liu *et al.* consider the effect of vehicle queue at signal intersection and calculate the travel time from the inductive loop to the stop line, including the waiting times for green light and queue clearance [18]. With the deepening of research, based on GPS data, Bie *et al.* predict the bus arrival time at a signalized intersection. The prediction model is built in various cases that depend on the signal phase when buses arrive at the detection zone [19].

II. OBJECTIVES AND CONTRIBUTIONS

In view of the need of real-time control strategy based on micro process simulation, it is worth predicting or evaluating bus arrival time at stops in complex traffic conditions. In order to improve the accuracy of prediction, the detailed running process of buses should be described and the space-time trajectory of buses influenced by various factors should also be clarified. However, existing surrounding traffic based approaches have shortcomings in describing the change of bus speed within running process of buses, in which bus speed is either zero or a fixed free-flow speed. According to the running process at intersection, that bus speed is zero expresses the waiting state in a vehicle queue and that bus speed is a fixed free-flow speed expresses a running condition after the queue has been cleared, while the transition from running to stopping and the influence of traffic flow are neglected. Meanwhile, the bus travel time from intersection to bus stop is ignored.

The objective of this research is to propose a prediction model for bus arrival time at bus stop under the influence of both upstream signalized intersection and surrounding traffic flow. The contributions of this paper are as follows:

1) Traffic density at different locations is evaluated to realize the linear change of speed based on the relationship between bus speed and traffic density, rather than a transformation between zero and a fixed free-flow speed.

2) Based on the multi-stage division of bus operating process, within each section, the microcosmic effects of different traffic flow on bus speed are considered.

The remainder of this paper is organized as follows: Section 3 presents the prediction framework and the formulation of corresponding proposed models; Section 4 provides a case study to test the prediction model in the real world scenarios in Jinan; Section 5 presents discussion, and Section 6 provides the conclusion and the further research.

III. MATHEMATICAL MODEL

A. PREDICTION FRAMEWORK

All input variables of the method are declared in Table 1.

TABLE 1. Input variable declaration.

Variables	Declaration
t_D	The recorded arrive time at the detection point
T_R	The time of red light
T_i^G	The green time in the current cycle i
T_i^R	The red time in the current cycle i
t_i^G	Starting time of green light
t_i^R	Starting time of red light
L_l	The distance between detection point and stop line
q_l	The traffic flow in the lane where the bus is located
N	The number of lanes that can accept buses
v_f	Free flow speed without any interference
K_j	Jam density
q_s	The saturation discharge rate
L_i	The geometric length of intersection that is the distance from the stop line to the exit lane
q_r	The traffic flow turning right into the exit lanes
L_s	The distance between the intersection and the bus stop

The assumptions concerning the formulation of the model are as follows:

(1) The signal phase is fixed, that is, the time of signal lights within each control cycle is the same.

(2) The speed of buses during the acceleration and deceleration depends on the density of the surrounding traffic and when the jam density occurs, the bus speed is 0.

(3) Flow rate of surrounding traffic is uniform throughout the prediction process.

(4) When the red light is on, the vehicles begin to queue, while there is no queue before the red light.

When bus arrives at an intersection, to obtain the arrival time at the downstream bus stop, it is necessary to determine the location of bus at the intersection and estimate the travel time between location and bus stop. Fig. 1 shows the running process of buses before arriving at the bus stop and the prediction of bus arrival time at a bus stop.

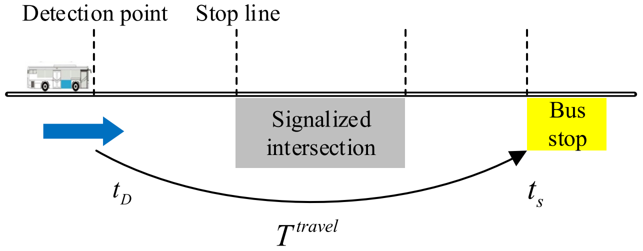


FIGURE 1. Prediction of arrival time at bus stop with traffic control.

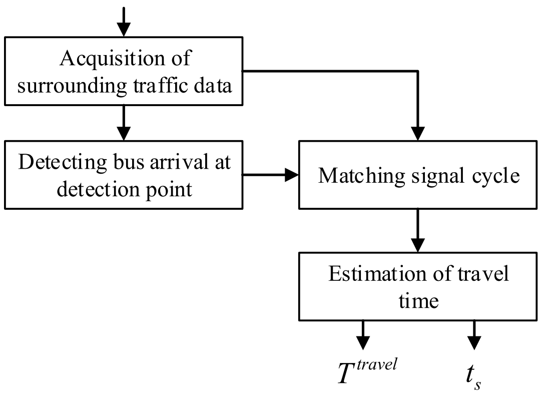


FIGURE 2. Flow chart for the prediction process.

There is a detection point in the approach lane of the intersection to record the arrive time at the intersection. Then, based on arrival time at the detection point, signal control and traffic flow, the travel time from the detection point to the downstream bus stop can be estimated. Hence, the arrival time at the downstream bus stop can be obtained following the running process of buses.

$$t_s = t_D + T^{travel} \tag{1}$$

$$L_D > T_R \cdot q_e \cdot \Delta h/N + \Delta L \tag{2}$$

where t_D is the recorded time when bus arrive at the detection point, which is a current time stamp; t_s is the predicted arrival time at the downstream bus stop; T^{travel} is the estimated travel time from the detection point to the downstream bus stop; and L_D is the distance between detection point and stop line of intersection. In order to ensure the accuracy of prediction, the location of detection point should be far away from queuing vehicles at intersections following the Eq. (2). q_e and N are total vehicle arrival rate and the number of lanes that can accept buses, respectively. Δh is average length per vehicle. $T_R \cdot q_e \cdot \Delta h/N$ expresses the length of vehicle queue during red light, and ΔL is a buffer distance between the location of detection point and the maximum queue point, which is a constant and can be set to 2∼4m.

Fig. 2 show a flow chart for the prediction process. When bus arrives at the detection point, the arrival time t_D at detection point and the current signal phase are determined, simultaneously. In this research, green time in the current cycle i is defined as T_i^G and the red time is defined as T_i^R. t_i^G and t_i^R are starting time of green light and red light, respectively. Fig. 3 shows a relationship between the arrival time t_D and the signal phase which can be used to express the spatio-temporal trajectories of buses and predict the bus travel time under the influence of signal control.

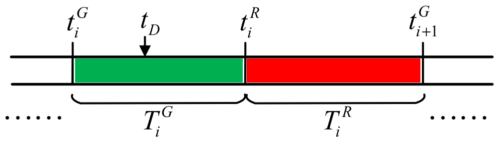


FIGURE 3. Relationship between the arrival time at detection point and the signal phase.

According to running process of buses before arriving at the bus stop in Fig. 1, after the arrival time is recorded at the detection point, bus will be heading for the stop line under the influence of signal control, cross the intersection, and then pass a road section before the bus stop. Hence, travel time from the detection point to the downstream bus stop can be divided into three parts: i) travel time (T_l) from the detection point to the stop line; ii) travel time (T_i) of crossing the intersection; iii) travel time (T_s) from the intersection to the bus stop. Thus, the travel time (T^{travel}) in Eq. (1) can be expressed by

$$T^{travel} = T_l + T_i + T_s \tag{3}$$

In addition to the influence of signal control, these three travel time are related to traffic flow, in which traffic density at different locations will cause corresponding speed changes. Hence, considering signal control and surrounding traffic flow, travel time (T_l, T_i, and T_s) can be illustrated as follows.

B. TRAVEL TIME FROM THE DETECTION POINT TO THE STOP LINE

Given a signal control, the estimation of travel time from detection point to stop line show a complexity. Bie *et al.* set up a time window to illustrate the relationship between travel time and current signal phase [19]. However, the running process of the vehicle is simplified, in which acceleration and deceleration of buses are neglected. In fact, the bus will slow down slowly until it stops when the current vehicle starts queuing for the green light, i.e., traffic density is high enough. Therefore, as the bus moves forward, the bus speed should be constantly updated based on the relationship between vehicle speed and traffic density.

Under the influence of signal control and surrounding traffic flow, bus may arrive directly at the stop line or need to queue up to the stop line. The vehicle queuing case is discussed first, in which arrival time with the mean speed of surrounding traffic but without signal control at stop line (denoted as v_o) is greater than the starting time of red light, i.e., $t_D + \frac{L_l}{v_o} \geq t_i^R$. Thus, in this case, the travel time from the detection point to the stop line can be expressed by

$$T_l = \max[t_i^R - t_D,\ 0] + \int_{L_1}^{L_1+L_2} \frac{dl}{v_{L_2}(l)} + \Delta T + \int_{L_1+L_2}^{L_l} \frac{dl}{v_{L_l}(l)} \tag{4}$$

where L_1 is distance between the detection point and the location of bus at the beginning of red light; L_2 is running distance from the location of bus at the beginning of red light to the end of queuing vehicles; and ΔT is waiting time in vehicle queue.

The first term in Eq. (4) denotes the bus running time unaffected by red light that can be determined by the difference between starting time of red light t_i^R and arrival time at detection point t_D. In addition, it is equal to zero if bus arrives at the detection point during red time, i.e., $t_i^R < t_D$. If $t_i^R > t_D$, the bus is assumed to be travelling with the mean speed of surrounding traffic. Hence, the distance L_1 can be expressed by

$$L_1 = v_o \cdot \max[t_i^R - t_D,\ 0] \tag{5}$$

Based on linear model proposed by Greenshields in traffic flow theory [20], the relationship between the traffic flow q_l, density K and speed v_o is expressed as.

$$q_l = v_o K \tag{6}$$

Speed-density relationship model proposed by Greenshields is

$$v_o = v_f \cdot [1 - \frac{K}{K_j}] \tag{7}$$

Thus, the mean speed of surrounding traffic v_o can be derived as $\frac{v_f}{2} \pm \sqrt{(v_f/2)^2 - \frac{q_l}{K_j} \cdot v_f}$, in which "±" is related to crowded and non-crowded traffic conditions. "−" is adopted under crowded condition, while "+" is adopted under non-crowded condition. In this paper, according to the vehicle running process before the beginning of red light, vehicles can travel freely and it presents a non-crowded traffic condition. Therefore, "+" is adopted in the above expression. Thus, the mean speed of surrounding traffic v_o is derived as

$$v_o = \frac{v_f}{2} + \sqrt{(v_f/2)^2 - \frac{q_l}{K_j} \cdot v_f} \tag{8}$$

Meanwhile, q_l can be expressed by

$$q_l = \frac{q_e}{N} \tag{9}$$

The second term in Eq. (4) denotes running time of bus after the beginning of red light until it reaches the end of the queue, in which bus decelerates to the vehicle queue at speed v_o. $L_l - L_1$ is distance between the location of bus at the beginning of red light and the stop line. Before the red light, vehicles running within this distance $(L_l - L_1)$ at speed v_o will form the queue in front of the bus. Hence, L_2 is the part of L_l out of L_1 and queue length. Therefore, L_2 can be calculated by

$$L_2 = L_l - (L_1 + \frac{L_l - L_1}{v_o} \cdot q_l \cdot \Delta h) \tag{10}$$

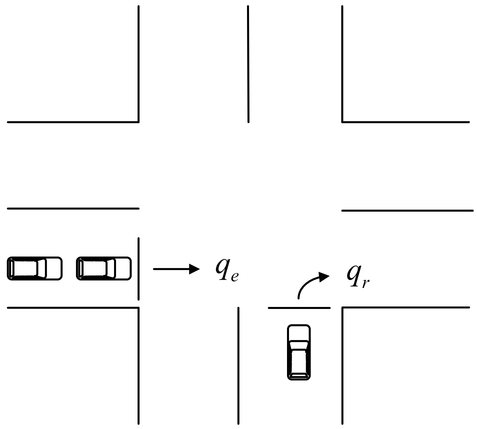


FIGURE 4. Composition of mixed traffic flow.

where $\frac{L_l - L_1}{v_o} \cdot q_l$ denotes the number of vehicles that need to stop and queue up after the beginning of red light.

The variable $v_{L_2}(l)$ in Eq. (4) is a function that presents the relationship between speed and distance. Based on speed-density model, $v_{L_2}(l)$ can be derived as

$$v_{L_2}(l) = v_f \cdot [1 - \frac{k(l)}{K_j}] \tag{11}$$

where $k(l)$ is traffic flow density at different locations l. Parameter l is a continuous variable with a range from the location of bus at the beginning of red light to the end of queuing vehicles, i.e., $l \in [L_1,\ L_1 + L_2]$. When the red light is on, the number of queuing vehicles in front of the bus is $\frac{L_l - L_1}{v_o} \cdot q_l$. Meanwhile, the distance between the stop line and the location of bus is $L_l - l$. Density $k(l)$ can be expressed by the ratio of the two, as the following Equation.

$$k(l) = \frac{\frac{L_l - L_1}{v_o} \cdot q_l}{L_l - l} l \in [L_1,\ L_1 + L_2] \tag{12}$$

Following Eq. (11) and (12), the relationship between bus speed and locations l can be derived as

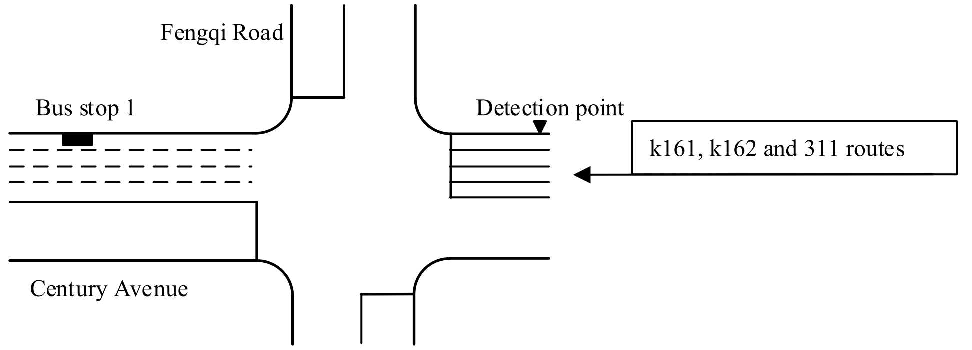
$$v_{L_2}(l) = v_f \cdot (1 - \frac{\frac{L_l - L_1}{v_o} \cdot q_l}{L_l - l} \cdot \frac{1}{K_j}) l \in [L_1,\ L_1 + L_2] \tag{13}$$

The third term ΔT in Eq. (4) presents bus waiting time in the queue and consists of waiting time during red light and queue dissipation time, in which waiting time during the red light is the time from the starting queue of the bus to the end of the red light and the queue dissipation time is the time from the end of the red light (the start of green light) to the departure of queuing vehicles in front of the bus. ΔT can be calculated as follow

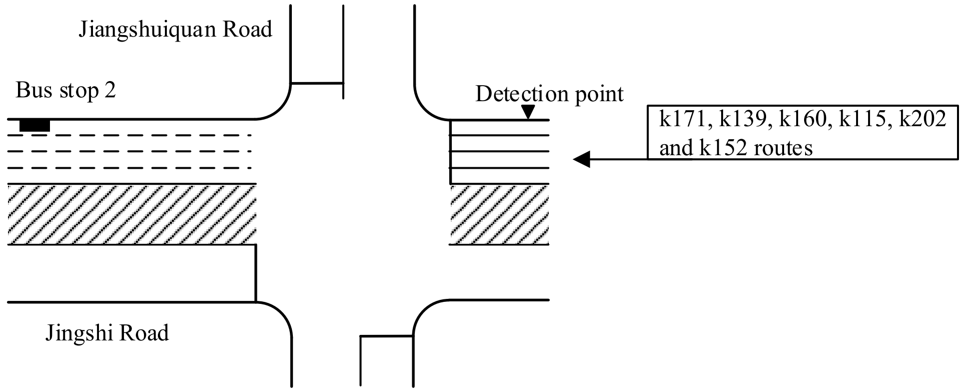
$$\Delta T = T_i^R + \frac{L_l - L_1}{v_o} \cdot q_l / q_s \tag{14}$$

The fourth term in Eq. (4) denotes running time with saturation discharge rate. Following Eq. (8), the speed $v_{L_l}(l)$ in Eq. (4) can be expressed as

$$v_{L_l} = \frac{v_f}{2} + \sqrt{(v_f/2)^2 - \frac{q_s}{K_j} \cdot v_f} \tag{15}$$



(a) Century avenue fengqi road (bus stop) incorporating Century Avenue and Fengqi Road intersection



(b) Jingshi road jiangshuiquan road (bus stop) incorporating Jingshi Road and Jiangshuiquan Road intersection

FIGURE 5. Investigated bus stops incorporation intersection.

The above is the analysis of vehicle queuing case. Next is the analysis of the situation that there is no queuing of vehicles, in which bus runs from the detection point to the stop line during the green light period, i.e., $t_D + \frac{L_l}{v_o} < t_i^R$. Thus, in this case, the bus speed is v_0 and the travel time is

$$T_l = \frac{L_l}{v_o} \tag{16}$$

C. CROSSING TIME THROUGH THE INTERSECTION

Bus travel time of crossing the intersection is related to cases of bus arriving at stop line. When a vehicle is inside an intersection, it usually does not change lanes. If bus passes through the stop line in the form of saturated discharge, i.e., $t_D + \frac{L_l}{v_o} \geq t_i^R$, the bus speed of crossing the intersection can be set to $v_{L_l}(l)$. Meanwhile, if bus passes through the stop line with the mean speed of surrounding traffic but without signal control, i.e., $t_D + \frac{L_l}{v_o} < t_i^R$, the bus speed can be set to v_o. Hence, the travel time of crossing the intersection can be expressed by

$$T_i = \begin{cases} \dfrac{L_i}{v_{L_l}}, & t_D + \dfrac{L_l}{v_o} \geq t_i^R \\ \dfrac{L_i}{v_o}, & t_D + \dfrac{L_l}{v_o} < t_i^R \end{cases} \tag{17}$$

In addition, Eq. (14) express the crossing time through the intersection from the stop line, which is based on the current signal status. And t_i^R in Eq. (17) is the starting time of red light in the current cycle i when bus arrives at the detection point. Thus, the value of t_i^R is consistent with the above prediction of travel time from the detection point to the stop line.

D. TRAVEL TIME FROM THE INTERSECTION TO THE BUS STOP

After bus departs from the intersection, the bus will be mixed with traffic flow in other directions. Fig. 4 shows the surrounding mixed traffic flow.

According to surrounding mixed traffic flow, following Eq. (8), the bus speed from the intersection to the bus stop can be calculated as

$$v_{L_s} = \frac{v_f}{2} + \sqrt{(v_f/2)^2 - \frac{q_b}{K_j} \cdot v_f} \tag{18}$$

and

$$q_b = \frac{q_e + q_r}{N_b} \tag{19}$$

where N_b is the number of lanes in mixed section and q_b is the traffic flow in the lane where the bus is located within the mixed traffic flow.

Given the distance L_s between the intersection and the bus stop, travel time from the intersection to the bus stop can be calculated as

$$T_s = \frac{L_s}{v_{L_s}} \tag{20}$$

IV. CASE STUDY

A. EXPERIMENTAL SURROUNDINGS

The proposed prediction model in part 2 is applied to the real world scenarios in Jinan, China. We choose two bus stops on different routes: (a) Century avenue fengqi road (bus stop 1) incorporating Century Avenue and Fengqi

Road intersection; and (b) Jingshi road jiangshuiquan road (bus stop 2) incorporating Jingshi Road and Jiangshuiquan Road intersection, as shown in Fig.5. Buses of k161, k162 and 311 routes pass through Century Avenue and Fengqi Road intersection and arrive at the bus stop 1. And six bus routes operating the bus stop 2 incorporating intersection includes k171, k139, k160, k115, k202 and k152 routes.

To test the validity of bus arrival time prediction model, the surrounding traffic data and the bus arrival and travel time should be collected. The surrounding traffic data in case study include the green time and red time in a signal cycle, dimension of intersection, arrival rate and speed of surrounding traffic. Among them, arrival rate and speed data of surrounding traffic is used to count the average values and then the density can be derived. Based on the relationship between traffic flow, speed and density, the speed of the bus in the surrounding traffic at different locations can be calculated. Combined with the arrival time of the bus at detection point, the arrival time at downstream stop can be predicted and then is compared with observed arrival time. Hence, the video surveys are carried out to collect the bus sequence, traffic flow in the approach lane of the intersection, green and red time of signal control, arrival time at detection point and so on. The buses are marked through the route number and the tail number of the license plate, and then bus sequence is obtained. There are two positions of the cameras including detection point and stop line. To obtain the traffic information in the downstream path, an investigator is focusing on the starting time of red light when the test starts. Based on the fixed signal phase, after the arrive time of following bus is recorded, the difference between the arrive time of first detected bus at the detection point and the starting time of red light can be inferred. Difference value between the arrive time t_D and starting time of red light t_i^R in the current cycle when buses are detected at the detection point are presented in Table 2. The downstream traffic flow during the current cycle, including q_e and q_b, when the bus is recorded to calculate the traffic density and speed in the downstream path.

Table 3 shows parameters of two investigated bus stops incorporating intersection, including the signal time, dimension of intersection, free flow speed, and jam density. 100 cycles at intersections 1 and 50 cycles at intersections 2 are observed. According to observation of queue length, maximum queue lengths appeared in 7 cycles at intersections 1 and 3 cycles at intersections 2. The maximum queue lengths at intersections 1 and 2 are 48m (8 vehicles) and 66m (11 vehicles), respectively. To ensure the accuracy of detection, the detection point is set away from the end of the queue. Based on the maximum queue length, a buffer distance ΔL is added and set to 2m. Thus, distances between the detection point and the stop line for two intersections are set to 50m and 68m. The surveyed buses have the same size and performance. Note that all surveys are carried out during the morning peak (7:30 – 9:30).

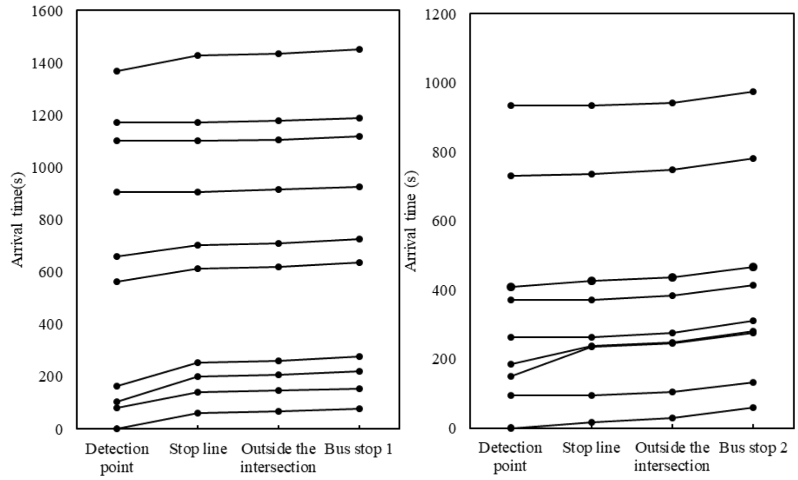
In the survey, the arrival time of 29 and 28 buses at two bus stops are recorded, respectively. The average travel time from detection point to bus stop 1 is 65.8s, while, related to bus stop 2, the average travel time is 71.4s. In order to show the space-time trajectory of buses clearly, a part of the observed buses is intercepted from all buses. Fig. 6 shows the observed arrival time of the intercepted buses at the stops 1 and 2,

TABLE 2. Difference value between the arrive time t_D and starting time of red light t_i^R in the current cycle.

Bus No.	Difference value(s)	Bus No.	Difference value(s)
Stop 1			
1	1	16	3
2	6	17	3
3	1	18	9
4	3	19	2
5	5	20	2
6	2	21	1
7	2	22	7
8	3	23	2
9	9	24	2
10	3	25	5
11	3	26	8
12	10	27	3
13	2	28	1
14	21	29	17
15	1		
Stop 2			
Bus No.	Difference value(s)	Bus No.	Difference value(s)
1	9	15	0
2	0	16	12
3	10	17	2
4	8	18	8
5	2	19	1
6	7	20	18
7	2	21	3
8	1	22	8
9	11	23	2
10	4	24	7
11	0	25	1
12	9	26	10
13	5	27	0
14	14	28	9

TABLE 3. Parameters of two investigated bus stops.

Parameters	Bus stop 1	Bus stop 2
L_l (m)	50	68
L_i (m)	65	60
L_s (m)	100	253
T^R (s)	57	63
T^G (s)	26	29
v_f (m/s)	14.3	15.3
K_j (pcu/m)	0.14	0.15
Δh (m)	6	



(a) Bus stop 1 (b) Bus stop 2

FIGURE 6. **The observed arrival time at bus stops 1 and 2.**

in which each line represents the space-time trajectory of a vehicle. Among them, four arrival time (i.e., at detection point, stop line, outside the intersection and bus stop) are recorded and the corresponding travel time of each section can be deduced. It can be seen that the crossing time through the intersection (T_l) and the travel time from the intersection to the bus stop (T_s) are relatively consistent, while the travel time of different buses from the detection point to the stop line (T_l) varies greatly because of the influence of signal control.

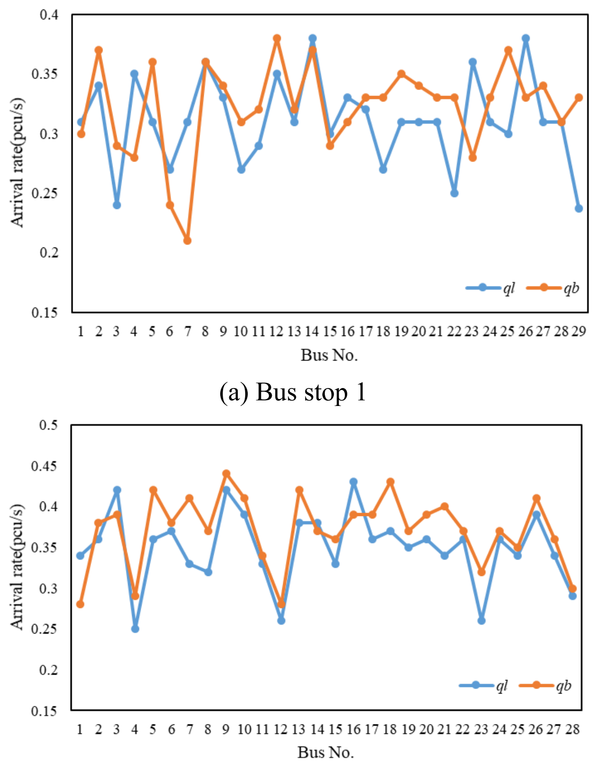
B. EXPERIMENTAL RESULTS AND ANALYSIS

When bus arrives at the detection point, the arrival time at the downstream bus stop is predicted through the proposed model. To test the model performance, the observed and predicted travel time are compared and the numerical difference should be measured. The mean absolute error (E_{MAE}) indicates the average numerical difference. The mean relative error (E_{MRE}) indicates the average level of error to offset dimensional influence. The maximum relative error (E_{MaxRE}) presents a maximum extent of prediction error. Among them, both the mean absolute error (E_{MAE}) and the mean relative error (E_{MRE}) reflect the prediction accuracy, which can be used to judge whether the model is effective. The maximum relative error (E_{MaxRE}) can present the limitation of the prediction model. Therefore, these three error measurements are collected to test the model performance and can be calculated as follows:

$$E_{MAE} = \frac{1}{M} \sum \left| T_m^{travel} - \hat{T}_m^{travel} \right| \tag{21}$$

$$E_{MRE} = \frac{1}{M} \sum \frac{\left| T_m^{travel} - \hat{T}_m^{travel} \right|}{T_m^{travel}} \tag{22}$$

$$E_{MaxRE} = \max_{m=1,2,\cdots,M} \left[\frac{\left| T_m^{travel} - \hat{T}_m^{travel} \right|}{T_m^{travel}} \right] \tag{23}$$

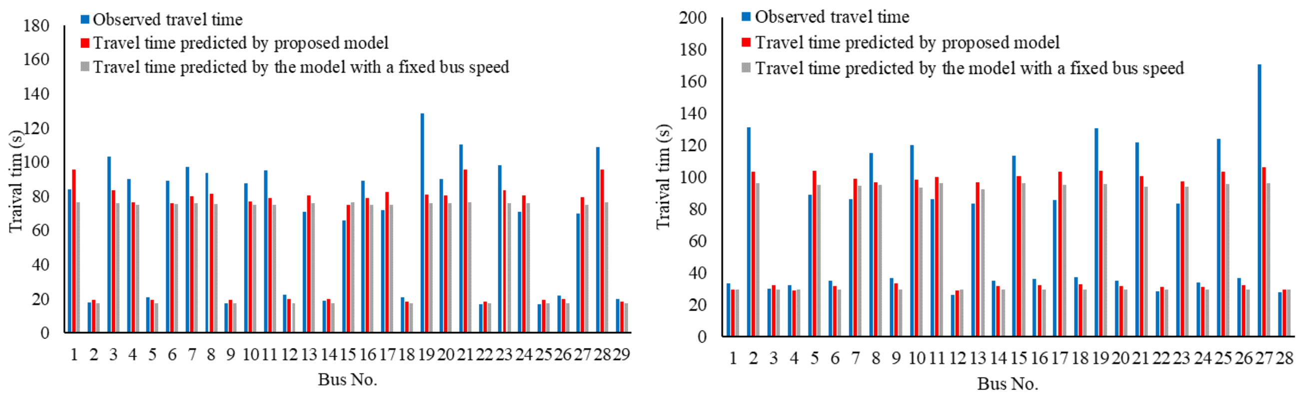


(b) Bus stop 2

FIGURE 7. **The traffic flows q_l and q_b in each cycle when the bus is detected.**

where M is the number of the observed buses; T_m^{travel} is the observed travel time of bus m; and $\hat{T}_m^{travel}$ is the predicted travel time of bus m.

One of the main contributions of this paper is to realize the linear change of speed to improve the accuracy of prediction. Thus, two prediction method are tested: (1) the proposed model, and (2) the model with a fixed bus speed. In the model with a fixed bus speed, the bus speed is between zero



(a) Bus stop 1 (b) Bus stop 2

FIGURE 8. Comparison of the observed and predicted travel time.

and an average speed (12.4m/s at stop 1 and 13.1 m/s at stop 2), in which bus speed is zero expresses the waiting state in a vehicle queue and bus speed is a fixed average speed expresses a running condition after the queue has been cleared. Under the same conditions, the parameters at two stops input the two predicted models. The traffic flows q_l and q_b which are inputted to predicted the travel time are shown in Fig. 7 in the revised version. Figure 8 shows a comparison of the observed travel time, travel time predicted by proposed model, and travel time predicted by the model with a fixed bus speed. Figs. 8(a) and 8(b) present the travel time of buses that arrive at bus stops 1 and 2, respectively. It can be seen that the travel times predicted by the model with a fixed bus speed are smaller than those predicted by proposed model. The reason may be that the model with a fixed bus speed does not consider reduction of bus speed the transition from running to stopping for waiting in a queue, and, in addition, there is a decrease in bus speed under the influence of downstream traffic flow.

From the two figures, the predicted travel time can follow the change of the observed travel time, while there is a difference between them. Table 4 shows the performance of the two prediction model. The results summarized in Table 4 show that the proposed model has much fewer average and maximum estimation errors compared with the model with a fixed bus speed. This performance improvement may be attributed to the estimation of the linear change of speed and the multi-stage division of bus operating process that reflects the microcosmic effects of different traffic flow on bus speed.

According to the mean absolute errors, it can be found that although the bus operation process is divided and simulated in detail, there are still errors. The reason may be that the linear change of bus speed can not completely simulate the randomness of operation of buses and surrounding traffic. Nevertheless, the mean relative errors are small and acceptable, which means that the prediction model is effective. Meanwhile, compared with mean relative errors, the maximum relative errors reach 0.37 and 0.38 at two stops. Excluding accidental factors, the maximum error may be caused by factors not considered. From survey data, there may be a big deviation

TABLE 4. Performance of the prediction model.

	Bus stop 1	
	Proposed model	Model with a fixed bus speed
E_{MAE} (s)	9.99	12.25
E_{MRE}	0.13	0.15
E_{MaxRE}	0.37	0.41
	Bus stop 2	
	Proposed model	Model with a fixed bus speed
E_{MAE} (s)	12.34	13.68
E_{MRE}	0.14	0.16
E_{MaxRE}	0.38	0.44

due to the randomness of queue dissipation time when vehicle queue occurs, accompanied by long travel time. But this is not absolute. This method is available as long as the vehicle queue does not exceed the detection point.

In order to reflect the influence of vehicle queuing on travel time prediction, the travel time with and without vehicle queue is extracted from the survey data and the relative errors are shown in Fig. 9, which reflects that the relative errors with a vehicle queue are generally larger than that without vehicle queue, while the relative error with vehicle queue can still be accepted. It shows that although this method can well simulate the process of vehicle queuing under the condition that the arrival rate of vehicles is a certain value, there is still a gap with the actual situation that surrounding vehicle arrival at stop line is random. Therefore, overcoming the influence of randomness on vehicle queuing can further improve the prediction accuracy.

V. DISCUSSION

This paper proposes a short-term prediction method for bus arrival time from the signalized intersection to the downstream bus stop. Under the influence of signal control, surrounding traffic and vehicle queue, the operation process of buses presents complexity. As a part of modular research, this paper focuses on the accuracy of method of simulating the bus operation process. On the premise of assumption that

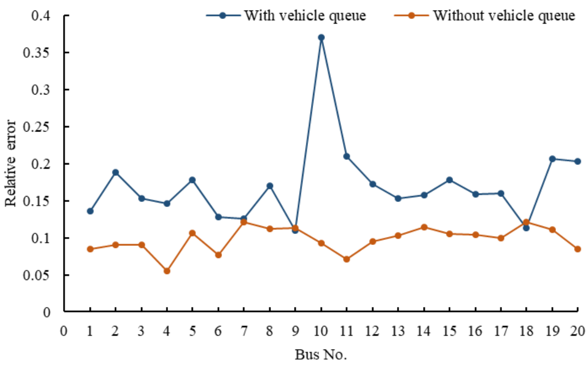


FIGURE 9. Comparison of the relative error of travel time with and without vehicle queue.

the signal phase is fixed and queued vehicles are released with a saturated flow rate, experimental results show that the proposed prediction model has a small error which can be accepted. However, in fact, during the practical applications, there is an inevitable time delay because of the information acquisition and processing, i.e., coordination between systems and communication delay [21], [22]. In this case, the accuracy of the prediction model will decrease. In addition, the proposed prediction methods can also be extended to non-signalized intersection, in which the signal light is all green that the parameter red time set to zero.

VI. CONCLUSIONS

This paper presents a bus arrival time prediction model which takes into account the influence of signalized intersection. Based on running process between the intersection and the downstream bus stop, the arrival times are segmented and predicted by three models, which include travel time from detection point to stop line, crossing time through the intersection, and travel time from intersection to bus stop. Meanwhile, to reflect the change of bus speed under the surrounding traffic flow, the speed is constantly updated according to the traffic density at different locations. To test the performance of the proposed prediction models, two bus stops incorporating signalized intersection are chosen as the object of investigation. The video and on-board surveys are carried out to collect observed travel time and traffic data. Results show that mean relative errors do not exceed 0.2 and models can predict travel time well. However, it can be found that vehicle queue in front of stop line affects the accuracy of prediction because of the randomness of traffic flow. Therefore, further research will focus on the travel time estimation with stochastic traffic flow to improve the accuracy of real-time prediction. In addition, in view of the complexity and importance of signal control, method of networked signal control or adaptive signal control will also be given attention based on the existing researches [23].

ACKNOWLEDGMENT

The authors confirm that there are no conflicts of interest regarding the publication of this manuscript. All data included in this study is available upon request by contact with the corresponding author.

REFERENCES

[1] Y. Gu, X. Fu, Z. Liu, X. Xu, and A. Chen, "Performance of transportation network under perturbations: Reliability, vulnerability, and resilience," *Transp. Res. E, Logistics Transp. Rev.*, vol. 133, Jan. 2020, Art. no. 101809.

[2] C. E. Cortés, D. Sáez, F. Milla, A. Nañez, and M. Riquelme, "Hybrid predictive control for real-time optimization of public transport systems' operations based on evolutionary multi-objective optimization," *Transp. Res. C, Emerg. Technol.*, vol. 18, no. 5, pp. 757–769, Oct. 2010.

[3] D. Sáez, C. E. Cortés, F. Milla, A. Nñez, A. Tirachini, and M. Riquelme, "Hybrid predictive control strategy for a public transport system with uncertain demand," *Transportmetrica*, vol. 8, no. 1, pp. 61–86, Jan. 2012.

[4] H. Chang, D. Park, S. Lee, H. Lee, and S. Baek, "Dynamic multi-interval bus travel time prediction using bus transit data," *Transportmetrica*, vol. 6, no. 1, pp. 19–38, Jan. 2010.

[5] S. I.-J. Chien, Y. Ding, and C. Wei, "Dynamic bus arrival time prediction with artificial neural networks," *J. Transp. Eng.*, vol. 128, no. 5, pp. 429–438, Sep. 2002.

[6] Y. Ding and I. J. Chien, "The prediction of bus arrival times with link-based artificial neural networks," in *Proc. Int. Conf. Comput. Intell. Neurosci.* 2000, pp. 730–733.

[7] Z. Wall and D. J. Dailey, "An algorithm for predicting the arrival time of mass transit vehicles using automatic vehicle location data," in *Proc. 78th Annu. Meeting Transp. Res. Board*, 1999, pp. 1–11.

[8] S. Liang, M. Ma, S. He, H. Zhang, and Z. Tang, "Influence of bus stop location upon traffic flow," in *Proc. Municipal Eng.*, 2019, pp. 1–22.

[9] J. Amita, S. S. Jain, and P. K. Garg, "Prediction of bus travel time using ANN: A case study in delhi," *Transp. Res. Procedia*, vol. 17, pp. 263–272, Oct. 2016.

[10] J. Ma, J. Chan, G. Ristanoski, S. Rajasegarar, and C. Leckie, "Bus travel time prediction with real-time traffic information," *Transp. Res. C, Emerg. Technol.*, vol. 105, pp. 536–549, Aug. 2019.

[11] B. Yu, W. H. K. Lam, and M. L. Tam, "Bus arrival time prediction at bus stop with multiple routes," *Transp. Res. C, Emerg. Technol.*, vol. 19, no. 6, pp. 1157–1170, Dec. 2011.

[12] J. Patnaik, S. Chien, and A. Bladikas, "Estimation of bus arrival times using APC data," *J. Public Transp.*, vol. 7, no. 1, pp. 1–20, Mar. 2004.

[13] B. Bian, N. Zhu, S. Ling, and S. Ma, "Bus service time estimation model for a curbside bus stop," *Transp. Res. C, Emerg. Technol.*, vol. 57, pp. 103–121, Aug. 2015.

[14] R. P. Guenthner and K. Hamat, "Distribution of bus transit on-time performance," *Transp. Res. Rec. J. Transp. Res. Board*, vol. 1202, pp. 1–8, 1988.

[15] H. Zhang, S. Zhao, Y. Cao, H. Liu, and S. Liang, "Real-time integrated limited-stop and short-turning bus control with stochastic travel time," *J. Adv. Transp.*, vol. 2017, pp. 1–9, 2017.

[16] Z. Liu, Y. Yan, X. Qu, and Y. Zhang, "Bus stop-skipping scheme with random travel time," *Transp. Res. C, Emerg. Technol.*, vol. 35, pp. 46–56, Oct. 2013.

[17] A. H. F. Chow, S. Li, and R. Zhong, "Multi-objective optimal control formulations for bus service reliability with traffic signals," *Transp. Res. B, Methodol.*, vol. 103, pp. 248–268, Sep. 2017.

[18] H. Liu, A. Skabardonis, W.-B. Zhang, and M. Li, "Optimal detector location for bus signal priority," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 1867, no. 1, pp. 144–150, Jan. 2004.

[19] Y. Bie, D. Wang, and H. Qi, "Prediction model of bus arrival time at signalized intersection using GPS data," *J. Transp. Eng.*, vol. 138, no. 1, pp. 12–20, Jan. 2012.

[20] M. J. Huber, "Traffic flow theory," *Transp. Res. Board Special Rep.*, vol. 62, nos. 1–2, pp. 5–7, 1976.

[21] B. Wang, W. Chen, B. Zhang, and Y. Zhao, "Regulation cooperative control for heterogeneous uncertain chaotic systems with time delay: A synchronization errors estimation framework," *Automatica*, vol. 108, Oct. 2019, Art. no. 108486.

[22] C. Deng, M. J. Er, G.-H. Yang, and N. Wang, ''Event-triggered consensus of linear multiagent systems with time-varying communication delays,'' *IEEE Trans. Cybern.*, vol. 50, no. 7, pp. 2916–2925, Jul. 2020.

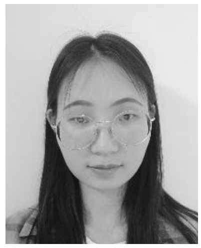
[23] L. Ma, X. Huo, X. Zhao, and G. Zong, ''Adaptive fuzzy tracking control for a class of uncertain switched nonlinear systems with multiple constraints: A small-gain approach,'' *Int. J. Fuzzy Syst.*, vol. 21, no. 8, pp. 2609–2624, Nov. 2019.

HU ZHANG received the Ph.D. degree from the College of Transportation, Jilin University. He is currently a Postdoctoral Researcher with the Business School, University of Shanghai for Science and Technology, Shanghai, China. His research interests include public transport management and intelligent transportation systems.

SHIDONG LIANG received the Ph.D. degree from the Transportation College, Jilin University, Changchun, China, in 2017. He is currently a Lecturer with the Business School, University of Shanghai for Science and Technology, Shanghai, China. His research interests include urban signal control and public transit management and control.

YIN HAN received the Ph.D. degree from the Transportation College, Jilin University, Changchun, China, in 2003. He is currently a Professor with the Business School, University of Shanghai for Science and Technology, Shanghai, China. His research interests include intelligent public transportation systems, traffic flow model, and modern transportation planning.

MINGHUI MA received the Ph.D. degree from the Transportation College, Jilin University, Changchun, China, in 2016. She is currently a Lecturer with the College of Automobile Engineering, Shanghai University of Engineering Science, Shanghai, China. Her research interests include traffic signal control and traffic data analysis.



RONGMENG LENG is currently pursuing the B.S. degree with the Business School, University of Shanghai for Science and Technology, Shanghai, China. Her work has focused on public transit management and intelligent transportation systems. After she graduates, she is interested in pursuing the M.S. or Ph.D. degree in the same field. She has completed the work of bus data handling with the Traffic Control Laboratory, Shanghai.