

Research Article

Queue Intensity Adaptive Signal Control for Isolated Intersection Based on Vehicle Trajectory Data

Juyuan Yin,¹ Peng Chen,² Keshuang Tang,³ and Jian Sun ¹

¹Department of Traffic Engineering, Key Laboratory of Road and Traffic Engineering, Ministry of Education, Tongji University, Shanghai 201804, China

²School of Transportation Science and Engineering, Beijing Key Laboratory for Cooperative Vehicle Infrastructure Systems and Safety Control, Beihang University, Beijing 100191, China

³Department of Transportation Information and Control Engineering, Key Laboratory of Road and Traffic Engineering, Ministry of Education, Tongji University, Shanghai 201804, China

Correspondence should be addressed to Jian Sun; sunjian@tongji.edu.cn

Received 16 July 2020; Revised 15 December 2020; Accepted 23 December 2020; Published 28 January 2021

Academic Editor: Shamsunnahar Yasmin

Copyright © 2021 Juyuan Yin 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.

With recent development of mobile Internet technology and connected vehicle technology, vehicle trajectory data are readily available and exhibit great potential to be used as an alternative data source for urban traffic signal control. In this study, a Queue Intensity Adaptive (QIA) algorithm is proposed, using vehicle trajectory data as the only input to perform adaptive signal control. First, a Kalman filter-based method is employed to estimate real-time queue state with vehicle trajectories. Then, based on queue intensity that quantifies queuing pressure, five control situations are defined, and different min-max optimization models are designed correspondingly. Last, a situation-aware signal control optimization procedure is developed to adapt intersection's queue intensity. QIA algorithm optimizes phase sequence and green time simultaneously. One case study was conducted at a field intersection in Shenzhen, China. The results show that provided with 7.4% penetrated vehicle trajectories, QIA algorithm effectively prevented queue spillback by constraining temporal percentage of queue spillback under 2.4%. The performance of QIA was also compared with the algorithm in Synchro and Max Pressure (MP) method. It was found that compared with Synchro, the extreme queue intensity, temporal percentage of queue spillback, delay, and stops were decreased by 54.7%, 97%, 22.3%, and 45.1%, respectively, and compared with MP the above four indices were decreased by 16%, 61.5%, -1.8%, and 49.4%, respectively.

1. Introduction

Traffic signal control has always been a hot topic due to its essential role in urban traffic operations. Adaptive signal control systems such as SCATS [1], OPAC [2], SCOOT [3], and RHODES [4] have been widely deployed, and they are mostly dependent on the data from fixed-location detectors such as inductive loop detectors and video cameras. However, considerable cost arising from installation and maintenance of detectors limits the wide-scale implementation of adaptive control systems, especially in developing countries [5]. Moreover, the erroneous data caused by frequently malfunctioned detectors in practice have significant impact on signal control performance [6].

With recent development of mobile Internet technology and connected vehicle (CV) technology, there is a growing trend for application of GPS devices in daily traveling such as vehicle-embedded GPS, cell phone GPS, and CV communication system. Vehicle trajectory data, sampled periodically from GPS devices on moving vehicles, contain a series of spatial-temporal points along the traveling route. Most of the studies associated with vehicle trajectory data have been focused on traffic state estimation, including traffic volume [7], density [8], travel time [9, 10], delay [11], emissions [12], and queue length [13–15]. It has been well recognized that compared with fixed detector data, vehicle trajectory data have advantages of being continuous, reliable, and low cost [16]. However, the major challenge for trajectory data to be

further utilized and explored comes from the low penetration rate of sampled vehicles currently and even in the near future, e.g., under 10% [17]. On the contrary, most of the vehicle trajectory-based traffic applications require relatively higher penetration rates, e.g., over 20% for delay estimation [11], lane-level speed estimation [17], and traffic signal control [18], and over 10% for arterial performance measurement [19]. This critical issue motivates researchers to develop methods for traffic state estimation under low-penetration conditions.

Nowadays, urban intersections are facing heavy traffic pressure and some typical adaptive control systems such as SCOOT and SCATS were reported to have deteriorated performance under congested conditions [20]. Despite previous studies which have developed signal control algorithms for undersaturated [21, 22] and oversaturated conditions [23, 24], respectively, the intersection with high traffic demands may experience a complex critical state where undersaturation and oversaturation alternate uncertainly. Even the intersection is not fully oversaturated for the whole period, since traffic can always fluctuate, any instantaneous sharp growth on queues may easily lead to queue spillback, especially when queue storage space is temporally insufficient. Therefore, real-time and accurate identification of traffic states is important, and moreover, timely applying appropriate signal control strategies for varying traffic situations is critical.

Because of desirable spatial-temporal characteristics, vehicle trajectory data can be used for providing an accurate and real-time estimation of traffic states, which will serve as a basis for developing new signal control algorithms and eventually help improve the operational performance of intersections. Moreover, using queue dynamics instead of traffic volumes as input to the signal control algorithm appears promising for intersections with high traffic demands and long queues. The reason is that, for queue spillback prevention, it will be more straightforward and effective to investigate queue dynamics directly with signal timings and estimated traffic states.

In this study, we propose Queue Intensity Adaptive (QIA) signal control algorithm using vehicle trajectory data as the only input to perform real-time estimation of queue state and adaptive signal control. The main contributions of this study are

- (1) A Kalman filter-based method is applied to estimate the queue state from vehicle trajectory data and provides real-time input to signal control optimization.
- (2) Queue intensity is presented to quantify the queuing pressure to be discharged within the limited space. It is modeled as a function of queue state and signal timings for distinguishing control situations as well as indicating signal control performance.
- (3) Min-max optimization models are designed for different control situations, and the worst-case queue intensity of the intersection is minimized.
- (4) A situation-aware signal control optimization procedure is developed to identify real-time control

situation and then apply appropriate models to optimize phase sequence and green time simultaneously, with the purpose of adapting the intersection queue intensity to the best control situation.

- (5) The proposed algorithm is evaluated based on a field intersection, and it is also compared with the algorithm embedded in Synchro and Max Pressure (MP) method.

The rest of the paper is organized as follows. Section 2 provides a brief review of related works on traffic signal control with vehicle trajectory data. Section 3 describes the overview of the QIA algorithm. Sections 4 and 5 explain the queue state estimation module and signal control optimization module, respectively. Section 6 provides a case study to validate the proposed QIA algorithm. Section 7 gives conclusions and future work for this study.

2. Literature Review

There have been several studies developing traffic adaptive signal control algorithms based on either Vehicle-to-Infrastructure (V2I) or CV communication data consisting of vehicle positions, speeds, and timestamps, which can be considered as a sort of vehicle trajectory data.

Feng et al. [6] proposed a real-time adaptive signal control algorithm utilizing vehicle location and speed data from CVs as the input to optimize phase sequence and duration by solving a two-level optimization problem. Two objective functions, minimizing total vehicle delay and minimizing queue length, were implemented and an algorithm called EVLS (Estimation of Location and Speed) was developed to estimate vehicle states of unequipped vehicles based on CV data.

Priemer and Friedrich [18] developed a decentralized adaptive traffic signal control algorithm with V2I communication data. The algorithm had a phase-based strategy and optimized phase sequence every 5 seconds in order to reduce the total queue length. The methods of dynamic programming and complete enumeration were used to solve the problem. It has been demonstrated that the algorithm shows better performance than TRANSYT-7F when the penetration rate was greater than 20%.

He et al. [25] presented a unified platoon-based formulation called PAMSCOD to optimize traffic signals for multiple travel modes at arterials and networks under V2I environment. A headway-based platoon recognition algorithm was developed to identify existing queues and significant platoons approaching each intersection. A mixed-integer linear program was formulated to solve the problem with flexible cycle lengths and offsets. The result showed that a 40% penetration rate was critical for effectively applying the algorithm.

Lee et al. [26] developed a cumulative travel-time responsive (CTR) real-time intersection control algorithm with CV data. The CTR algorithm adopted a Kalman filter to estimate cumulative travel time (CTT) with CV data. Based on real-time estimated CTT, the phase with the highest CTT was selected to be actuated for a fixed-time interval. The CTR

was proved to improve intersection mobility when the penetration rate of CV is greater than 30%.

Goodall et al. [27] developed a rolling horizon traffic signal control algorithm called predictive microscopic simulation algorithm. The algorithm used individual CV locations, headings, and speeds to predict an objective function such as delay, stops, and decelerations using microscopic simulation, and the phasing was chosen to optimize the objective function.

The above methods present innovative ideas of traffic adaptive signal control for the future environment with relatively high-penetration CV data. At the moment, how to effectively apply low-penetration vehicle trajectory data for traffic signal control remains the critical issue to be addressed. A representative study based on vehicle trajectory data from DiDi, which is a major ride-sharing company in China providing vehicle hailing service via smartphone applications, was made by Zheng et al. [16]. They proposed a feedback traffic signal control method for arterials using DiDi vehicle trajectory data as input. The method first estimated directional traffic volumes and evaluated arterial operational performance. Then, time of day schedule was determined by clustering approaches and signal parameters such as cycle length, green spilt, and offset which were optimized. The method was proved to be able to reduce arterial delay during the whole analysis period, whereas it may be difficult to accommodate real-time traffic state estimation and sensitive signal control under varying traffic conditions, considering that the trajectories were aggregated into relatively longer periods (e.g., 30 min).

3. Overview of the QIA Algorithm

QIA algorithm uses vehicle trajectory data consisting of instantaneous vehicle IDs, positions, and timestamps as the only input to perform real-time queue state estimation and adaptive signal control for the intersection. It mainly contains two modules: queue state estimation module (QSE) and signal control optimization module (SCO). QSE module provides queue states as necessary information for SCO module, which designs the optimization models and generates optimal signal timing plans.

The overview of QIA algorithm is shown in Figure 1. In the QSE module, a method based on the integration of Kalman filter (KF) and shockwaves is applied to estimate real-time queue state with vehicle trajectory data. In the SCO module, the signal timing plan specified by phase sequence and green time over a prospective fixed-length horizon is determined. Based on queue states estimated from the QSE module, queue intensity is presented to quantify the queuing pressure to be discharged within the limited space. It is modeled as a function of signal timings for distinguishing different control situations, such as short-term oversaturation, long-term oversaturation, and high queue intensity. Accordingly, particular optimization models are designed for different control situations in consideration of different expectations for signal control performance. A situation-aware signal control optimization procedure is developed to identify real-time control situation and determine the optimal timing

plan, aiming to adapt the queue intensity of the intersection to the best situation.

It is worth noting in this study that queue state instead of traffic volume serves as the basis for signal optimization. The reason is twofold. First, in low-penetration vehicle trajectory environment, the queue state will be relatively easier to estimate than other traffic states such as volume and delay. Second, for intersections confronting the challenge of high traffic demands and long queues, it is more straightforward and effective to relate queue length and queue forming speed directly with signal timings for the prevention of queue spillback.

4. QSE Module

In our previous study, we proposed a KF-based method using vehicle trajectory data to perform offline queue length estimation [28]. In the QSE module, the previous method is introduced and extended to meet the requirement of online operation by simplifying the allocation of joining points and leaving points, simplifying the estimation of queue discharging speed, and providing real-time queue states as input to signal control optimization. The main ideas of the QSE module are illustrated in Figure 2. The previous method [28] is only briefly described in this section for understanding the completeness of the whole work.

Vehicle trajectory data can be described as a series of spatial-temporal points (t_i^n, x_i^n) , where x_i^n is current position, t_i^n is timestamp, i is vehicle ID, and n is number of trajectory point. The sampling intervals are assumed to be under 3 seconds, which agree with the sampling frequency of DiDi vehicle trajectory data updated from cell phone GPS. The penetrated vehicle trajectories (see the blue lines in Figure 2) are used as the only input to QIA algorithm.

The basic ideas of the QSE module are that joining points and leaving points (see the red and green circles in Figure 2, respectively) are identified from vehicle trajectories in real time as the measurement input of KF; the queue state characterized by two variables, back of queue (see the magenta squares in Figure 2) and queue forming speed (see the magenta arrows in Figure 2), is estimated with KF update; the real-time queue states are produced for the SCO module.

4.1. State-Space Model for the Discrete-Time Queue Forming System. An appropriate state-space model, described by the state equation in equation (1) and the measurement equation in equation (2), is employed to characterize the queue forming dynamics and prepare for the application of KF:

$$\mathbf{X}_k = \Phi_{k|k-1} \mathbf{X}_{k-1} + \mathbf{w}_k, \quad (1)$$

$$\mathbf{Z}_k = \mathbf{H}_k \mathbf{X}_k + \mathbf{v}_k, \quad (2)$$

where \mathbf{X}_k is state vector, \mathbf{Z}_k is measurement vector, $\Phi_{k|k-1}$ is state transition matrix, \mathbf{H}_k is measurement matrix, \mathbf{w}_k is system noise vector, and \mathbf{v}_k is measurement noise vector.

Queue forming is considered as a discrete stochastic process, and queue state at a certain time step k is

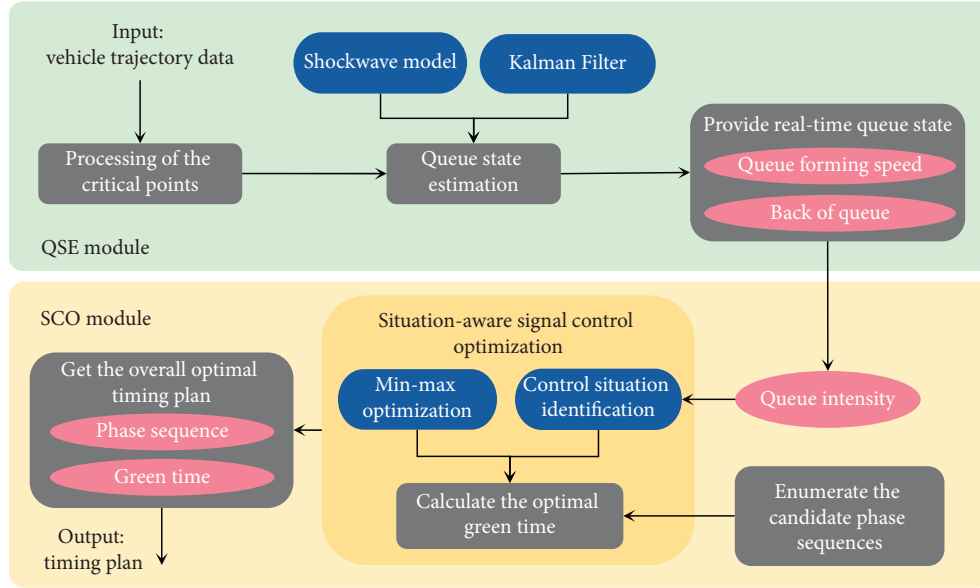


FIGURE 1: Overview of the QIA algorithm.

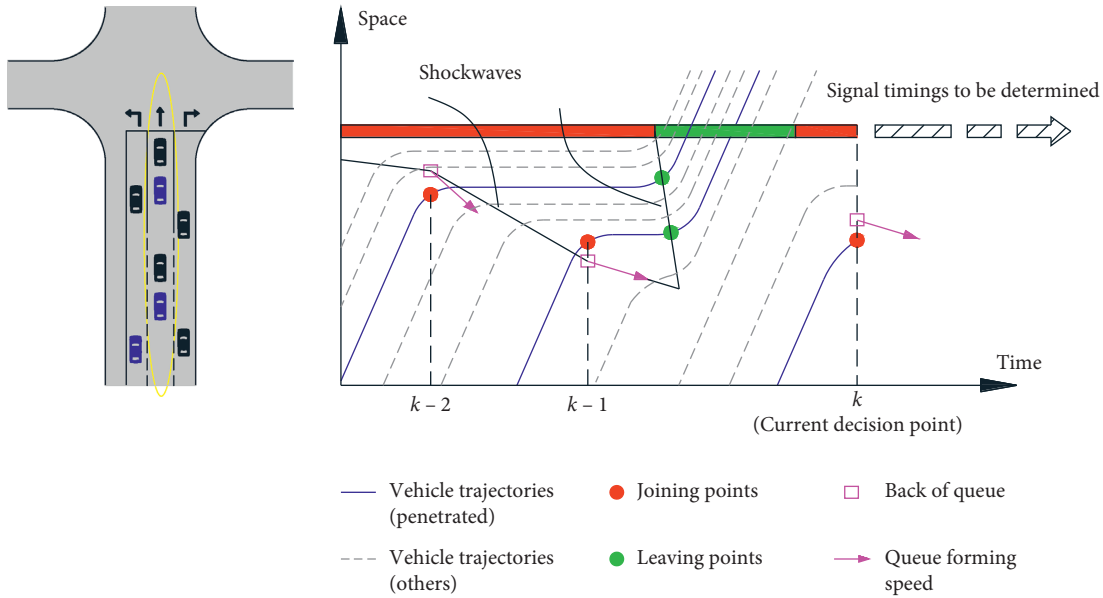


FIGURE 2: Illustrations for basic ideas of the QSE module.

characterized by two state variables: the back of queue QX_k and the queue forming speed VF_k . The acceleration of queue forming a_k is assumed to be a white Gaussian noise. Then, the state equation of the queue forming system for each movement can be represented by equation (3) with $\mathbf{X}_k = \begin{bmatrix} QX_k \\ VF_k \end{bmatrix}$, $\Phi_{k|k-1} = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix}$, $\mathbf{w}_k = \begin{bmatrix} (T^2/2) \\ T \end{bmatrix} a_k$, and T referring to the sampling interval. As QX_k is observable, \mathbf{H}_k from equation (2) can be given $[1 \ 0]$:

$$\begin{bmatrix} QX_k \\ VF_k \end{bmatrix} = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix} \begin{bmatrix} QX_{k-1} \\ VF_{k-1} \end{bmatrix} + \begin{bmatrix} (T^2/2) \\ T \end{bmatrix} a_k. \quad (3)$$

Once the characteristics of system noise and measurement noise are given, which are usually predetermined from field measurement, the state-space model for the discrete-time queue forming system is completed. The state-space models of different movements are independent.

4.2. Processing of the Critical Points

Step 1 (classification of vehicles' motion status): for each trajectory point (t_i^k, x_i^k) at current time step k , vehicle motion status M_i^k , which is defined to be either queuing ($M_i^k = 1$) or moving ($M_i^k = 0$), is simply classified by comparing the calculated point speed \bar{v}_i^k with a threshold \bar{v} :

$$M_i^k = \begin{cases} 1, & \text{if } v_i^k \leq \bar{v}, \\ 0, & \text{if } v_i^k > \bar{v}, \end{cases} \quad (4)$$

$$v_i^k = \frac{x_i^k - x_i^{k-1}}{t_i^k - t_i^{k-1}}.$$

Step 2 (identification of the critical points): the critical points refer to the joining points and the leaving points. The joining point is identified when vehicle motion status transfers from moving into queuing; similarly, the leaving point is identified when a transition of vehicle motion status from queuing to moving occurs. Use IP_i^k to indicate which class the point belongs to: NCP for noncritical point, JP for joining point, and LP for leaving point, and it can be determined by

$$IP_i^k = \begin{cases} \text{LP}, & \text{if } M_i^k = 0 \text{ and } M_i^{k-1} = 1 \text{ and } k \geq 2, \\ \text{JP}, & \text{if } M_i^k = 1 \text{ and } M_i^{k-1} = 0 \text{ and } k \geq 2, \\ \text{NCP}, & \text{otherwise.} \end{cases} \quad (5)$$

Step 3 (allocation for the critical points): for the critical point (τ^k, ξ^k) , if we project it to the stop bar at queue discharging speed VD , the crossover point written as $(\tau_{\text{projection}}^k, \xi_{\text{projection}}^k)$ can be calculated by equations (6) and (7). Then, (τ^k, ξ^k) belongs to the cycle that covers $\tau_{\text{projection}}^k$:

$$\xi_{\text{projection}}^k = x_{\text{stop}}, \quad (6)$$

$$\tau_{\text{projection}}^k = \tau^k + \frac{x_{\text{stop}} - \xi^k}{VD}. \quad (7)$$

4.3. Queue State Estimation with KF Update. Based on the processing of critical points, every time the joining point (τ^k, ξ^k) of a movement is newly identified, it serves as a new measurement input, $\mathbf{Z}_k = \xi^k$. Then, the KF performs state update and measurement update to estimate $\hat{\mathbf{X}}_k$ on that movement (see [28], for the details of KF update).

4.4. Real-Time Input to SCO Module. At the moment t_0 when the SCO module is triggered, $QX(t_0)$ and $VF(t_0)$ from all the movements are the inputs for the SCO module. They are calculated from $\hat{\mathbf{X}}_k = \begin{bmatrix} QX_k \\ VF_k \end{bmatrix}$, which is the queue state estimation coming from the latest KF update at τ^k . If t_0 and τ^k belong to the same cycle, $QX(t_0)$ and $VF(t_0)$ are determined by equations (8) and (9). If t_0 and τ^k belong to cycle n and m , respectively, $n > m$, $(t_{\text{MQ}}, x_{\text{MQ}})$, and $(t_{\text{RQ}}, x_{\text{RQ}})$ from cycle m to $n-1$ can be calculated by equations (11)–(17); then, $VF(t_0)$ and $QX(t_0)$ are calculated by equation (8) and (10):

$$VF(t_0) = VF_k, \quad (8)$$

$$QX(t_0) = QX_k + VF_k(t_0 - \tau^k), \quad (9)$$

$$QX(t_0) = x_{\text{RQ}}^{(n-1)} + VF_k(t_0 - t_{\text{RQ}}^{(n-1)}). \quad (10)$$

As most of the queued vehicles leave the queue at a saturation flow rate, queue discharging speed VD varies slightly over the cycles. Thus, VD is assumed to be a constant that can be predetermined by a linear regression method with historical vehicle trajectory data [28]. Besides, travel speed VT is also regarded as a predetermined constant.

5. SCO Module

In SCO module queue intensity, quantifies queuing pressure is presented as a function of signal timings, and it can also indicate signal control performance. Accordingly, five control situations are defined and different optimization models are designed correspondingly. The essential objective is to minimize queue intensities of all the movements over the decision horizon. Hence, it can be considered as a multiobjective optimization problem. The method of min-max optimization is applied to deal with worst-case scenarios, which means the movement with the largest queue intensity. A situation-aware signal control optimization procedure is developed to adapt intersection queue intensity via identifying real-time control situation and then applying appropriate models.

The SCO module is designed to be triggered when KF update is performed for any movement of the intersection, which means that a joining point is newly identified as mentioned in Section 4.3. The SCO module is also triggered if any signal interval starts. Once the SCO module is triggered, current timestamp is called decision point, and the timing plan specified by phase sequence and green time over a prospective fixed-length horizon, i.e., next two cycles from current decision point, is to be determined. The following key steps are related (1) to determine the set of all the candidate phase sequences, (2) to identify the control situation, calculate the optimal green time under each candidate phase sequence, and obtain the solution set after all the candidate phase sequences have been enumerated, and (3) to select the optimal phase sequence and corresponding green time from the solution set. The optimal timing plan will be executed until the next time SCO module is triggered.

5.1. Phase Sequence and Green Time. Signal timing plan can be described by phase sequence and green time, and they are considered as control variables in this optimization problem.

With ϕ_i denoting phase i , $\mathbf{S}^{(c)}$ ($n \times n$ 0-1 matrix) and $\mathbf{g}^{(c)}$ ($n \times 1$ vector, $\mathbf{g}^{(c)} = [g_1^{(c)}, g_2^{(c)}, \dots, g_n^{(c)}]^T$) are used to represent the variables of the phase sequence and the green time of all the n phases in cycle c ($c = 1, 2$), respectively. For the element $S_{i,j}$ of \mathbf{S} , $S_{i,j} = 1$ indicates that ϕ_j is actuated ahead of ϕ_i , otherwise, $S_{i,j} = 0$. The timestamp of green start $\mathbf{GS}^{(c)}$ ($\mathbf{GS}^{(c)} = [GS_1^{(c)}, GS_2^{(c)}, \dots, GS_n^{(c)}]^T$) is calculated as follows:

$$\mathbf{GS}^{(1)} = gS_0 + \mathbf{S}^{(1)}(\mathbf{g}^{(1)} + I), \quad (11)$$

$$\mathbf{GS}^{(2)} = gS_0 + \frac{1}{2} \sum_{i=1}^n (g_i^{(1)} + I) + \mathbf{S}^{(2)}(\mathbf{g}^{(2)} + I), \quad (12)$$

where I is interval length and gS_0 is timestamp of the first green start in cycle 1.

Note that \mathbf{g} is continuous (integer constraint is not considered), whereas \mathbf{S} is discrete. To simplify the problem, some proper rules are made so that \mathbf{S} can be determined only from a finite set of candidate phase sequences (Ω_S). The rules are specified as follows: first, each phase should be activated only once in each cycle of the decision horizon; second, phase sequence remains unchanged over the decision horizon, i.e., $\mathbf{S}^{(1)} = \mathbf{S}^{(2)}$. With such rules, all the candidate phase sequences can be easily enumerated. A typical four-approach intersection, for example, there are eight signalized movements: WBT, WBL, EBT, EBL, NBT, NBL, SBT, and SBL, and eight phases corresponded to each movement, denoted by ϕ_1 to ϕ_8 . With phase pair denoted by (ϕ_i, ϕ_j) , which is equal to (ϕ_j, ϕ_i) , there are totally eight feasible phase pairs: (ϕ_1, ϕ_2) , (ϕ_1, ϕ_3) , (ϕ_2, ϕ_4) , (ϕ_3, ϕ_4) , (ϕ_5, ϕ_6) , (ϕ_5, ϕ_7) , (ϕ_6, ϕ_8) , and (ϕ_7, ϕ_8) . Under the above two rules, a candidate phase sequence is simplified to consist of four phase pairs without any repeated phase, e.g., $\mathbf{S}: (\phi_1, \phi_2) \longrightarrow (\phi_3, \phi_4) \longrightarrow (\phi_5, \phi_6) \longrightarrow (\phi_7, \phi_8)$.

According to two different conditions of decision point, candidate phase sequences can be enumerated as follows.

Condition 1. The decision point is within the activation of (ϕ_p, ϕ_q) , which is the leading phase pair in the first cycle of the decision horizon, and only the following three phase pairs in this cycle are to be determined. If (ϕ_p, ϕ_q) refers to (ϕ_1, ϕ_2) , the following three can only be (ϕ_3, ϕ_4) , (ϕ_5, ϕ_6) , (ϕ_7, ϕ_8) or (ϕ_3, ϕ_4) , (ϕ_5, ϕ_7) , (ϕ_6, ϕ_8) , and each of them can form 3! permutations, i.e., totally 12 candidate phase sequences ($3! \times 2 = 12$).

Condition 2. The decision point is within the all-red interval following (ϕ_p, ϕ_q) . No phase has been activated, and all the four phase pairs in the first cycle are to be determined. If (ϕ_p, ϕ_q) refers to (ϕ_7, ϕ_8) , only the phase sequences with (ϕ_1, ϕ_2) , (ϕ_1, ϕ_3) , (ϕ_2, ϕ_4) , (ϕ_3, ϕ_4) , or (ϕ_5, ϕ_6) in the lead are considered as candidates, i.e., totally 60 candidate phase sequences ($5 \times 12 = 60$).

From the above, \mathbf{S} can be enumerated and the original problem is divided into independent subproblems. For each subproblem, one certain \mathbf{S}_k is selected from Ω_S as known, and only \mathbf{g} is to be determined, i.e., searching the optimal solution $\mathbf{g}^*(\mathbf{S}_k)$. \mathbf{g} should satisfy some primary constraints. First, green time should be higher than the minimum green G_{\min} and lower than the maximum green G_{\max} as shown in equation (13). Second, if the current decision point is in Condition 1, which means that (ϕ_p, ϕ_q) has been actuated

from $GS_i^{(1)}$, $i = p, q$, this part of green time $t_0 - GS_i^{(1)}$ is deterministic and must be satisfied as shown in equation (14). Third, the red time should not exceed the maximum red R_{\max} , as shown in equation (15), preventing an excessive red caused by the changeable phase sequence:

$$G_{\min} \leq g_i^{(c)} \leq G_{\max}, \quad i = 1, 2, \dots, n; c = 1, 2, \quad (13)$$

$$g_i^{(c)} \geq t_0 - GS_i^{(c)}, \quad i = 1, 2, \dots, n; c = 1, 2, \quad (14)$$

$$GS_i^{(c)} - RS_i^{(c)} \leq R_{\max}, \quad i = 1, 2, \dots, n; c = 1, 2, \quad (15)$$

where RS is the timestamp of red start and c denotes the cycle number.

After solving all the subproblems, the solution set $\{\mathbf{g}^*(\mathbf{S}_1), \mathbf{g}^*(\mathbf{S}_2), \dots, \mathbf{g}^*(\mathbf{S}_N)\}$ is obtained. The overall optimal \mathbf{S} and \mathbf{g} can be easily selected from the set, i.e., $\mathbf{g}^*(\mathbf{S}^*)$.

5.2. Queue Intensity. In this study, the signal control algorithm is essentially based on queue dynamics, and we present queue intensity $QI_i(t)$ to describe the relative size of queue for a certain movement i at t . $QI_i(t)$ is defined as the ratio of the queue length $QL_i(t)$ to the queue threshold QT_i as follows:

$$QI_i(t) = \frac{QL_i(t)}{QT_i}. \quad (16)$$

Queue intensity is presented because the absolute queue length may not reveal the actual queuing pressure within the limited space, e.g., for the same 200 meter queues of two movements but with different queue thresholds ($QL_1 = QL_2 = 200$ m, $QT_1 = 250$ m, and $QT_2 = 180$ m), the second one is obviously more urgent to be discharged because of queue spillback occurrence ($QI_1 < 1 < QI_2$). The queue threshold QT_i could be either a physical queue storage space (in length) or a subjective value for decision purposes.

The shockwave, which is usually used for depicting the process of queue formation and dissipation, is employed to relate the queue dynamics with signal control. At current moment t_0 , the parameters of each movement, back of queue QX , queue forming speed VF , queue discharging speed VD , and travel speed VT , are needed and assumed to be constant for the next several cycles. Once the signal timings are determined, the queue length from t_0 is a piecewise linear function of time, and the shockwaves can be built (see the blue lines in Figure 3). For each cycle, the end of maximum queue and residual queue (see (t_{MQ}, x_{MQ}) and (t_{RQ}, x_{RQ}) in Figure 3, respectively) can be derived according to equations (17)–(21). The maximum queue length (MQL) and the residual queue length (RQL) are the distances from stop bar x_{stop} to x_{MQ} and x_{RQ} , respectively. For simplicity, signal statuses include only green and red, and a signal cycle begins with a red start as shown in equation (21):

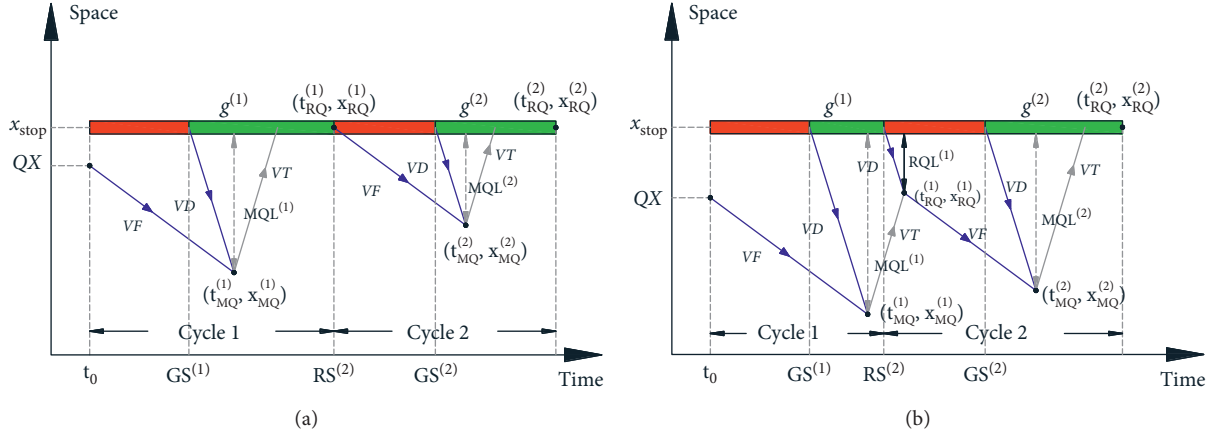


FIGURE 3: Illustrations for queue intensity. (a) Undersaturated condition. (b) Oversaturated condition.

$$t_{MQ}^{(c)} = \begin{cases} \frac{VF \cdot t_0 - VD \cdot GS^{(c)} + x_{stop} - QX}{VF - VD}, & c = 1, \\ \frac{VF \cdot t_{RQ}^{(c-1)} - VD \cdot GS^{(c)} + x_{stop} - x_{RQ}^{(c-1)}}{VF - VD}, & c \geq 2. \end{cases} \quad (17)$$

$$x_{MQ}^{(c)} = \begin{cases} VF \cdot (t_{MQ}^{(c)} - t_0) + QX, & c = 1, \\ VF \cdot (t_{MQ}^{(c)} - t_{RQ}^{(c-1)}) + x_{RQ}^{(c-1)}, & c \geq 2, \end{cases} \quad (18)$$

$$t_{RQ}^{(c)} = \max \left\{ RS^{(c+1)}, \frac{VT \cdot t_{MQ}^{(c)} - VD \cdot RS^{(c+1)} + x_{stop} - x_{MQ}^{(c)}}{VT - VD} \right\}, \quad (19)$$

$$x_{RQ}^{(c)} = x_{stop} + VD \cdot (t_{RQ}^{(c)} - RS^{(c+1)}), \quad (20)$$

$$RS^{(c)} = GS^{(c-1)} + g^{(c-1)}. \quad (21)$$

From the above definition of queue intensity, the maximum queue intensity (MQI) and the residual queue intensity (RQI), corresponding to MQL and RQL, respectively, can be easily calculated by equation (16). By combining equations (8) and (9) and equations (17)–(21), the queue intensities for all the movements of the intersection over the decision horizon are correlated to the control variables \mathbf{S} and \mathbf{g} .

Figure 3 illustrates the relationship between signal timings and queue intensity. The performance of signal control can be characterized in terms of MQI and RQI. For a single movement in a cycle, RQI = 0 indicates that queued vehicles are totally cleared at the end of the cycle (see Figure 3(a)), i.e., undersaturation; otherwise, RQI > 0 indicates that a portion of queued vehicles are not cleared in the cycle (see Figure 3(b)), i.e., oversaturation; MQI > 1 means that the maximum queue length through a cycle exceeds the queue threshold, and it may cause severe consequences such as queue spillback. It can also be observed that MQL and RQL are related to delay, e.g., the area enclosed by

shockwaves and stop bar exactly represents the total delay. However, delay is not directly optimized as the objective function in this study because it is difficult to estimate under oversaturation and simply minimizing delay may cause unacceptable long queues. Instead, we present queue intensity as the indicators for signal control performance. QX and VF, which are the key parameters for queue intensity, can be easily estimated by the proposed method in Section 4.

5.3. Situation-Aware Signal Control Optimization. MQI and RQI for a single movement in a cycle are expected to be small enough for stable queue and low delay. However, when considering the whole intersection over the decision horizon, minimizing MQI and RQI of all the movements across the cycles is a multiobjective optimization problem, of which the objectives are competing. For example, if one movement is assigned additional green time to clear the queue, i.e., reducing RQI, it will increase the red time of the following phases as well as their MQI. Minimizing the sum of multiple weighted objectives may be unreasonable because it is complex to assign individual weights of multiple objectives. Moreover, the movement with the maximum queue intensity of the intersection is critical, especially for intersections with a high risk of queue spillback. Thus, the worst-case queue intensity is always considered, i.e., $\max_{i \in \{1, 2, \dots, n\}} \{RQI_i\}$ and $\max_{i \in \{1, 2, \dots, n\}} \{MQI_i\}$, and the method of min-max optimization is applied.

In practice, signal control objectives are usually varying in different situations. For example, it is unnecessary to further reduce $\max_{i \in \{1, 2, \dots, n\}} \{MQI_i\}$ at the cost of higher $\max_{i \in \{1, 2, \dots, n\}} \{RQI_i\}$ when $\max_{i \in \{1, 2, \dots, n\}} \{MQI_i\}$ is far below a predefined safety margin γ ($\gamma < 1$). In such case, oversaturation will cause worse consequences, and minimizing $\max_{i \in \{1, 2, \dots, n\}} \{RQI_i\}$ with the constraints of $\max_{i \in \{1, 2, \dots, n\}} \{MQI_i\} \leq \gamma$ will be adequate in this case. Especially, for the intersection with highly fluctuated traffic, the optimization model is supposed to accommodate varying traffic conditions.

Therefore, the control situation is defined to characterize different traffic conditions for signal control optimization. A situation-aware signal control optimization procedure is

developed to determine the real-time control situation and apply appropriate optimization models for searching the optimal timing plan, aiming to adapt the queue intensity of the intersection to the best situation.

5.3.1. Control Situation. The control situation is described by a series of satisfied constraints. There are four constraints:

Constraint 0 (C0): primary constraints for \mathbf{g} , as shown in equations (13)–(15).

Constraint 1 (C1): all queued vehicles are cleared at the end of cycle 2, denoted by

$$\max_{i \in \{1, 2, \dots, n\}} \{RQI_i^{(2)}\} \leq 0. \quad (22)$$

Constraint 2 (C2): MQIs of all the movements over the decision horizon are below safety margin γ ($\gamma < 1$), denoted by

$$\begin{cases} \max_{i \in \{1, 2, \dots, n\}, i \neq p, q} \{MQI_i^{(1)}\} \leq \gamma, \\ \max_{i \in \{1, 2, \dots, n\}} \{MQI_i^{(2)}\} \leq \gamma. \end{cases} \quad (23)$$

Since $GS_p^{(1)}$ and $GS_q^{(1)}$ are known at any decision point ($GS_p^{(1)} = GS_q^{(1)} = gs_0$, (ϕ_p, ϕ_q) represents the leading phase pair), $MQI_p^{(1)}$ and $MQI_q^{(1)}$ are always independent of control variables according to equations (17) and (18), and they should be removed in this constraint as shown in equation (23).

Constraint 3 (C3): all queued vehicles within cycle 1 are cleared at the end of this cycle, denoted by

$$\max_{i \in \{1, 2, \dots, n\}} \{RQI_i^{(1)}\} \leq 0. \quad (24)$$

Then, the control situation is determined according to the satisfaction of different constraints, which are prioritized, and they represent different expectations for the signal control optimization. There are five control situations denoted by CS taking values of 0, 1, 2, 3, and 4.

CS = 0 (invalid situation): C0 has the highest priority, for it represents some basic requirements that must be satisfied for a valid timing plan. If C0 cannot be satisfied, this situation is invalid.

CS = 1 (long-term oversaturation): with C0 satisfied, C1 has a higher priority than C2 and C3, i.e., to clear all queues at the end of the next two cycles. If C1 cannot be satisfied, this situation is defined as long-term oversaturation, which means that vehicles experience too much delay, and growing queue intensities may soon lead to queue spillback.

CS = 2 (high queue intensity): with C0 and C1 satisfied, C2 has a higher priority than C3, i.e., to limit

queue intensities of the whole intersection under the safety margin over the next two cycles, ensuring that long-term oversaturation is prevented. If C2 cannot be satisfied, this situation is defined as high queue intensity, which means that long-term oversaturation can be prevented, though the intersection faces the risk of queue spillback.

CS = 3 (short-term oversaturation): with C0, C1, and C2 satisfied, C3 is to be satisfied, i.e., to clear all queues at the end of the next cycle, ensuring that long-term oversaturation is prevented, and queue intensities are limited to the safety margin. If C3 cannot be satisfied, this situation is defined as short-term oversaturation, which means that oversaturation will not last more than one cycle, and queue intensities are below safety margin.

CS = 4 (undersaturation and low queue intensity): if all the constraints can be satisfied, the intersection keeps undersaturated over the next two cycles, and queue intensities are below safety margin.

5.3.2. Optimization Procedure. Based on the above-defined control situations, the situation-aware signal control optimization procedure consists of a series of optimization problems with constraints from loose to tight, aiming to search the optimal timing plan with the best control situation, i.e., the highest CS. The whole optimization procedure can be depicted in Figure 4 and is specified as follows:

Step 1: start the procedure with the set of candidate phase sequences Ω_S prepared.

Step 2: select one of the undetermined phase sequences \mathbf{S}_k from Ω_S .

Step 3: solve the optimization problem 1 (OP 1). Minimizing $f_1 = \max_{i \in \{1, 2, \dots, n\}} \{RQI_i^{(2)}\}$ subject to C0, as shown in equation (25), then check if the feasible solution exists, if yes, go to step 4; if not, $CS(\mathbf{S}_k) = 0$, no solution for \mathbf{S}_k ; thus, \mathbf{S}_k is not feasible; then, go to step 8:

$$\begin{aligned} \min \quad & f_1 = \max_{i \in \{1, 2, \dots, n\}} \{RQI_i^{(2)}\} \\ \text{s.t.} \quad & \text{C0: } \begin{cases} G_{\min} \leq g_i^{(c)} \leq G_{\max}, \\ g_i^{(c)} \geq t_0 - GS_i^{(c)}, \\ GS_i^{(c)} - RS_i^{(c)} \leq R_{\max}, \quad i = 1, 2, \dots, n; c = 1, 2. \end{cases} \end{aligned} \quad (25)$$

Step 4: with f_1^* and \mathbf{g}_1^* denoting the optimal objective value and solution of OP 1, respectively, check if $f_1^* \leq 0$; if yes, go to step 5; if not, $CS(\mathbf{S}_k) = 1$, and the optimal objective value and solution for \mathbf{S}_k are $f^*(\mathbf{S}_k) = f_1^*$ and $\mathbf{g}^*(\mathbf{S}_k) = \mathbf{g}_1^*$, respectively; then, go to step 8.

Step 5: solve OP 2. Minimizing $f_2 = \max_{i, j \in \{1, 2, \dots, n\}, i \neq p, q} \{MQI_i^{(1)}, MQI_j^{(2)}\}$ subject to C0 and C1, as shown in equation (26), check if $f_2^* \leq \gamma$; if

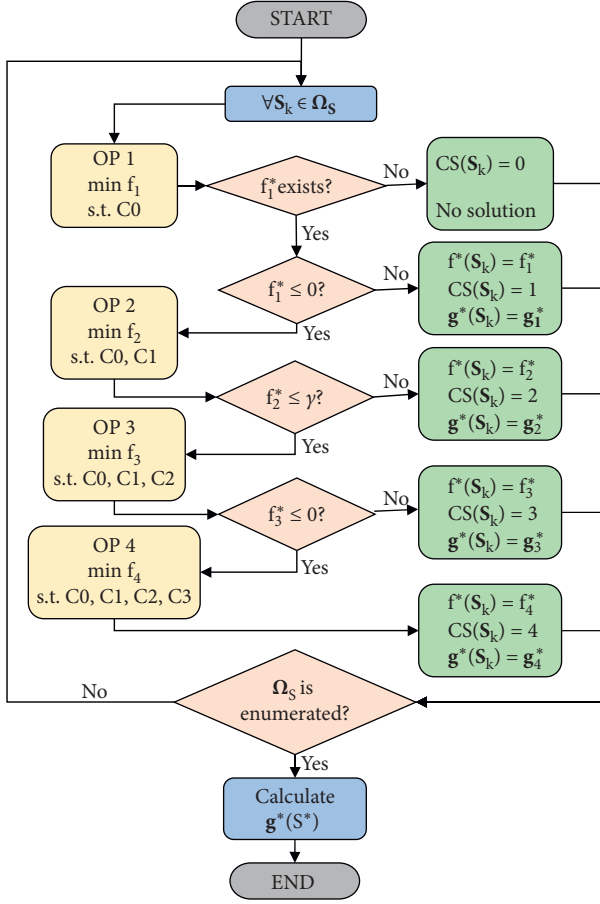


FIGURE 4: Flowchart of signal control optimization procedure of the SCO module.

yes, go to step 6; if not, $CS(S_k) = 2$, and the optimal objective value and solution for S_k are $f^*(S_k) = f_2^*$ and $g^*(S_k) = g_2^*$, respectively; then, go to step 8:

$$\min f_2 = \max_{i,j \in \{1,2,\dots,n\}, i \neq p,q} \{MQI_i^{(1)}, MQI_j^{(2)}\}$$

$$C0: \begin{cases} G_{\min} \leq g_i^{(c)} \leq G_{\max}, \\ g_i^{(c)} \geq t_0 - GS_i^{(c)}, \\ GS_i^{(c)} - RS_i^{(c)} \leq R_{\max}, \end{cases} \quad (28)$$

$$C1: RQI_i^{(2)} \leq 0,$$

$$i, j = 1, 2, \dots, n; c = 1, 2.$$

Step 6: solve OP 3. Minimizing $f_3 = \max_{i \in \{1,2,\dots,n\}} \{RQI_i^{(1)}\}$ subject to C0, C1, and C2, as shown in equation (27), check if $f_3^* \leq 0$; if yes, go to step 7; if not, $CS(S_k) = 3$, and the optimal objective value and solution for S_k are $f^*(S_k) = f_3^*$ and $g^*(S_k) = g_3^*$, respectively; then, go to step 8:

$$\min f_3 = \max_{i \in \{1,2,\dots,n\}} \{RQI_i^{(1)}\}$$

$$C0: \begin{cases} G_{\min} \leq g_i^{(c)} \leq G_{\max}, \\ g_i^{(c)} \geq t_0 - GS_i^{(c)}, \\ GS_i^{(c)} - RS_i^{(c)} \leq R_{\max}, \end{cases} \quad (28)$$

$$C1: RQI_i^{(2)} \leq 0,$$

$$C2: \begin{cases} MQI_j^{(1)} \leq \gamma, \\ MQI_i^{(2)} \leq \gamma, \end{cases}$$

$$C3: RQI_i^{(1)} \leq 0,$$

$$i, j = 1, 2, \dots, n; j \neq p, q; c = 1, 2.$$

where p and q are indices of (ϕ_p, ϕ_q) , which is the leading phase pair of the first cycle.

Step 7: solve OP 4. Minimizing $f_4 = \max_{i,j \in \{1,2,\dots,n\}, i \neq p,q} \{MQI_i^{(1)}, MQI_j^{(2)}\}$ subject to C0, C1, C2, and C3, as shown in equation (28). $CS(S_k) = 4$, and the optimal value and solution for S_k are $f^*(S_k) = f_4^*$ and $g^*(S_k) = g_4^*$, respectively; then, go to step 8:

$$\min f_4 = \max_{i,j \in \{1,2,\dots,n\}, i \neq p,q} \{MQI_i^{(1)}, MQI_j^{(2)}\}$$

$$C0: \begin{cases} G_{\min} \leq g_i^{(c)} \leq G_{\max}, \\ g_i^{(c)} \geq t_0 - GS_i^{(c)}, \\ GS_i^{(c)} - RS_i^{(c)} \leq R_{\max}, \end{cases} \quad (28)$$

$$C1: RQI_i^{(2)} \leq 0,$$

$$C2: \begin{cases} MQI_j^{(1)} \leq \gamma, \\ MQI_i^{(2)} \leq \gamma, \end{cases}$$

$$C3: RQI_i^{(1)} \leq 0,$$

$$i, j = 1, 2, \dots, n; j \neq p, q; c = 1, 2.$$

Step 8: check if all the phase sequences in Ω_S has been enumerated; if yes, the optimal objective values, solutions, and CS for each phase sequence in Ω_S have been determined, i.e., $f^*(S_k)$, $g^*(S_k)$, and $CS(S_k)$, $\forall S_k \in \Omega_S$, and then go to step 9; if not, go to step 2 and start a new iteration.

Step 9: find the set Λ_S of phase sequences with the highest CS, as shown in equation (29); then, the optimal phase sequence $S^* \in \Lambda_S$ has the minimum objective value, as shown in equation (30), and the corresponding optimal solution is $g^*(S^*)$:

$$\Lambda_S = \arg \max_{S_k \in \Omega_S} \{CS(S_k)\}, \quad (29)$$

$$S^* = \arg \min_{S_k \in \Lambda_S} \{f^*(S_k)\}. \quad (30)$$

Note that $f_4 = f_2 = \max_{i,j \in \{1,2,\dots,n\}, i \neq p,q} \{MQI_i^{(1)}, MQI_j^{(2)}\}$ is because MQI can be further reduced when $\max_{i \in \{1,2,\dots,n\}, c=1,2} \{RQI_i^{(c)}\}$ has been constrained to be 0.

The procedure shows that searching the optimal green time under each phase sequence is a process of improving the control situation, by which situation-aware signal control optimization is characterized. If the control situation cannot be further improved, the optimal timing plan is also determined. Generally, a timing plan is optimized to satisfy as many constraints as possible and achieve lower objective values so that $\mathbf{g}^*(\mathbf{S}^*)$ with the highest CS and the lowest objective value is overall optimal.

6. Case Study and Result Analysis

A field intersection, Huanggang-Fuzhong intersection in Shenzhen, China, is selected to validate the QIA algorithm. Figure 5 shows the layout of the investigated intersection. It has four approaches and eight signalized movements (including through and left-turn traffic flow). The hourly and 15 min volumes from 8:00 to 9:00 are surveyed and shown in Table 1. The Peak Hour Factor (PHF), which indicates the short-term traffic fluctuation, is also computed and presented. The PHFs of all the movements range from 0.66 to 0.83, indicating varying traffic conditions. Worse yet, the intersection faces a high risk of queue spillback due to high traffic demand with significant fluctuation and limited queue storage space, e.g., only 260 m on the northbound approach (SBT and SBL).

Our previous study showed that, at this intersection, there were about 7.4% DiDi vehicles, which uploaded their trajectory data including instantaneous positions and timestamps with a resolution of three to six seconds [28]. To explore the potential of QIA algorithm to be implemented in current data environment, 7.4% penetrated vehicles are simulated to upload their locations with 10-meter errors every three seconds according to the field situation, and these trajectories are used as the only input for QIA algorithm in the case study.

6.1. Experiment Design. To evaluate the QIA algorithm under the traffic conditions of high demands and short-term fluctuations, scenarios with different PHFs, 0.65, 0.7, 0.75, 0.8, and 0.85, are simulated. In each scenario with $PHF = \rho$, V_{hour} denotes hourly volumes in Table 1, and peak 15 min volume and nonpeak 15-min volume are set $(V_{hour}/4\rho)$ and $(V_{hour} - V_{hour}/4\rho)/3$, respectively.

Note that, since we aim to evaluate any possible queue length in the case study, the queue storage space of each movement in simulation is set much longer than the physical truth. We still use 260 m as the queue threshold ($QT = 260$ m) for each movement, which means that QT only serves for control and evaluation purposes. Even if queue is longer than QT, it literally means queue spillback and will not impact the traffic like gridlock in simulation.

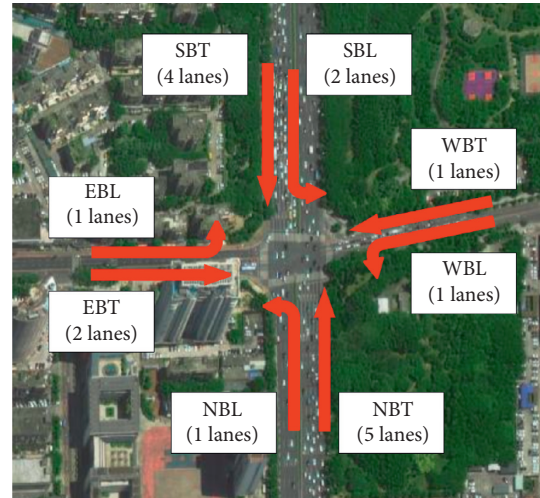


FIGURE 5: Layout of Huanggang-Fuzhong intersection.

6.2. Comparison of Different Methods. The proposed QIA algorithm is compared with the other three methods: QIA-Lite, fixed-time optimization by Synchro, and MP control [29].

QIA-Lite refers to the version of QIA without phase sequence optimization, which means that the phase sequence is fixed and only green time are optimized in QIA-Lite, i.e., steps 1, 2, 8, and 9 in the QIA optimization procedure are skipped. Synchro is a widely used signal timing tool, using hourly volume and PHF as input to generate a fixed timing plan that minimizes average delay and stops. MP is a distributed adaptive control policy which has attracted attention in recent years for notable performance on stabilizing network queues. It is also capable of isolated intersection, at which the pressure of a movement refers to the queue length multiplied by the corresponding saturation rate. Based on real-time measurement of queue lengths, at each decision point, MP selects the phase pair with the maximum pressure to be activated for a specific time interval, e.g., 5 s, and then MP is called again. “Algorithm 3: MP with minimum green time” in [30] is implemented as MP in this case, and MP is called every 5 s which is also suggested in [30]. The queue length input to MP is provided by the QSE module, which means that the data source and the data processing of MP, QIA-Lite, and QIA are identical. The related parameters in this case study are listed in Table 2.

6.3. Results and Discussion. QIA, QIA-Lite, and MP were coded in MATLAB, and the fixed-time method was achieved by Synchro. The studied four methods were simulated with several scenarios in VISSIM. Each scenario was simulated for 2 hours. The computational efficiency of QIA was able to meet the requirements of actual speed simulation in a PC with 8-GB RAM and an Intel i7-7700 CPU.

Figure 6 illustrates the details of QIA optimization at a certain decision point during the simulation of $PHF = 0.65$. At $t = 569$ s, the activated phase pair (ϕ_5, ϕ_7) terminated and

TABLE 1: Traffic volumes at Huanggang-Fuzhong intersection.

	WBT	WBL	EBT	EBL	NBT	NBL	SBT	SBL	
15-min volume (pcu)	1st	47	58	63	107	431	25	797	67
	2nd	50	96	75	82	618	14	610	94
	3rd	62	76	124	69	500	18	614	145
	4th	43	111	84	128	499	17	583	78
Hourly volume (pcu/h)	202	341	346	386	2048	74	2604	384	
PHF	0.81	0.77	0.7	0.75	0.83	0.74	0.82	0.66	

TABLE 2: Parameters for the case study.

Movement	q	R	\mathbf{P}_1	VF_1 (m/s)
WBT	0.002	100	$\begin{bmatrix} 62.5 & 0 \\ 0 & 0.1 \end{bmatrix}$	-0.37
WBL	0.003	100	$\begin{bmatrix} 93.75 & 0 \\ 0 & 0.15 \end{bmatrix}$	-0.63
EBT	0.003	111	$\begin{bmatrix} 93.75 & 0 \\ 0 & 0.15 \end{bmatrix}$	-0.37
EBL	0.004	100	$\begin{bmatrix} 125 & 0 \\ 0 & 0.2 \end{bmatrix}$	-0.64
NBT	0.01	135	$\begin{bmatrix} 312.5 & 0 \\ 0 & 0.5 \end{bmatrix}$	-1.06
NBL	0.002	100	$\begin{bmatrix} 62.5 & 0 \\ 0 & 0.1 \end{bmatrix}$	-0.16
SBT	0.01	142	$\begin{bmatrix} 312.5 & 0 \\ 0 & 0.5 \end{bmatrix}$	-1.42
SBL	0.003	111	$\begin{bmatrix} 93.75 & 0 \\ 0 & 0.15 \end{bmatrix}$	-0.4

Notes: the following parameters have the same values for all the movements: $VD = -6$ m/s, $VT = 10$ m/s, $\bar{v} = 1$ m/s, $G_{\min} = 10$ s, $G_{\max} = 80$ s, $R_{\max} = 260$ s, $I = 5$ s, $QT = 260$ m, and $\gamma = 0.7$; q , \mathbf{R} , and \mathbf{P}_1 are parameters for KF, see [28], for details; q , \mathbf{R} , \mathbf{P}_1 , VF_1 , and VD are measured from a presimulation based on volumes in Table 1 and original signal control.

triggered the optimization procedure of QIA; thus, current decision point was $t = 569$ s (see the dark dashed lines). Current queue intensity and its growth speed (VF/QT) of each movement (see the blue arrows) were estimated from the QSE module. Since the decision point is in Condition 2 as defined in Section 5.1, there were 60 candidate phase sequences to be enumerated, and the optimal green time of all the phase sequences were calculated. The results of two candidate phase sequences are analyzed: one is the current phase sequence $S_0: (\phi_6, \phi_8) \rightarrow (\phi_2, \phi_4) \rightarrow (\phi_1, \phi_3) \rightarrow (\phi_5, \phi_7)$ and the other is $S_k: (\phi_6, \phi_8) \rightarrow (\phi_5, \phi_7) \rightarrow (\phi_2, \phi_4) \rightarrow (\phi_1, \phi_3)$.

For S_0 (LHS in Figure 6), $f_1 = \max_{i \in \{1, 2, \dots, 8\}} \{RQI_i^{(2)}\}$ subject to C0 was first minimized in OP 1 by equation (22). Since the optimized RQIs (see the dark circles) of several movements at the end of cycle 2 were still greater than zero, i.e., C1 could not be satisfied, the optimization procedure under S_0 ended with $\mathbf{g}^*(S_0) = \mathbf{g}_1^*(S_0)$ and $CS(S_0) = 1$. For S_k (RHS in Figure 6), RQIs of all the movements were minimized to zero, i.e., C1 could be satisfied, then $f_2 = \max_{i, j \in \{1, 2, \dots, 8\}, i \neq j} \{MQI_i^{(1)}, MQI_j^{(2)}\}$ subject to C0, and C1 was minimized in OP 2 by equation (23). However, optimized MQIs (see the magenta squares) of WBL, EBL, and SBT within cycle 2 were higher than the safety margin γ ($\gamma = 0.7$ in this case), i.e., C2 could not be satisfied. Thus, the

optimization procedure under S_k ended with $\mathbf{g}^*(S_k) = \mathbf{g}_2^*(S_k)$ and $CS(S_k) = 2$.

For both phase sequences, QIA provided their optimal green time to adapt the queue intensities, whereas $\mathbf{g}^*(S_k)$ was considered better than $\mathbf{g}^*(S_0)$ for $CS(S_k) > CS(S_0)$. Compared with $\mathbf{g}^*(S_0)$, $\mathbf{g}^*(S_k)$ sets (ϕ_5, ϕ_7) ahead of (ϕ_2, ϕ_4) and (ϕ_1, ϕ_3) , and it shortened the red time of (ϕ_5, ϕ_7) so that the queue intensity of SBT which had the highest queue forming speed was further reduced. As a result, despite of increased residual queues on some movements in cycle 1, all queues were eventually cleared at the end of cycle 2, i.e., $\mathbf{g}^*(S_k)$ prevented long-term oversaturation at the cost of worsening short-term oversaturation, which revealed the tradeoff logic of QIA. Actually, S_k was found to be optimal for this decision after all the candidate phase sequences were enumerated. Therefore, $\mathbf{g}^*(S_k)$ would be executed in QIA, whereas $\mathbf{g}^*(S_0)$ was the only choice for QIA-Lite as the phase sequence was fixed.

During the simulation, the queue intensity of each movement per second was recorded. We took the maximum value over all the movements $\max_{i \in \{1, 2, \dots, 8\}} \{QI_i(t)\}$ as the queue intensity of the whole intersection. The variation of the queue intensity of the whole intersection over the study period (2 hours) is shown in Figure 7.

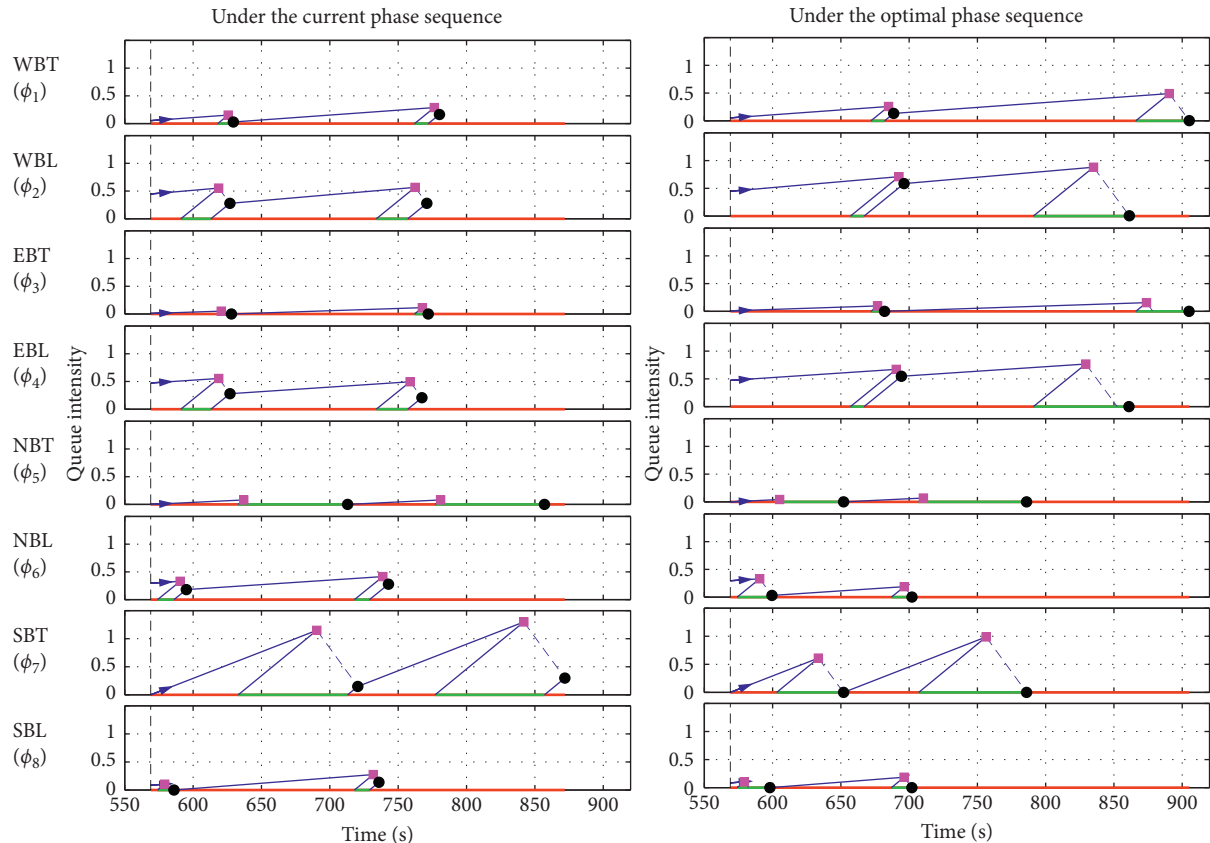


FIGURE 6: Illustrations of QIA optimization under two typical phase sequences at decision point $t = 569$ s.

From Figure 7, under the short-term traffic fluctuation, queue intensity varied significantly with time. The highest queue intensity of fixed time and QIA-Lite always occurred at the first 15 min in each hour, and queue spillback happened when $PHF < 0.85$. Because SBT, the movement with the highest hourly demand of the intersection (651 pcu/h/lane and 4 lanes), reached its peak volume and the intersection was probably oversaturated in the first 15 min, it was hard for the two methods to limit the growing queues.

MP kept the queue intensity under a relatively low level in the first 30 min for each hour. Since the pressure is proportional to the saturation rate, SBT and NBT, and the two main movements of the intersection got more priorities according to MP policy. However, the queue intensity of MP reached high, and queue spillback was occasional in the last 30 min in each hour. That is because movements with few lanes could not get enough services when their volumes reached the peak.

Compared with the other three methods, QIA basically kept the queue intensity lower and more robust during the whole period except for slight queue spillback at the extremely fluctuated condition ($PHF = 0.65$).

As short-term fluctuation was enhanced with decreasing PHF, the queue intensity of fixed time grew much more significantly than the other three methods. The queue intensity of QIA had the least growth, which means that QIA stabilized the queue intensity and limited the extreme value in the highly fluctuated situation.

In order to quantitatively evaluate their performances, the comprehensive results of the case study are shown in Figure 8. In Figure 8(a), we use the extreme queue intensity (EQI, the 95th percentile of queue intensities over the study periods) to represent the worst situation over the traffic fluctuated period. In Figure 8(b), to further analyze the severity of queue spillback, we present the temporal percentage of queue spillback (TPQS) referring to the percentage of the sum of t with $\max_{i \in \{1, 2, \dots, 8\}} \{QI_i(t)\} \geq 1$ over the study period. Delay and stops are also presented in Figures 8(c) and 8(d), respectively. Some detailed results are listed in Table 3.

From the results, the four indices of fixed time grew sharply as fluctuation increased. Especially, at $PHF = 0.65$, EQI exceeded 2 and TPQS was beyond 30%. Average EQI, TPQS, delay, and stops over all scenarios were 1.5, 16.9%, 59.7 s, and 2.26, respectively, representing serious queue spillback and poor performance. The other three methods achieved much better performances; even they used only 7.4% penetrated vehicle trajectories as data input, and the fixed-time optimization of Synchro used accurate hourly demand and PHF value. However, as fluctuations were close to zero, vehicle trajectories at such a low penetration were hard to ensure absolute superiorities of adaptive methods because of estimation errors of the queue state.

The results show that QIA achieved the best performance among the four methods. The average EQI, TPQS, delay, and stops of QIA were 0.68, 0.5%, 46.4 s, and 1.24, respectively.

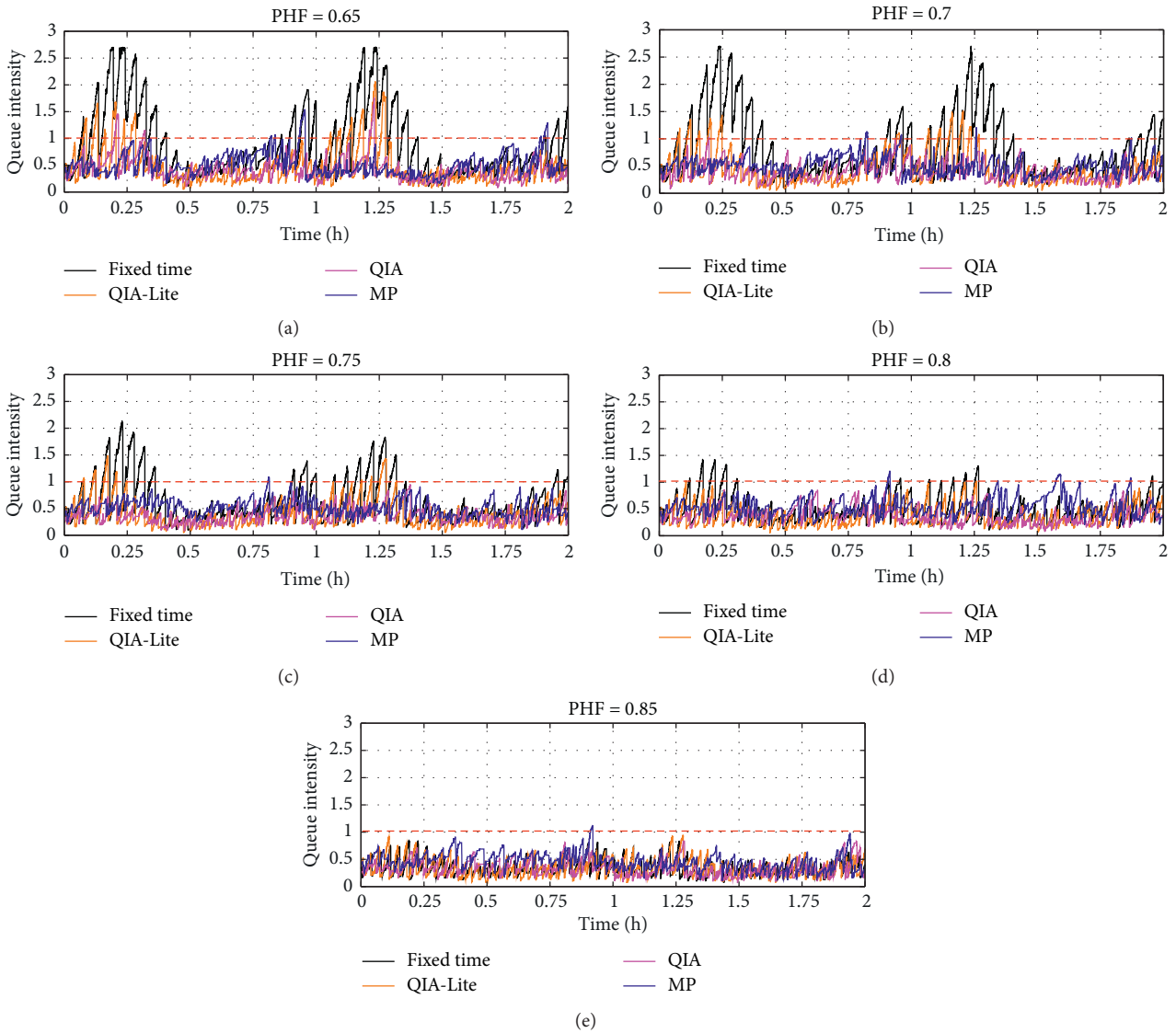


FIGURE 7: Queue intensity variation over the study period.

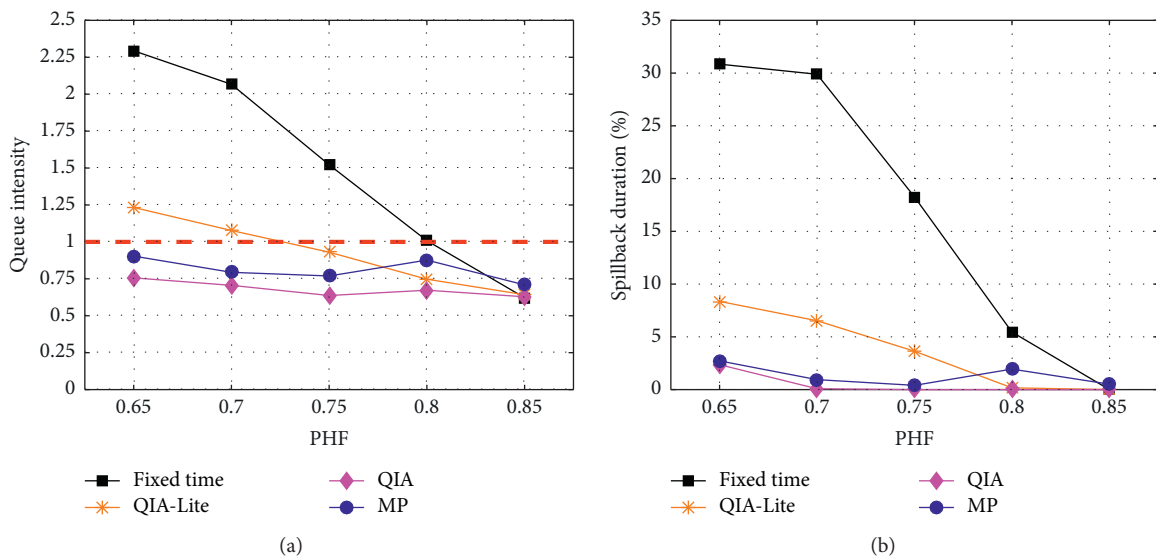


FIGURE 8: Continued.

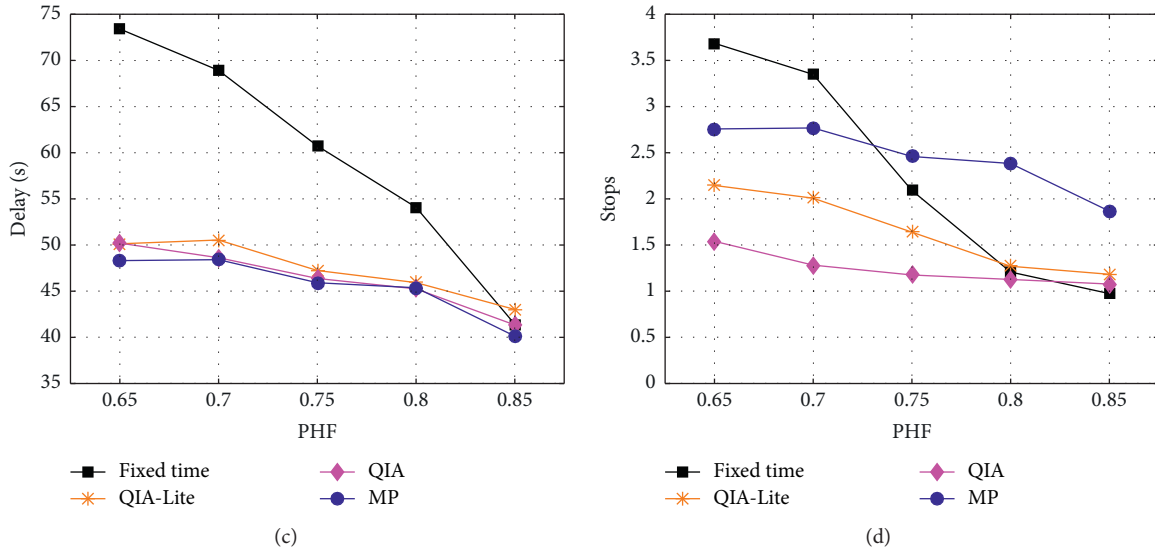


FIGURE 8: The comprehensive results of four methods over study period. (a) Extreme queue intensity (EQI). (b) Temporal percentage of queue spillback (TPQS). (c) Delay per vehicle. (d) Stops per vehicle.

TABLE 3: The comprehensive results of four methods.

Methods	EQI		TPQS		Delay (s)		Stops	
	Avg	Max	Avg	Max	Avg	Max	Avg	Max
Fixed time	1.50	2.29	16.9%	30.9%	59.7	73.4	2.26	3.69
QIA-Lite	0.93	1.23	3.7%	8.4%	47.4	50.1	1.65	2.15
QIA	0.68	0.76	0.5%	2.4%	46.4	50.2	1.24	1.54
MP	0.81	0.90	1.3%	2.7%	45.6	48.3	2.45	2.76

Even at the extremely fluctuated condition (PHF = 0.65), TPQS was only 2.4%, meaning that QIA almost prevented queue spillback. Except for delay, the other three indices of QIA were the best among the four methods. Actually, there were quite small differences on delay of QIA-Lite, MP, and QIA.

It should be noted that, in this case, only one isolated intersection was built, and queue spillback would not really happen in simulation as mentioned before, so the real impacts of queue spillback such as gridlock were not considered. As a result, the delay shown in Figure 8(c) and Table 3 may be underestimated against reality. Therefore, it can be inferred that if the impacts of queue spillback were considered, QIA would achieve quite better performance in delay against the other three methods because queue spillback would rarely happen.

The results of MP prove its effective performance at isolated intersection. According to the policy of MP, frequently switching the activated phases according to the feedback of varying pressures benefits MP in the lowest delay. The comparison of MP and QIA shows that the proposed QIA algorithm is also meaningful. QIA optimizes green time and phase sequence based on reasonable models and control situations, and it adapts the queue intensities of the whole intersection without neglecting those movements with few lanes. As a result, drops of 49%, 16% and 61.5% on stops, EQI and TPQS, respectively, and only 0.8 s growth on delay were made by QIA against MP.

The comparison of QIA, QIA-Lite, and Synchro shows that both the optimization of green time and the optimization of phase sequence are significant. QIA-Lite may also be meaningful, considering that optimizing only green time is more acceptable and achievable for practice than adjusting phase sequence frequently.

Generally, over all the high-demand scenarios of different fluctuations, QIA effectively prevented queue spillback by constraining TPQS under 2.4%, and it achieved much better performance than Synchro fixed-time optimization with drops of 54.7%, 97%, 22.3%, and 45.1% on EQI, TPQS, delay, and stops, respectively. Compared with MP the above four indices were decreased by 16%, 61.5%, -1.8%, and 49.4%, respectively.

It is noticeable that, with the continuously growing vehicle trajectory data, which means lower estimation errors and better control performances, vehicle trajectories-based adaptive methods such as QIA may be promising.

7. Conclusion

This study proposed a Queue Intensity Adaptive (QIA) signal control algorithm for isolated intersection. QIA algorithm used real-time vehicle trajectory data consisting of instantaneous vehicle positions and timestamps as the only input to perform real-time queue state estimation and adaptive signal control. The comparison of QIA, QIA-Lite, and Synchro shows that both the optimization of green time and the optimization of phase sequence are significant. The conclusions of this study are summarized as follows:

- (1) The queue state estimation method based on the integration of Kalman filters and shockwaves effectively provided real-time queue state and supported signal

control optimization in low-penetration vehicle trajectory data environment.

- (2) Based on the queue intensity that is presented to quantify the queuing pressure, five control situations, including invalid situation, long-term oversaturation, high queue intensity, short-term oversaturation, and undersaturation and low queue intensity, were defined to characterize different traffic conditions.
- (3) The min-max optimization model was applied to minimize the worst-case queue intensity of the intersection, and the robustness was ensured under fluctuated traffic conditions.
- (4) The situation-aware signal control optimization procedure was developed, and it progressively adjusted models until the intersection's queue intensity was adapted to the best control situation; thus, it accommodated varying traffic conditions.
- (5) The case study conducted at a field intersection showed that QIA algorithm with only 7.4% penetrated vehicle trajectories prevented queue spillback effectively by constraining the temporal percentage of queue spillback duration under 2.4%.
- (6) Compared with the algorithm embedded in Synchro, the extreme queue intensity, temporal percentage of queue spillback, delay, and stops were decreased by 54.7%, 97%, 22.3%, and 45.1%, respectively, and compared with Max Pressure, the above four indices were decreased by 16%, 61.5%, -1.8%, and 49.4%, respectively.

Nevertheless, the study still has limitations that the phase sequence optimization is simplified to reduce the computational complexity. In addition, the proposed method needs to be validated in more scenarios with different traffic demands and varying penetration rates. It is expected in our future work that this study will be further improved by utilizing vehicle trajectory data over multiple intersections to simultaneously control the critical paths and prevent spillback on a large scale for arterials and networks.

Data Availability

All data, models, and code generated or used during the study are included within the submitted article.

Conflicts of Interest

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

Acknowledgments

This research was jointly sponsored by the National Key Research and Development Program of China (2018YFB16005), National Natural Science Foundation of China (U1764261, 61873018), and "Dawn" Program of Shanghai Education Commission (18SG21).

References

- [1] A. G. Sims and K. W. Dobinson, "The sydney coordinated adaptive traffic (SCAT) system philosophy and benefits," *IEEE Transactions on Vehicular Technology*, vol. 29, no. 2, pp. 130–137, 1980.
- [2] N. Gartner, "OPAC: a demand-responsive strategy for traffic signal control," *Transportation Research Record*, vol. 906, pp. 75–81, 1983.
- [3] P. B. Hunt, D. I. Robertson, R. D. Bretherton, and M. C. Royle, "The scoot on-line traffic signal optimization technique," *Traffic Engineering and Control*, vol. 23, pp. 190–192, 1982.
- [4] P. Mirchandani and L. Head, "A real-time traffic signal control system: architecture, algorithms, and analysis," *Transportation Research Part C: Emerging Technologies*, vol. 9, no. 6, pp. 415–432, 2001.
- [5] Y. Yin, "Robust optimal traffic signal timing," *Transportation Research Part B: Methodological*, vol. 42, no. 10, pp. 911–924, 2008.
- [6] Y. Feng, K. L. Head, S. Khoshmaghani, and M. Zamanipour, "A real-time adaptive signal control in a connected vehicle environment," *Transportation Research Part C: Emerging Technologies*, vol. 55, pp. 460–473, 2015.
- [7] J. Zheng and H. X. Liu, "Estimating traffic volumes for signalized intersections using connected vehicle data," *Transportation Research Part C: Emerging Technologies*, vol. 79, pp. 347–362, 2017.
- [8] T. Seo, T. Kusakabe, and Y. Asakura, "Estimation of flow and density using probe vehicles with spacing measurement equipment," *Transportation Research Part C: Emerging Technologies*, vol. 53, no. 6, pp. 134–150, 2015.
- [9] M. Ramezani and N. Geroliminis, "On the estimation of arterial route travel time distribution with markov chains," *Transportation Research Part B: Methodological*, vol. 46, no. 10, pp. 1576–1590, 2012.
- [10] A. Hofleitner, R. Herring, P. Abbeel, and A. Bayen, "Learning the dynamics of arterial traffic from probe data using a dynamic bayesian network," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 4, pp. 1679–1693, 2012.
- [11] X. Ban, R. Herring, P. Hao, and A. M. Bayen, "Delay pattern estimation for signalized intersections using sampled travel times," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2130, no. 1, pp. 109–119, 2009.
- [12] D. Sun, K. Zhang, and S. Shen, "Analyzing spatiotemporal traffic line source emissions based on massive didi online car-hailing service data," *Transportation Research Part D: Transport and Environment*, vol. 62, pp. 699–714, 2018.
- [13] X. Ban, P. Hao, and Z. Sun, "Real time queue length estimation for signalized intersections using travel times from mobile sensors," *Transportation Research, Part C: Emerging Technologies*, vol. 19, no. 6, pp. 1133–1156, 2011.
- [14] M. Ramezani and N. Geroliminis, "Queue profile estimation in congested urban networks with probe data," *Computer-Aided Civil and Infrastructure Engineering*, vol. 30, no. 6, pp. 414–432, 2015.
- [15] G. Comert, "Queue length estimation from probe vehicles at isolated intersections: estimators for primary parameters," *European Journal of Operational Research*, vol. 252, no. 2, pp. 502–521, 2016.
- [16] J. Zheng, W. Sun, H. X. Liu, S. Huang, S. Shen, and C. Yu, "Traffic signal optimization using crowdsourced vehicle trajectory data," in *Proceedings of the 97th Annual Meeting of Transportation Research Board*, Washington, DC, USA, 2018.

- [17] H. Rim, C. Oh, K. Kang, and S. Kim, "Estimation of lane-level travel times in vehicle-to-vehicle and vehicle-to-infrastructure-based traffic information system," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2243, no. 1, pp. 9–16, 2011.
- [18] C. Priemer and B. Friedrich, "A decentralized adaptive traffic signal control using V2I communication data," in *Proceedings of the 12th International IEEE Conference on Intelligent Transportation Systems*, St. Louis, MI, USA, 2009.
- [19] M. Li, Z. Zou, F. Bu, and W. Zhang, "Application of vehicle infrastructure integration data on real-time arterial performance measurements," in *Proceedings of the 87th Annual Meeting of Transportation Research Record*, Washington, DC, USA, 2008.
- [20] K. Aboudolas, M. Papageorgiou, A. Kouvelas, and E. Kosmatopoulos, "A rolling-horizon quadratic-programming approach to the signal control problem in large-scale congested urban road networks," *Transportation Research Part C: Emerging Technologies*, vol. 18, no. 5, pp. 680–694, 2010.
- [21] F. V. Webster, *Traffic Signal Settings*, Transport Research Laboratory, Crowthorne, UK, 1958.
- [22] S. Chen and D. J. Sun, "An improved adaptive signal control method for isolated signalized intersection based on dynamic programming," *IEEE Intelligent Transportation Systems Magazine*, vol. 8, no. 4, pp. 4–14, 2016.
- [23] T.-H. Chang and J.-T. Lin, "Optimal signal timing for an oversaturated intersection," *Transportation Research Part B: Methodological*, vol. 34, no. 6, pp. 471–491, 2000.
- [24] W. Sun, Y. Wang, G. Yu, and H. X. Liu, "Quasi-optimal feedback control for an isolated intersection under oversaturation," *Transportation Research Part C: Emerging Technologies*, vol. 67, pp. 109–130, 2016.
- [25] Q. He, K. L. Head, and J. Ding, "Pamscod: platoon-based arterial multi-modal signal control with online data," *Transportation Research Part C: Emerging Technologies*, vol. 20, no. 1, pp. 164–184, 2012.
- [26] J. Lee, B. Park, and I. Yun, "Cumulative travel-time responsive real-time intersection control algorithm in the connected vehicle environment," *Journal of Transportation Engineering*, vol. 139, no. 10, pp. 1020–1029, 2013.
- [27] N. J. Goodall, B. L. Smith, and B. Park, "Traffic signal control with connected vehicles," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2381, no. 1, pp. 65–72, 2013.
- [28] J. Yin, J. Sun, and K. Tang, "A kalman filter-based queue length estimation method with low-penetration mobile sensor data at signalized intersections," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2672, no. 45, pp. 253–264, 2018.
- [29] P. Varaiya, "Max pressure control of a network of signalized intersections," *Transportation Research Part C: Emerging Technologies*, vol. 36, pp. 177–195, 2013.
- [30] X. Sun and Y. Yin, "A simulation study on max pressure control of signalized intersections," *Transportation Research Record*, vol. 2672, no. 18, pp. 17–127, 2018.