Hindawi
Journal of Advanced Transportation
Volume 2017, Article ID 7696094, 10 pages
https://doi.org/10.1155/2017/7696094

Research Article

Relieving the Impact of Transit Signal Priority on Passenger Cars through a Bilevel Model

Ding Wang, Wenxin Qiao, and Chunfu Shao

MOE Key Laboratory for Urban Transportation Complex Systems Theory and Technology, School of Traffic and Transportation, Beijing Jiaotong University, Beijing, China

Correspondence should be addressed to Wenxin Qiao; wxqiao@bjtu.edu.cn

Received 1 March 2017; Revised 8 May 2017; Accepted 18 May 2017; Published 11 July 2017

Academic Editor: Dongjoo Park

Transit signal priority (TSP) is an effective control strategy to improve transit operations on the urban network. However, the TSP may sacrifice the right-of-way of vehicles from side streets which have only few transit vehicles; therefore, how to minimize the negative impact of TSP strategy on the side streets is an important issue to be addressed. Concerning the typical mixed-traffic flow pattern and heavy transit volume in China, a bilevel model is proposed in this paper: the upper-level model focused on minimizing the vehicle delay in the nonpriority direction while ensuring acceptable delay variation in transit priority direction, and the lower-level model aimed at minimizing the average passenger delay in the entire intersection. The parameters which will affect the efficiency of the bilevel model have been analyzed based on a hypothetical intersection. Finally, a real-world intersection has been studied, and the average vehicle delay in the nonpriority direction decreased 11.28 s and 22.54 s (under different delay variation constraint) compared to the models that only minimize average passenger delay, while the vehicle delay in the priority direction increased only 1.37 s and 2.87 s; the results proved the practical applicability and efficiency of the proposed bilevel model.

1. Introduction

Traffic congestion in urban area significantly undermines the mobility of the city [1]. Encouraging public transportation system has gradually been recognized as an effective way to relieve traffic congestions. Transit vehicles cause less congestion and pollution due to their larger occupancy compared with passenger cars; however, transit riders often suffer longer travel time than passenger car riders considering connection and waiting time. Therefore, in order to attract transit ridership, it is necessary to grant priority to transit vehicles.

Transit priority system mainly includes "space priority" and "time priority." Space priority is mainly involved with authorizing dedicated lane for transit; however, it is not always feasible because of limited space. Time priority mainly refers to providing signal priority for transit to cross the signalized intersection. Adopting advanced traffic signal control techniques to optimize signal timing plan is an effective way to enhance the efficiency of transit vehicles. Transit signal priority (TSP) is a time priority technology which has been used extensively to improve the efficiency of transit vehicles by providing them priorities at signalized intersections.

While the implementation of TSP has potential benefits such as reducing overall transit delay as well as passenger delays at intersection, it also shows some negative effects on passenger cars and vehicles on the side streets which have only a small volume of transit vehicles. For example, extending green time in the priority direction will consequently shorten the green time of the nonpriority direction, which in turn may significantly increase the vehicle delay in the nonpriority direction. In order to solve this problem, this paper proposed a bilevel optimization model to minimize the impact of transit signal priority on side streets, while ensuring the efficiency of transit signal priority. The proposed bilevel model provides a new method to reach a state of equilibrium between priority and nonpriority directions at a signalized intersection. This method overcomes the challenge born by the conventional transit signal priority which cares little about the impact of TSP strategy on side streets.

2. Literature Review

2.1. TSP Control Method. Towards the efficiency of transit operation system, transportation scholars have rich research achievements on transit signal priority (TSP) strategies. Wilbur Smith cooperated with the Los Angeles Department of Transportation, and for the first time the transit signal priority experiment proved that transit signal priority can effectively shorten the travel time [2]. Thereafter, transit signal priority researches have been conducted globally. Based on different control methods, TSP can be mainly divided into two categories: active priority and passive priority.

Active priority relies on the vehicle detector, when the bus is detected by a special vehicle detector, and the signal corresponding to the control strategy will provide the designated signal control methods (red reduction, green extension, etc.). Furthermore, based on different computation process, active priority can be categorized into model-based priority and rule-based priority.

Rule-based priority determines whether to give bus priority based on a series of constraints, such as bus delay and bus presence. The rule-based priority includes conditional and unconditional strategies based on the presence or absence of specific priority rules and algorithms. Unconditional TSP was the first to be proposed; the theory is that, in any case, buses should be given priority when they are near the intersection. Many scholars have devoted great efforts into the unconditional transit signal priority studies [3–5]. Later, some scholars have proposed conditional TSP; the theory is that different priorities should be given to the bus under different traffic conditions based on actual bus performance [6]. Many transit priority strategies are rule-based priority; however, it is always not enrolled in the optimization process [7].

In contrast to rule-based priority, model-based priority has become more popular in TSP strategy recently. Model-based priority needs data from detectors, like traffic volume, bus location, and so forth [8, 9]. With the real-time data, the objective function can be an optimal performance of all vehicles in the intersection and receive corresponding signal timing plans with high efficiency.

Generally, active priority reacts with real-time traffic flow to ensure efficient performance; however, the success of active priority strategies depends on high accuracy motion detection and communication network, which is expensive and requires regular maintenance. Active priority strategies also affect the nonpriority vehicles and may cause the failure of signal coordination in the progression of vehicle platoons [10].

Compared with active priority, passive priority is easy and inexpensive to implement. Passive priority often assumes bus arrival at a given distribution and is also implemented in a given cycle length and phase sequence; the signal timing plans are irrelevant to actual vehicle presence, but based on obtained historical data. Therefore, the success of passive priority strategies depends on low fluctuation of the traffic volumes and deterministic dwell times of buses at stops [10]. It is proved that passive priority is effective when bus proportion is relatively large [11].

2.2. Signal Optimization Model with TSP. It is important to clarify the objective function when we build the signal optimization model. The literature provides several performance measures, like travel time, vehicle delay, passenger delay, and so forth [12]. These measures can be used as the objective function for signal optimization model and we can also use the combination of more than two performance measures as the multiple objective functions.

Traffic delay is a crucial index to evaluate the traffic efficiency at the intersection. According to HCM 2010, control delay is defined as "the increase in travel time due to traffic signal control" [13]. Many scholars have been devoted to minimize the vehicle delay in the signal control model; the objective function mainly includes minimizing the total bus delay [14, 15] and minimizing the total delay for bus and car passengers [10, 16, 17]. However, there are still some issues to be addressed in the future.

First, many researches have been constrained to grant priority in one direction without considering the transit vehicle in the nonpriority direction [7]; this will make the model simple, but the applicability of these models may be limited. Moreover, many existing systems do not consider the capacity difference between transit vehicle and passenger cars. Apparently the transit vehicle has much larger passenger occupancies than passenger car, especially in the peak hour when there is large commuter traffic. Existing researches have dealt with this problem by minimizing the total passenger delay in the entire intersection; this was a milestone in developing transit priority, but this method may not work in every condition. For example, when much higher volume of transit vehicle traveled in one direction and no transit vehicle in the other direction, congestion may increase significantly in the direction with no transit vehicle because the passenger delay model will allocate very limited green time in this direction.

Traffic system in China is quite complex because of the mixed-traffic flow; therefore, the capacity of transit vehicles needs to be carefully considered when designing signal timing plans [18]. Many scholars have aimed at minimizing average person delay of the entire intersection [19, 20]. However, minimizing passenger delay alone may cause negative effects on the nonpriority direction, especially when there are already heavy passenger vehicle volumes on the nonpriority direction.

To address these shortcomings of transit signal priority, this paper proposed a bilevel optimization model which aims at relieving the impact of transit signal priority on vehicles in the nonpriority direction. This bilevel model is different from previous TSP studies which care little about the vehicle delay in the nonpriority direction. This study proposed a new method to reach a status of equilibrium between priority and nonpriority directions to ease the traffic congestion.

3. Problem Description and Model Development

Passenger cars and transit vehicles are usually treated equally in the conventional signal control method and many

researches were aimed at minimizing the average vehicle delay regardless of the high occupancy of transit vehicles. With the growth of traffic demand, urban cities developed severe congestion problems. Meanwhile, the development of public transportation shows an opportunity to reduce passenger cars on the road by attracting people to travel by bus. Many cities in the world have applied TSP to improve the service level of transit vehicles, due to the awareness of different occupancy between transit vehicles and passenger cars; minimizing average passenger delay becomes an important index to evaluate the performance of signal timing plans. However, transit signal priority granting the right-of-way to vehicles in the priority direction may consequently sacrifice the vehicles in the nonpriority direction; hence, ignoring the impact of TSP on the nonpriority direction does not guarantee system optimal. As a result, how to reach a state of equilibrium between priority and nonpriority directions to ultimately ease traffic congestion is a problem to be addressed in this paper.

Concerning the mixed-traffic flow pattern in China, this paper proposed a bilevel optimization model for signal timing plan. The bilevel optimization model in this paper aims at minimizing the impact of transit signal priority on vehicles in the nonpriority direction through two subprograms.

The lower-level subprogram in this paper aims at minimizing the average passenger delay. Many scholars devote great efforts in this field; however, minimizing passenger delay alone may cause negative impacts on vehicles in the nonpriority direction, especially when there is already heavy volume in the nonpriority direction. The upper-level subprogram was designed to address this concern.

The upper-level subprogram aims at minimizing the average vehicle delay in the nonpriority direction subject to acceptable average vehicle delay variation in the priority direction, which will ensure transit priority and minimize the average vehicle delay in the nonpriority direction at the same time.

The general stochastic bilevel mathematical model can be formulated as follows, where UP presents the upper-level subprogram and LP is the lower-level subprogram:

(UP)

$$\begin{aligned} \min \quad & \sum_{j=1}^{N} \widehat{d_j} \\ \text{S.t:} \quad & \frac{\sum_{j=1}^{P} \widehat{d_j} - \sum_{j=1}^{P} d_j}{\sum_{j=1}^{P} d_j} \le \alpha, \\ & C_{\min} \le \widehat{C} \le C_{\max}, \\ & \widehat{G_i} \ge G_{i\min}, \\ & \sum_{i=1}^{m} \widehat{G_i} + L = \widehat{C}, \end{aligned} \tag{1}$$

where $\sum_{j=1}^{P} d_j$ is calculated by the following:

(LP)

$$\begin{aligned} \min \quad & \frac{\sum_{j=1}^{A} d_j \times \left(v_{j,\mathrm{BUS}}\gamma_{\mathrm{BUS}} + v_{j,\mathrm{CAR}}\gamma_{\mathrm{CAR}}\right)}{\sum_{j=1}^{A} \left(v_{j,\mathrm{BUS}}\gamma_{\mathrm{BUS}} + v_{j,\mathrm{CAR}}\gamma_{\mathrm{CAR}}\right)} \\ \text{S.t:} \quad & C_{\min} \le C \le C_{\max} \\ & G_i \ge G_{i\min} \\ & \sum_{i=1}^{m} G_i + L = C, \end{aligned} \tag{2}$$

where A is the number of approach lanes in the entire intersection; N denotes the number of approach lanes in the nonpriority direction; P is the number of approach lanes in the priority direction ($A = N + P$). d_j is the average vehicle delay of the jth approach; $\sum_{j=1}^{p} d_j$ is the acquired vehicle delay in the priority direction after minimizing the average passenger delay in the entire intersection (LP level model) and $\sum_{j=1}^{P} \widehat{d_j}$ is the acquired vehicle delay in the priority direction after the UP level model; $\sum_{j=1}^{N} \widehat{d_j}$ is the minimal vehicle delay in the nonpriority direction after the UP level model; α is the acceptable delay variation in the transit priority direction and $\alpha \in [0, 1]$; G_i and $\widehat{G_i}$ are green time allocated to phase i (s); $G_{i\min}$ is minimum green time allocated to phase i (s); C and $\widehat{C}$ are cycle length (s) and $C_{\min}$ means minimum cycle length(s) and $C_{\max}$ means maximum cycle length(s); m is the total number of phases; γ_{CAR} is the average passenger loads of Car; γ_{BUS} is the average passenger loads of bus; $v_{\mathrm{j,BUS}}$ is transit volume on approach j; $v_{\mathrm{j,CAR}}$ is the volume of passenger cars on approach j; L is the total lost time in one cycle length.

4. Methodology

4.1. Delay Function. There are several indexes to measure the performance efficiency of intersections, such as queue length, traffic delays, and stop times. The delay is caused by factors such as traffic interference, traffic management, and control facilities. Current signal control strategies always focus on minimizing average vehicle delay; however, the philosophy behind developing TSP control strategies is to minimize the average passenger delay by operating traffic signals, which is an evolutionary step from current signal control strategies.

HCM2010 defined the Average Control Delay experienced by all vehicles that arrive during the analysis period. The control delay for a given lane group is computed by using equation as follows.

$$d = d_{\mathrm{uniform}} + d_{\mathrm{incremental}} + d_{\mathrm{initial}}, \tag{3}$$

where d is the control delay (s/veh); d_{uniform} is the uniform delay (s/veh); $d_{\mathrm{incremental}}$ is the incremental delay (s/veh); d_{initial} is the initial delay (s/veh) and it is usually ignored when

there are no initial queue vehicles at the beginning of each cycle.

$$d_{\text{uniform}} = \frac{0.5C\left(1 - g/C\right)^2}{1 - \left[\min\left(1, X\right) g/C\right]},$$

$$d_{\text{incremental}} = 900T\left[\left(X-1\right) + \sqrt{\left(X-1\right)^2 + \frac{8KIX}{\text{CAP} \times T}}\right], \tag{4}$$

where X is the average volume to capacity ratio; g is the green time; C is the cycle length; I is upstream filtering adjustment factor; K is incremental delay factor. Because the value of $d_{\text{incremental}}$ is too small to affect the result in this paper, only d_{uniform} is computed in this paper.

Average passenger delay is estimated by using the intersection control delay, traffic volume and transit proportion in each approach, and average occupancy of passenger car and transit. The passenger delay of the jth approach is

$$d_{j,p} = d_j \times \left(v_{j,\text{BUS}}\gamma_{\text{BUS}} + v_{j,\text{CAR}}\gamma_{\text{CAR}}\right), \tag{5}$$

where $d_{j,p}$ is the passenger delay of the jth approach and then the average passenger delay in the LP mathematical objective function should be built as follows:

$$\frac{\sum_{j=1}^{A} d_j \times \left(v_{j,\text{BUS}}\gamma_{\text{BUS}} + v_{j,\text{CAR}}\gamma_{\text{CAR}}\right)}{\sum_{j=1}^{A}\left(v_{j,\text{BUS}}\gamma_{\text{BUS}} + v_{j,\text{CAR}}\gamma_{\text{CAR}}\right)}. \tag{6}$$

4.2. Constraints

4.2.1. Cycle Length. Increasing the length of cycle time can reduce the number of phase changes per unit time, reducing the signal loss per unit time and improving the traffic capacity of the intersection. However, if the cycle length is too long, traffic capacity will decrease and then the delay is growing fast. Minimum green time of each phase is the limiting factor to minimize cycle time:

$$C_{\min} = \sum_{i=1}^{n} g_{i\min} + L, \quad L = \sum_{i=1}^{n}\left(L_s + I_i - A\right),$$
$$\sum_i g_i + L = C, \tag{7}$$

where L is the total loss time of each cycle, I_i is the green interval, L_s is the start-up lost time, and A is amber time; then we can get the constraint of cycle length:

$$C_{\min} \le C \le C_{\max}. \tag{8}$$

4.2.2. Green Time. Minimum green time is necessary to ensure that the pedestrian can safely cross the intersection. In addition, every phase has a minimum green time, which means no phase will be skipped and provide a safety buffer between the phases. Thus, constraining the green times for each lane group makes it possible to reduce delays more than if minimum green times for the phases were fixed. Here we use the pedestrian green time as the minimum green time of each phase:

$$g_{i\min} = g_{\text{WALK}}^{\min}. \tag{9}$$

According to the equation of *HCM2010*,

$$g_{\text{WALK}}^{\min} = \text{WALk} + \text{FDW} \quad \text{FDW} = \frac{L_C}{\left(S_{15}^{p}\right)}, \tag{10}$$

where WALk = 7 s, FDW is pedestrian green flash interval (s), L_C is the crosswalk length (m), and S_{15}^{p} is the walking speed of 4 ft/s. Then we can get the range of green time of phase i:

$$g_i \ge g_{i\min}. \tag{11}$$

4.2.3. Delay Variation. This upper-level subprogram in the proposed model will benefit vehicles in the nonpriority direction; however, we need to make sure the average person delay change in the bus priority direction is acceptable. Therefore, constraining the delay changes in the priority direction should be a premise. There should be a delay restriction:

$$\frac{\sum_{j=1}^{P}\widehat{d_j} - \sum_{j=1}^{P} d_j}{\sum_{j=1}^{P} d_j} \le \alpha, \tag{12}$$

where $\sum_{j=1}^{P} d_j$ is the acquired vehicle delay in the nonpriority direction after minimizing the average passenger delay in the entire intersection (LP level model) and $\sum_{j=1}^{P}\widehat{d_j}$ is the acquired vehicle delay in the priority direction after the UP level model.

5. Solution Procedure

With new technologies like radar monitor and inductive loop detector, we can get real-time traffic data. In this paper, the proposed model can also be used in real-time traffic control with data including the volume of transit and passenger cars, turning ratios, occupancy of transit, and passenger cars.

When the inductive loop detector is put far from the intersection in the upstream, the total traffic volume of transit and passenger cars can be obtained. The volume of traffic is a significant factor in this paper, but it has a close relationship with the capacity of the road; thus in this paper, the saturation value will be used to reflect the traffic demand condition in the road. The inductive loop detector can also be put at the specific approach near the intersection; this can obtain the turning ratios of different approaches in the intersection. Occupancy of transit and passenger cars is hard to get with the real-time data, but future studies of connected vehicle can address this problem. In this paper, the real-time data cannot be obtained in Beijing; thus historic data from the field survey are used in the study. The difference of occupancy of transit between the peak period (commuter traffic in the morning and evening) and nonpeak period is significant; thus the efficiency of the model in TSP should be addressed under different occupancies.

Nonlinear optimization problems are always hard to solve. In order to avoid dealing with nonlinearities, the proposed model in this paper was solved through MATLAB using exhaustive search-based method through two nesting loops; one of the loops is used to get the optimal value for the lower-level subprogram, and another loops will get the global optimality for the proposed model by using the optimal value obtained in the first loop. Because of the low computational complexity, the computation time in this paper is less than 5 s, which is acceptable.

(1) Start the first loop (lower-level subprogram).

  (a) Set the objective function = minimum average passenger delay.
  (b) Search through all possible sets of variables (cycle length and green time of each phase).
  (c) For a selected variable set, compute the corresponding average passenger delay.
  (d) Select an optimal set with a minimum average passenger delay.
  (e) Compute the corresponding vehicle delay in the priority direction and nonpriority direction.

(2) Start the second loop (upper-level subprogram).

  (a) Set the objective function = minimum average vehicle delay in the nonpriority direction.
  (b) Search through all possible sets of variables (cycle length and green time of each phase).
  (c) For a selected variable set, compute the corresponding average vehicle delay.
    (1) If the vehicle delay in the priority direction dissatisfies the delay variation constraint (α), do nothing; this is not acceptable for transit signal priority.
    (2) If the vehicle delay in the nonpriority direction is not a minimum average vehicle delay, do nothing; this is not the optimal solution.
    (3) Else, store associated variables and delay result; this is the optimal solution.

(3) Output the optimal solution.

6. Numerical Experiments

Transit signal priority may cause significant negative impact on the nonpriority direction, especially when traffic volume is already very heavy. Therefore, if signal timing is determined only considering minimizing the passenger delay at an intersection would be unfair to the vehicle in nonpriority direction. The proposed bilevel model in this paper will reduce the vehicle delay in nonpriority direction compared to the models that only minimize the total personal delay, and this delay reduction serves as an index to evaluate the efficiency of the bilevel model in this paper.

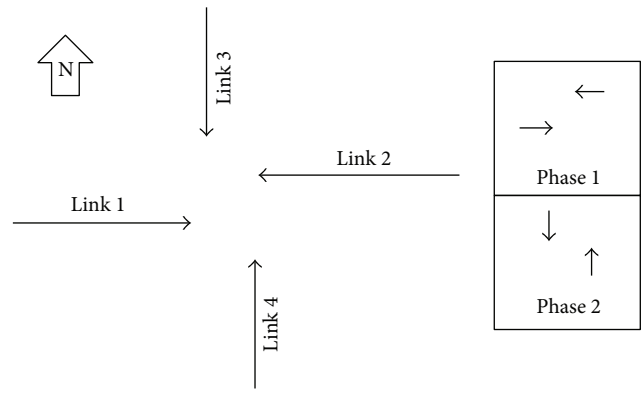


Figure 1: Test intersection.

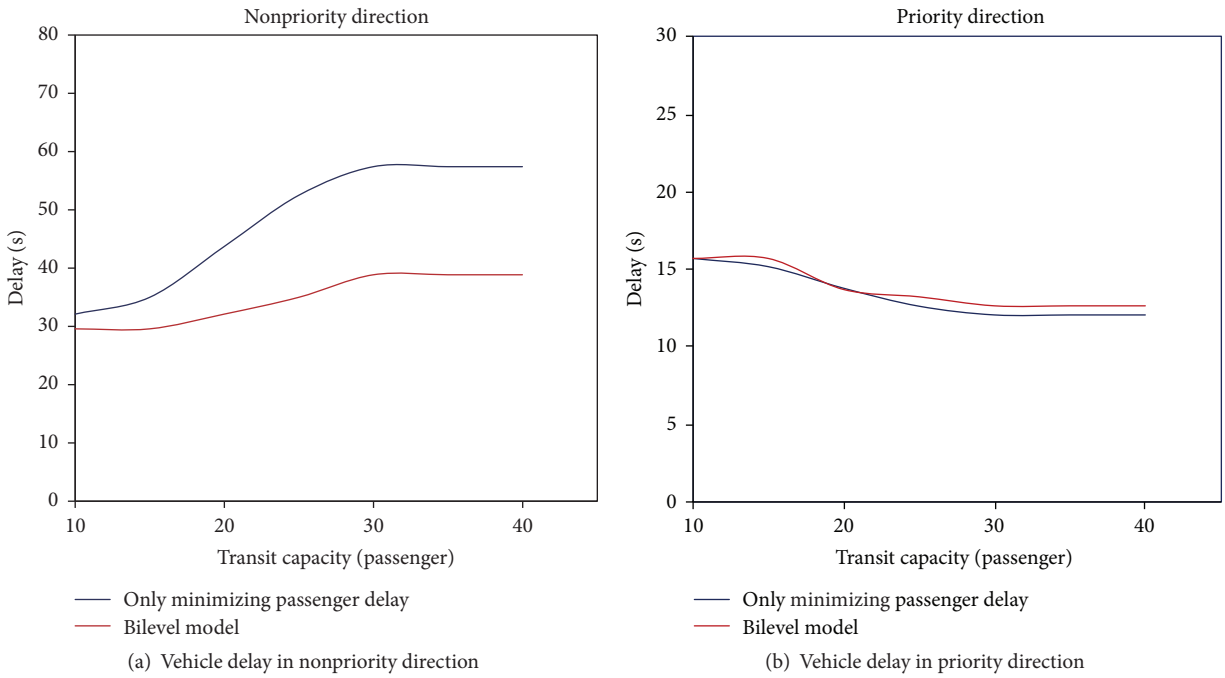
6.1. Hypothetic Intersection. This section analyzes the proposed transit signal optimizing model through a hypothetic intersection as an example. For illustration, consider a typical intersection with a two-phase signal control plan; Link 1 and Link 2 are in the transit priority direction and each has two lanes; Link 3 and Link 4 are in the nonpriority direction and each has only one lane (Figure 1). The road capacity is assumed to be the same as its saturation flow in this intersection and the loss time for each phase is 3 s.

There are several factors that will affect the efficiency of the proposed model in this paper, such as transit ratio, different saturations in two directions, passenger car, and transit loads. Saturation is an important index to reflect traffic condition, which is calculated as V/C, where V is the maximum traffic volume and C is the maximum capacity. Define $x1 \in (0, 1)$ as saturation index (V/C) in the transit priority direction and $x2 \in (0, 1)$ in the nonpriority direction. In order to reflect how each factor will affect the efficiency of the optimizing model, we assume some default values:

  (a) Minimum green time is 20 s for each phase.
  (b) Acceptable delay variation α in transit priority direction is 5%.
  (c) Passenger car occupancy is 1.5 (pax/veh) and transit occupancy is 30 (pax/veh).
  (d) Saturation $x1 = x2 = 0.5$.

6.1.1. The Occupancy of Passenger Car and Transit. Transit vehicles have a larger occupancy compared to passenger cars, regarding the total passenger delay in an entire intersection; the actual vehicle load is an important factor which will significantly affect the efficiency of the bilevel model in this paper. According to this consideration, we assume passenger car occupancy fixed at 1.5 (pax/veh) and let transit occupancy vary from 10 (pax/veh) to 40 (pax/veh); the result is shown in Figure 2.

Figure 2 shows a clear delay comparison between two directions. Figure 2(a) shows vehicle delay in the priority direction and Figure 2(b) shows vehicle delay in the nonpriority direction. The vertical distance between two lines is the delay reduction between minimizing passenger delay only



(a) Vehicle delay in nonpriority direction

(b) Vehicle delay in priority direction

Figure 2: Vehicle delays of two directions with different transit occupancy.

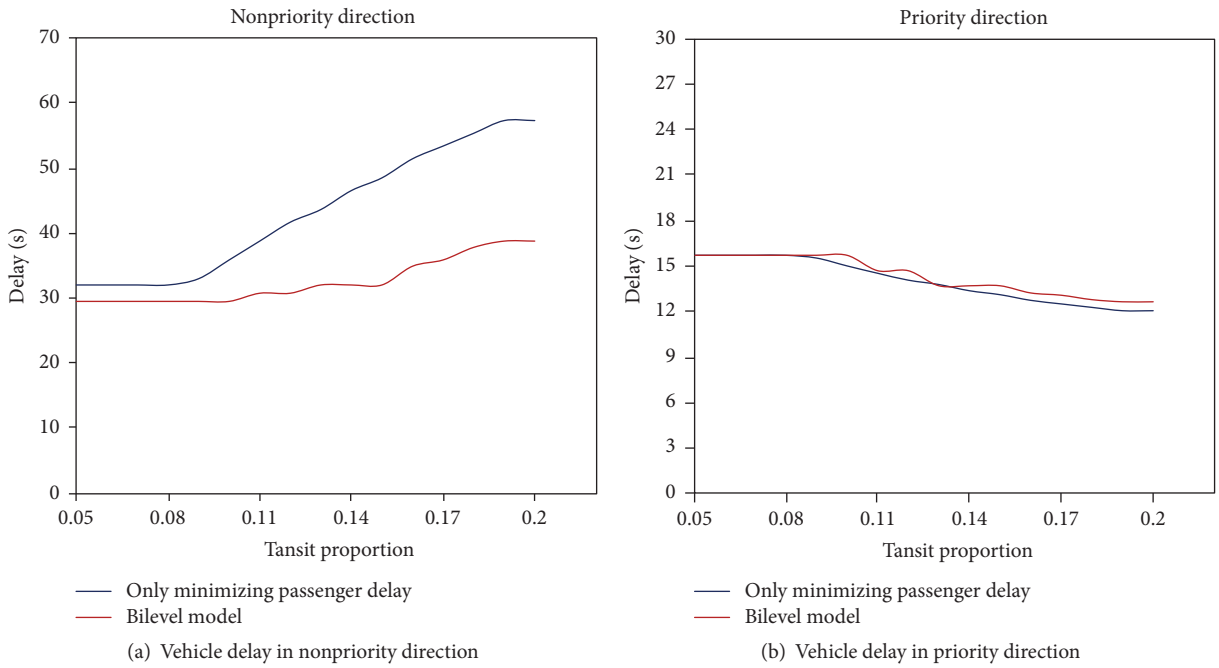
and using the bilevel model. These figures indicate that transit and passenger car loads have a noticeable influence on the efficiency of the bilevel model. When transit load is around 10 passengers, the delay reduction is only 7.91%; however, when transit occupancy is increased to 40 passengers, the delay reduction becomes 32.38%. Because we have restricted the delay variation in the priority direction and it can be easily seen from the right figure of Figure 2, these two figures have proved that the bilevel model in this paper could have a significant delay reduction in nonpriority direction and ensure the upper-level model has negligible impact on the transit priority direction.

6.1.2. Transit Proportion. Generally, transit vehicles have much larger occupancy than passenger cars; however, if there is only a small proportion of transit in the transit priority direction, the negative effect on the nonpriority direction will be too small to minimize. When the transit proportion increases, vehicles in the transit priority direction will be given more signal timing to minimize the total passenger delay, causing increased delay in the nonpriority direction. In order to analyze the variation of the delay reduction with the changes of transit proportion, Figure 3 shows different delay values between the two methods: minimizing passenger delay only and using the bilevel model proposed.

As the figures (Figure 3) shown above, the average vehicle delay reduction in the nonpriority direction is significant (Figure 3(a)). If we aim at minimizing passenger delay alone which will greatly increase the vehicle delay in the nonpriority direction, especially when transit proportion is very high, this negative effect will increase. In addition, we can easily draw a conclusion from the left figure: the higher transit proportion in the priority direction, the higher reduction of vehicle delay in the nonpriority direction, indicating the higher efficiency of the bilevel model proposed. At the same time, because of the restriction of vehicle delay variations in the priority direction, the bilevel model ensures the efficiency of TSP strategy, as shown in Figure 3(b); the average vehicle delay barely increases in the priority direction.

6.1.3. Saturation. The degree of saturation of intersection is a crucial factor to measure the level of service; obviously, different saturation rates will inevitably affect the efficiency of the bilevel model. In most situations, the transit signal priority strategies will be implemented on a main road and the saturation flow in the cross road is often in a lower degree; therefore, inappropriate implementation of TSP will easily cause congestion in the nonpriority direction. When taking saturation rates into consideration, we cannot ignore the relationship between different saturations in two directions. Figure 4 will show us the delay variation in the nonpriority direction with different saturation conditions.

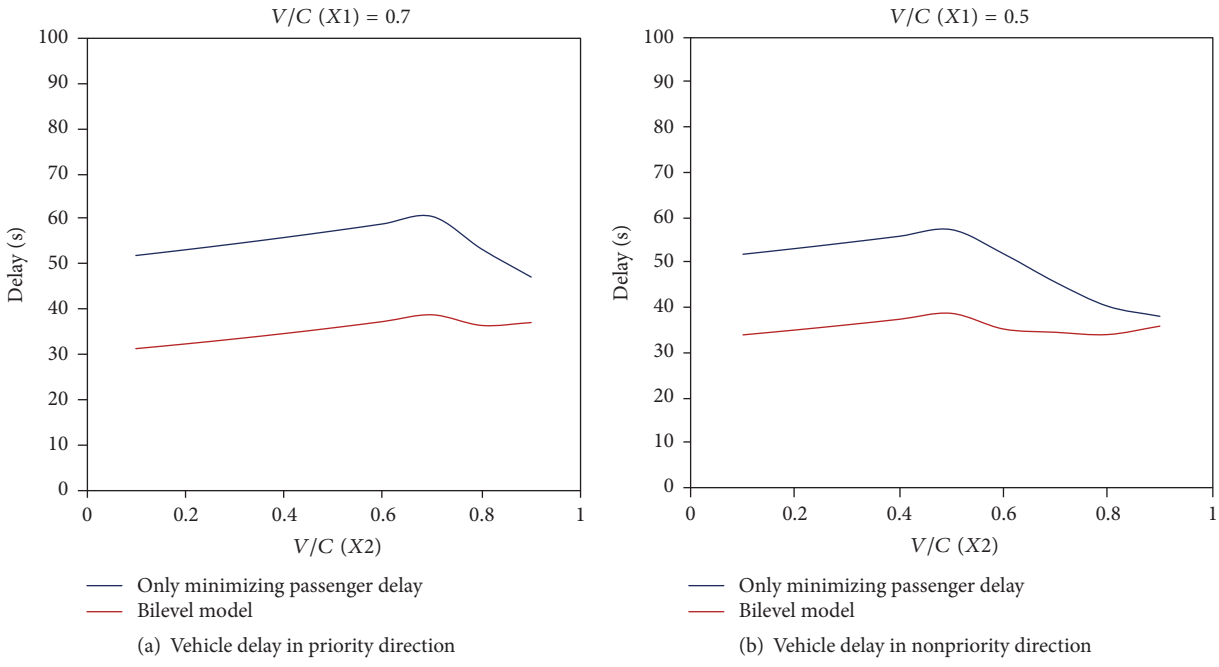
Based on the result in Figure 4, it is obvious that different saturations in priority direction (Figure 4(a)) and nonpriority direction (Figure 4(b)) will affect the efficiency of the bilevel model. It is worth mentioning that the vehicle delay in the nonpriority direction will firstly increase and then decrease; the change-point of $x2$ in Figure 4(a) is approximately 0.5 and in Figure 4(b) is approximately 0.7; however, this does not mean that the change-point will be the same as $x1$ but rather shows the change-point is closely related to $x1$. Before the change-point, delay reduction of the bilevel model is always significant; however, after the change-point, delay reduction becomes smaller.



(a) Vehicle delay in nonpriority direction

(b) Vehicle delay in priority direction

Figure 3: Vehicle delays of two directions with different transit occupancy.



(a) Vehicle delay in priority direction

(b) Vehicle delay in nonpriority direction

Figure 4: Vehicle delay (nonpriority direction) with different saturation of $X2$ and $X1$.

The actual road situation must be taken into consideration when deciding which method is more efficient to apply. In some situation, such as when the transit proportion is very low, the bilevel model proposed will not be efficient and its computation is more complex. However, when the transit proportion is very high, or transit occupancy is much larger than passenger cars, it is very beneficial to implement this bilevel model. Case 2 will introduce a practical example of an actual intersection.

6.2. Test Site. The optimizing model proposed in this paper is applied to a real-world intersection, which is located in downtown Beijing, where traffic volume is very large and dense. The intersection of Beilishi road and Fuchengmen

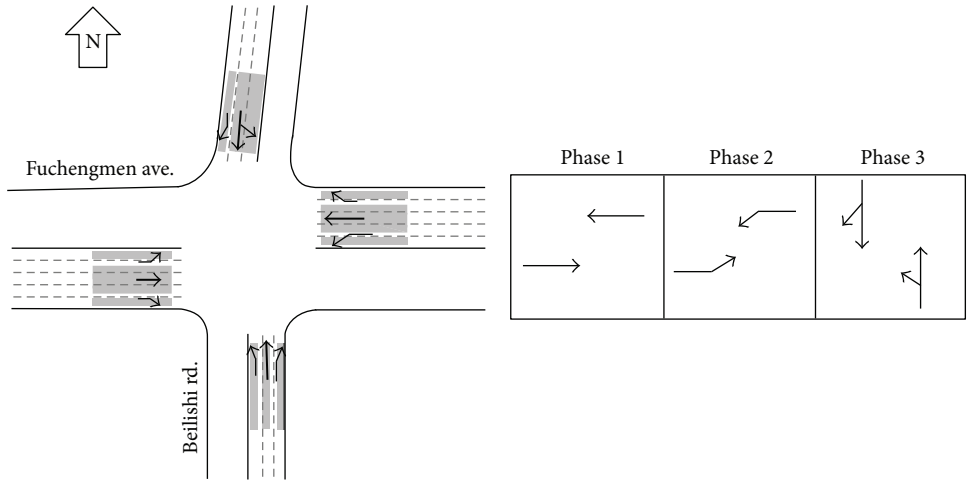


Figure 5: Study intersection.

Table 1: Parameters of the traffic volume.

Parameters	Value			
Direction	Volume	Turning ratio (left)	Transit proportion	
			Straight	Left
EB	2121 (veh/h)	0.05	0.1	0.04
WB	1142 (veh/h)	0.13	0.1	0.09
SB	266 (veh/h)	0.66	0.01	0.03
NB	460 (veh/h)	0.78	0.01	0.04
Transit occupancy	50 (pax/veh)			
Passenger car occupancy	1.8 (pax/veh)			
Cycle length	60 s~120 s			
Acceptable delay variation α	5% and 10%			

avenue has a lot of commuter traffic in the morning peak period in this area; Fuchengmen avenue has more than 20 bus lines; traffic congestion often occurs in this intersection. Figure 5 shows the lane configurations and the signal phase of the intersection. Fuchengmen avenue is a major road and Beilishi road is a minor road. The volumes have been transformed into equivalent unit traffic volumes. All traffic data in this study are collected from the morning peak period (7:00 a.m.~9:00 a.m.) on May 24, 2015 (Friday).

Commuting traffic is the major component of traffic flow in peak hour, often accompanied with congestion. The data on this actual intersection have been collected in the morning peak hour and the parameters of the traffic volume can be described as shown in Table 1.

As shown in Table 1, bus loads in the morning peak hour are more than 25 times the passenger cars; the transit proportion in the Fuchengmen avenue is about 10% of all vehicles and the transit proportion in the Beilishi road is only about 1%. Because the left turning vehicles and straight vehicles are in different phase and they have a different traffic composition, therefore, the restrictions of average vehicle delay in the priority direction were calculated separately. Average vehicle delay in the nonpriority direction refers to vehicles in all approaches in the nonpriority direction, and average person delay refers to the overall average person delay in the entire intersection. We use MATLAB to solve this bilevel model and the result is shown in Table 2.

Table 2 shows a comparison between two methods: minimizing passenger delay only and the bilevel model. When the acceptable delay variation α is 5%, the delay of straight vehicles in the priority direction has increased only 1.37 s or 4.41%; however, the average vehicle delay in the nonpriority direction has decreased 11.28 s or 8.82%. At this intersection, the delay of left-turn vehicles in the priority direction also decreases 5.91 s or 6.79% and the average person delay has little changes of 0.1 s or 0.52%. When the acceptable delay variation α is 10%, the benefits of the proposed bilevel model become more significant. For example, the delay of straight vehicles in the priority direction has increased only 2.87 s; however, the average vehicle delay in the nonpriority direction has decreased 22.54 s. The result in Table 2 has proved the practical applicability of the bilevel model.

7. Conclusion and Future Research

For conventional methods, traffic engineers are mostly focused on minimizing the vehicle delay only or minimizing the total passenger delay of the entire intersection. These

Table 2: The optimizing result compared with the conventional method.

Method and evaluation	Minimizing passenger delay	Bilevel model (5%)	Delay variation (s)	Bilevel model (10%)	Delay variation (s)
Cycle time C (s)	114	108		102	
Green g_1 (s)	60	54		48	
Green g_2 (s)	15	15		15	
Green g_3 (s)	30	30		30	
Average vehicle delay (s) (Straight-priority direction)	31.08	32.45	1.37 (4.41%)	33.95	2.87 (9.23%)
Average vehicle delay (s) (Left-turn-priority direction)	86.95	81.04	−5.91 (−6.79%)	75.15	−11.8 (−13.57%)
Average vehicle delay (s) (Nonpriority direction)	127.88	116.60	−11.28 (−8.82%)	105.34	22.54 (−17.63%)
Average person delay (s) (Entire intersection)	19.16	19.26	0.1 (0.52%)	19.42	0.26 (1.36%)

strategies can effectively reduce the delay of buses; however, they consequently sacrifice the right-of-way of vehicles from side streets. In this paper, a bilevel signal optimization model is developed to relieve the impact of transit signal priority on side streets. The proposed bilevel model provides a new method to reach a state of equilibrium between priority and nonpriority directions at a signalized intersection. This method overcomes the challenge born by the conventional TSP strategy which cares little about the impact of TSP strategy on side streets. To avoid dealing with nonlinearities, the proposed bilevel optimization model is solved by exhaustive search-based method through two nesting loops. Sensitivity analysis was conducted to test the efficiency of the bilevel model proposed, which provides an analytical method for traffic engineers to decide which signal timing program will be more efficient and practical based on the actual intersection. The results in the case study of an actual intersection were compared against the minimization of total person delay at the intersection, which has further proved the practical applicability of the bilevel model in this paper.

For future study, coordinated transit signal priority concerning vehicles from nonpriority direction will be studied. In this paper, single intersection was studied and multiple intersections with the coordinated timing plan will be an important research direction. In addition, with the new emerging connected vehicle technology, real-time traffic signal optimization strategy can be developed based on the proposed model. Finally, average passenger delay and vehicle delay have been adopted in this paper as the objective function, and future research can explore other important measurements, for example, emissions, travel time, and network reliability.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This paper is supported by "National Natural Science Foundation of China" (71621001) and "Fundamental Research Funds for the Central Universities" (2017JBM034).

References

[1] J. A. Lindley, "Urban freeway congestion: quantification of the problem and effectiveness of potential solutions.," *ITE Journal*, vol. 57, no. 1, pp. 27–32, 1987.

[2] W. Smith and Associates, *Study of Evolutionary Urban Transportation, Volumes I, II and III. U.S. Department of Housing and Urban Development*, Westinghouse Airbrake Co. and Institute of Public Administration, 1968.

[3] J. S. Ludwick Jr., "Simulation of an unconditional preemption bus priority system," *Transportation Research Record*, no. 536, pp. 1–10, 1975.

[4] F. Dion, H. Rakha, and Y. Zhang, "Integration of transit signal priority within adaptive traffic signal control systems," in *In 84th Annual Meeting of the Transportation*, Research Board, Washington, DC, USA, 2005.

[5] G. Zhou, "Design of transit signal priority at signalized intersections with queue jumper lanes," *Journal of Public Transportation*, vol. 12, no. 4, pp. 123–129, 2009.

[6] Z. Liu, W. Qiao, J. Liu, and D. Yang, "A rule-based signal compensation strategy for the non-bus-priority approaches at an intersection," in *Proceedings of the 15th COTA International Conference of Transportation Professionals: Efficient, Safe, and Green Multimodal Transportation, CICTP '15*, pp. 2007–2017, American Society of Civil Engineers, Beijing, China, July 2015.

[7] E. Christofa, I. Papamichail, and A. Skabardonis, "Person-based traffic responsive signal control optimization," *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 3, pp. 1278–1289, 2013.

[8] M. Strating, *Coordinated signal control for urban networks by using MFD Doctoral dissertation [M.S. thesis]*, Delft University of Technology, Delft, The Netherlands, 2010.

[9] W. Qiao, T. Zhu, X. F. Yang, and J. Liu, "Transit signal priority control algorithm with gaming theory: application in Beijing,"

in *Proceedings of the 95th Annual Meeting of the Transportation Research Board*, Washington, DC, USA, 2016.

[10] E. Christofa and A. Skabardonis, "Traffic signal optimization with application of transit signal priority to an isolated intersection," *Transportation Research Record*, no. 2259, pp. 192–201, 2011.

[11] N. Hounsell and J. Wu, "Public transport priority in real time traffic control systems," in *Asce Proceedings of the International Conference on Application of Advanced Technologies in Transportation Engineering*, vol. xviii, pp. 71–75, 1995.

[12] K. N. Balke and H. Charara, "Development of a traffic signal performance measurement system (TSPMS)," Texas Transportation Institute College Station Tx, vol. 42, no. 3, pp. 546–556, 2005.

[13] H. C. Manual, HCM2010, Transportation Research Board, National Research C, 2010.

[14] L. Head, D. Gettman, and W. Zhiping, "Decision model for priority control of traffic signals," *Transportation Research Record*, no. 1978, pp. 169–177, 2006.

[15] W. Ma, X. Yang, and Y. Liu, "Development and evaluation of a coordinated and conditional bus priority approach," *Transportation Research Record*, vol. 2145, no. 2145, pp. 49–58, 2010.

[16] G. L. Chang, M. Vasudevan, and C. C. Su, "Modeling and evaluation of adaptive bus-preemption control with and without AVL systems," in *Proceedings of the Vehicle Navigation and Information Systems Conference, '95*, pp. 305–316, 1995.

[17] M. Vasudevan, *Robust optimization model for bus priority under arterial progression [Ph.D. dissertation]*, 2005.

[18] E. Christofa, I. Papamichail, and A. Skabardonis, "Person-based traffic signal optimization for real-time applications," *In Transportation Research Board 91st Annual Meeting (No. 12-2228)*, January 2012.

[19] W. H. Zhang, L. U. Hua pu, Q. Shi, and Q. Liu, "Optimal signal-planning method of intersections based on bus priority," *Journal of Traffic and Transportation Engineering*, vol. 4, no. 3, pp. 49–53, 2004.

[20] Y. Feng, Y. L. Pei, and C. H. Cao, "Study of single bus priority signal planning under the mixed traffic flow," *Journal of Harbin Institute of Technology*, vol. 2, p. 031, 2007.