Transportation Research Part C 30 (2013) 93–115

Contents lists available at SciVerse ScienceDirect

Transportation Research Part C

journal homepage: www.elsevier.com/locate/trc

An adaptive control algorithm for traffic-actuated signals

Xing Zheng [a,*], Will Recker [b]

[a] Chongqing Transport Planning Institute, No. 18, Yangheercun, Jiangbei District, Chongqing 400020, China
[b] Institute of Transportation Studies, University of California at Irvine, Irvine, CA 92697, USA

ARTICLE INFO

Article history:
Received 15 March 2011
Received in revised form 17 February 2013
Accepted 17 February 2013

Keywords:
Actuated control
Adaptive control
Optimization
Microscopic simulation

ABSTRACT

A real-time, on-line control algorithm is proposed that aims to maintain the adaptive functionality of actuated controllers while improving the performance of traffic-actuated signal control system. To be consistent with the operation logic of existing signal control devices, only those four basic control parameters that can be found in modern actuated controllers are considered: phase sequence, minimum green, unit extension and maximum green. Microscopic simulation is used to test and evaluate the proposed control algorithm comparing with free-mode actuated, actuated-coordinated and volume–density control in a calibrated signalized network. Simulation results indicate that the proposed algorithm has the potential to improve the performance of the network at different traffic demand levels.

1. Introduction

Traditionally, traffic signals at a signalized intersection operate in one of two different control modes: either pre-timed or traffic-actuated (both semi-actuated and full-actuated). In pre-timed control, all of the control parameters, including cycle length, phase splits and phase sequence, are preset offline based on an assumed deterministic demand level at different time periods of day. This control mode has only a limited ability to accommodate the traffic fluctuations that are commonly found in reality. In traffic-actuated control, cycle length, phase splits and even phase sequence can be changed in response to the real-time vehicle actuations registered at loop detectors or other traffic sensors, but these changes are still subject to a set of predefined, fixed control parameters (e.g., minimum green, unit extension and maximum green, etc.) that are not accordingly responsive to the varying traffic condition.

Alternatively, an on-line control algorithm that decides real-time signal operation parameters offers a potential improvement to actuated control. Existing on-line control algorithms that have been deployed are typified by the so-called third-generation urban traffic control systems (UTCS), which can be further categorized into centralized traffic-responsive systems and distributed traffic-adaptive systems. In the centralized system, such as SCATS (Sims and Dobinson, 1979; Lowrie, 1992) and SCOOT (Hunt et al., 1981; Robertson and Bretherton, 1991), a master computer is used to adjust the cycle lengths, offsets and splits of all signals on a cycle-by-cycle basis, such that a "best" timing plan can be found for the operation of the entire network. In the distributed system, such as OPAC (Gartner, 1983) and RHODES (Head et al., 1992; Mirchandani and Head, 2001), separate calculations are taken to determine the phase sequences and durations of each signal, with an attempt to optimize the performance of each local intersection. To some extent, actuated controllers are themselves "adaptive" in view of their ability to vary the same set of outcomes as in adaptive control, but as mentioned previously, this adaptability is restricted by a set of predefined, fixed control parameters that are not adaptive to current conditions. To achieve the functionality of truly adaptive controllers, a set of online optimized phasing and timing parameters are needed.

* Corresponding author.
E-mail addresses: xingzheng1980@gmail.com (X. Zheng), wwrecker@uci.edu (W. Recker).

0968-090X/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.trc.2013.02.007

In general, three issues must be addressed in formulating an on-line optimal control problem: (1) development of a mathematical model that represents the current, or expected, traffic condition of the controlled system; (2) specification of the real-time control objective that can be expressed as a certain performance index; and (3) design of an appropriate optimization technique such that the controlled system meets the specified criteria. Mathematical models that represent the traffic condition of the controlled system, which in this research corresponds to the signalized intersections, can be classified into the following three generalized categories (Pavlis and Recker, 2004): (1) store-and-forward models (e.g., Gazis and Potts, 1963); (2) dispersion-and-store models (e.g., Pacey, 1956; Robertson, 1969); and (3) kinematic wave models (e.g., Stephanedes and Chang, 1993; Lo, 2001).

The most common objective for real-time signal control strategies is to design a signal timing plan that minimizes total (or, average) intersection control delay, which serves as the primary performance indicator of level of service at signalized intersections. Existing intersection control delay models are summarized into four categories (Dion et al., 2004): (1) deterministic queuing models (e.g., Cox and Smith, 1961), (2) shock wave models (e.g., Rorbech, 1968; Stephanopoulos and Michalopoulos, 1979; Michalopoulos et al., 1980), (3) steady-state stochastic models (e.g., Webster, 1958; Newell, 1960, 1965; McNeil, 1968; Heidemann, 1994), and (4) time-dependent stochastic models (e.g., Akcelik, 1981, 1988; Brilon and Wu, 1990; Akcelik and Rouphail, 1993; Fambro and Rouphail, 1997).

Two fundamental approaches are commonly used for on-line optimizations (Lin and Vijayakumar, 1988): binary choice approach and sequencing approach. In binary choice approach, time is divided into successive small intervals (2–4 s) and a decision is made in each interval either to extend or to terminate the current signal phase. Examples of this approach include Miller's algorithm (Miller, 1963), Traffic Optimization Logic (TOL) (Bang, 1976), Modernized Optimized Vehicle Actuation strategy (MOVA) (Vincent and Young, 1986), and Stepwise Adjustment of Signal Timing (SAST) (Lin et al., 1987). In sequencing approach, the decision-making interval is relatively longer and usually long-term optimal settings, such as phase splits and phase switching points, are specified. Typical of this approach is the Optimization Policies for Adaptive Control (OPAC) (Gartner, 1983), which incorporates a rolling horizon scheme to determine optimal signal switching sequences for a future timing period.

Additionally, the development and adoption of on-line control procedures have been hampered by two fundamental impediments to their successful implementation: (1) the theoretically-sound algorithms generally are specified in terms of those parameters and control options that are not simply within the lexicon of control devices and typically involve complex programming formulations (e.g., mixed-integer and/or piece-wise functions) that do not lend themselves to real-time solution, and (2) the practically feasible algorithms that do manipulate those parameters employed in modern control devices almost universally are formulated based on highly simplified approximations and assumptions (e.g., steady-state condition) to both control response and traffic measurement. In consideration of these aspects, this paper, as an extension of the previous work (Zheng et al., 2010), introduces an on-line control algorithm that aims to maintain the adaptive functionality of actuated controllers while improving the performance of traffic-actuated signal control system. In the interest of facilitating deployment, this algorithm is developed based on the timing protocol of the standard NEMA (National Electrical Manufacturers Association) eight-phase full-actuated dual-ring controller (ITE, 2001). In formulating the optimal control problem, a flow prediction model is developed to estimate the future vehicle arrivals at the target intersection, the traffic condition at the target intersection is described as "over-saturated" through the whole timing process, and the optimization objective is specified as to minimize total cumulative vehicle queue as an equivalent to minimizing total intersection control delay. According to the implicit timing features of actuated control, a modified rolling horizon scheme is devised to optimize the values of four basic control parameters—phase sequence, minimum green, unit extension and maximum green—based on the future flow estimation, and these optimized parameters serve as available signal timing data for further optimizations. This dynamically recursive optimization procedure properly reflects the functionality of truly adaptive controllers.

In the following section, we will address how the general issues regarding online optimal control are considered in this research. Then, an online optimization procedure that corresponds to the proposed adaptive control algorithm is introduced in detail. Next, microscopic simulation is used to test and evaluate the performance of the proposed algorithm in a calibrated network consisting of thirty-eight signals. Last, conclusions and potential future works are presented. The notations used in the model formulation appear in the Appendix.

2. Consideration of the general issues regarding online optimal control

2.1. Modeling the traffic condition at signalized intersections

Generally, the operational state of a signal control phase varies alternatively between "in service" and "not in service" within the complete timing period. Each in-service period is defined as phase split, and each not-in-service period defined as the red duration, which is actually the in-service period for those conflicting phases within the same ring. From another viewpoint, the phase state changes alternatively between "effective green" and "effective red". Effective green refers to the period from the end of start-up lost time to the beginning of clearance lost time within the same phase, and effective red refers to the period preceding effective green, which consists of total lost time and red duration. The relationships between these timing periods are shown in Fig. 1, and can be expressed by

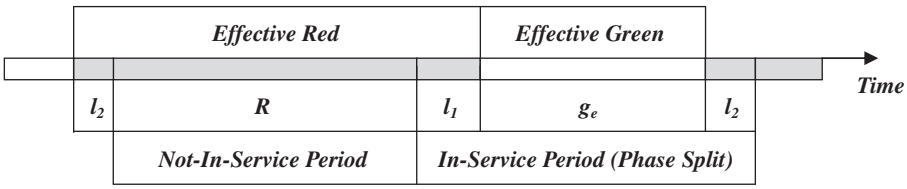

Fig. 1. Phase state.

$$r_{ei} = l_{2i} + R_i + l_{1i} = R_i + L_i \tag{1}$$

$$G_i = l_{1i} + g_{ei} + l_{2i} = g_{ei} + L_i \tag{2}$$

$$R_i = \sum_{i'} G_{i'} \tag{3}$$

where i' is the index of the preceding phases that conflict with phase i within the same ring.

Based on the definitions of effective green and effective red, the traffic condition for each phase, both under-saturated and over-saturated, can be illustrated with the queue formation and dissipation process as shown in Fig. 2. $A_i(t)$ and $D_i(t)$ are two general (smooth) functions of time that represent the cumulative number of vehicle arrivals and departures respectively for phase i. (Here, smooth functions are adopted to use differential calculus in the optimization process.) The vertical distance between the arrival and departure curves represents the number of vehicle queue at time t and equals to $A_i(t) - D_i(t)$.

By extending this approach, a combined arrival–departure curve pattern can be developed to illustrate the real-time traffic condition of the entire intersection. Fig. 3 shows an example of the combined queue accumulation curves within a signal

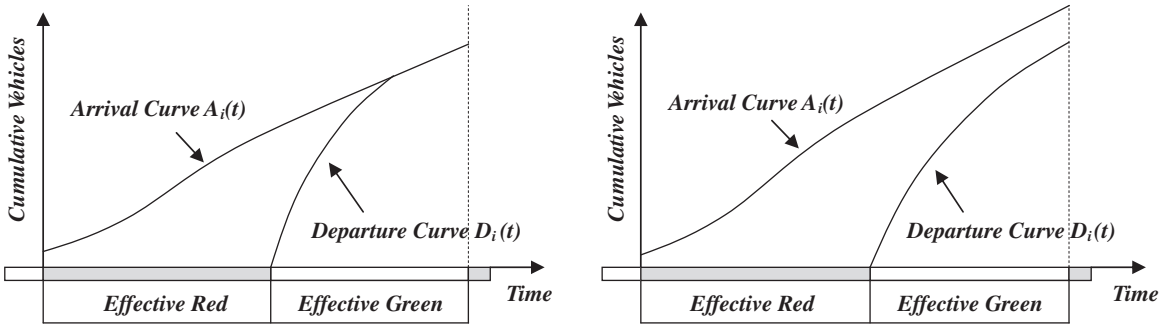

Fig. 2. Queue accumulation curves.

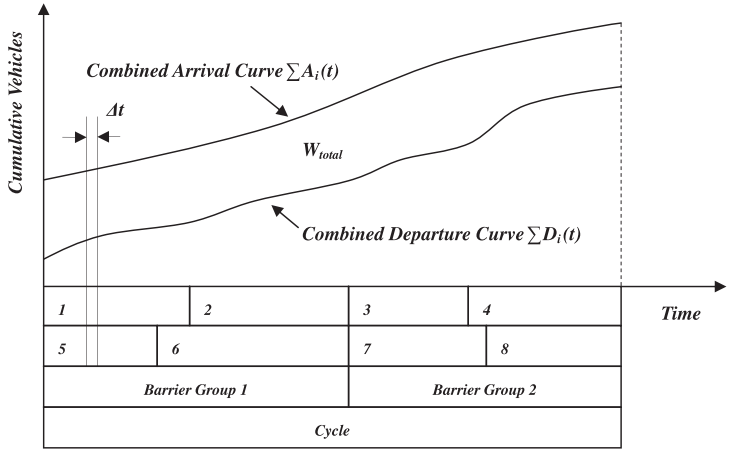

Fig. 3. Combined queue accumulation curves.

control cycle. (Note that a signal control cycle is comprised of two barrier groups and, in actuated control, both barrier groups and the cycle length are variable.) $\sum A_i(t)$ and $\sum D_i(t)$ are two general functions of time that represent the total cumulative number of vehicle arrivals and departures respectively for all eight phases ($i = 1, 2, \ldots, 8$). The vertical distance between the combined arrival and departure curves represents the total number of vehicle queue at the intersection at time t and equals to $\sum A_i(t) - \sum D_i(t)$. Theoretically speaking, since only two phases can be in service at a time (e.g., Δt), the combined departure curve will never intersect the combined arrival curve as long as conflicting phases have an arrival flow rate greater than zero. (In practice, the combined departure curve will finally meet the combined arrival curve when no conflicting phase is actuated and the signal stays in green for current phases.) Therefore, it is safe to presume that the signalized intersection is "over-saturated" through the whole timing process (i.e., $\sum A_i(t) - \sum D_i(t) > 0$).

2.2. Specifying the real-time control objective

In this research, only one objective is considered for optimization—to minimize total intersection control delay. Different from other well-known delay expressions, herein the delay experienced by each phase is considered simultaneously. Based on the combined queue accumulation curves as shown in Fig. 3, the total intersection control delay during a certain period, W_{total}, is equivalent to the area between $\sum A_i(t)$ and $\sum D_i(t)$ over this period, which can be determined by

$$W_{total} = \int_{t_1}^{t_2} \left[\sum_i A_i(t) - \sum_i D_i(t) \right] dt \tag{4}$$

where t_1 is the start moment of the target period and t_2 is the end moment of the target period.

Then, the control objective can be expressed by

$$\min W_{total} = \min \int_{t_1}^{t_2} \left[\sum_i A_i(t) - \sum_i D_i(t) \right] dt \tag{5}$$

2.3. Designing the on-line optimization technique

Since the advent of Miller's algorithm (Miller, 1963), rolling horizon scheme (or, variations of it) has become increasingly attractive regarding on-line control procedures (e.g., Robertson and Bretherton, 1974; Gartner, 1983; Bell, 1990; Heydecker, 1990). In this scheme, a project horizon is predetermined which generally consists of N time intervals, as shown in Fig. 4. Flow data are measured for the first H intervals (head portion) and are estimated from a model for the next $N - H$ intervals (tail portion). Optimal phase splits and/or phase switching points are specified based on these data so as to minimize the total delay (or, optimize other performance indices) over an upcoming target period (e.g., a length of time equal to the horizon). Then, the project horizon is shifted into the future by R intervals (roll period) and the same process repeats for the next iteration. Usually, the roll period is equal to the length of the head portion, and it can be as small as one time interval (say, 2–4 s).

Without distorting the principle, in this section we introduce a modified rolling horizon scheme that is developed based on the implicit timing features of actuated control (refer to Fig. 5 for example):

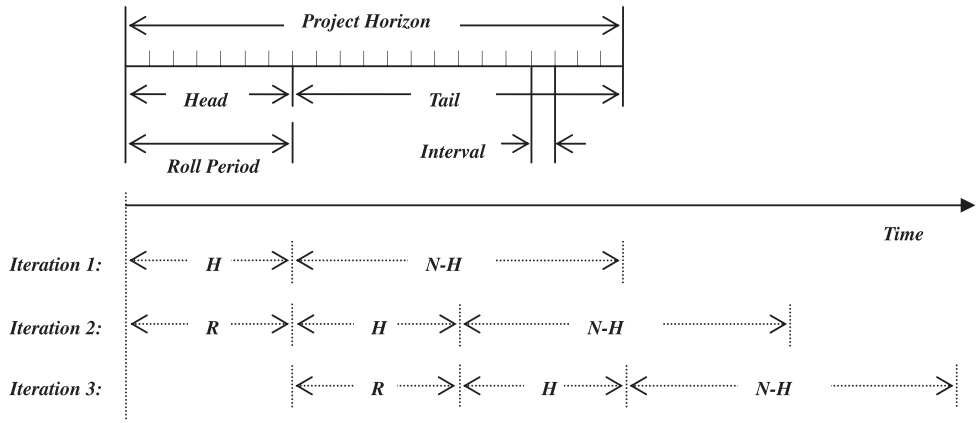

Fig. 4. Rolling horizon scheme.

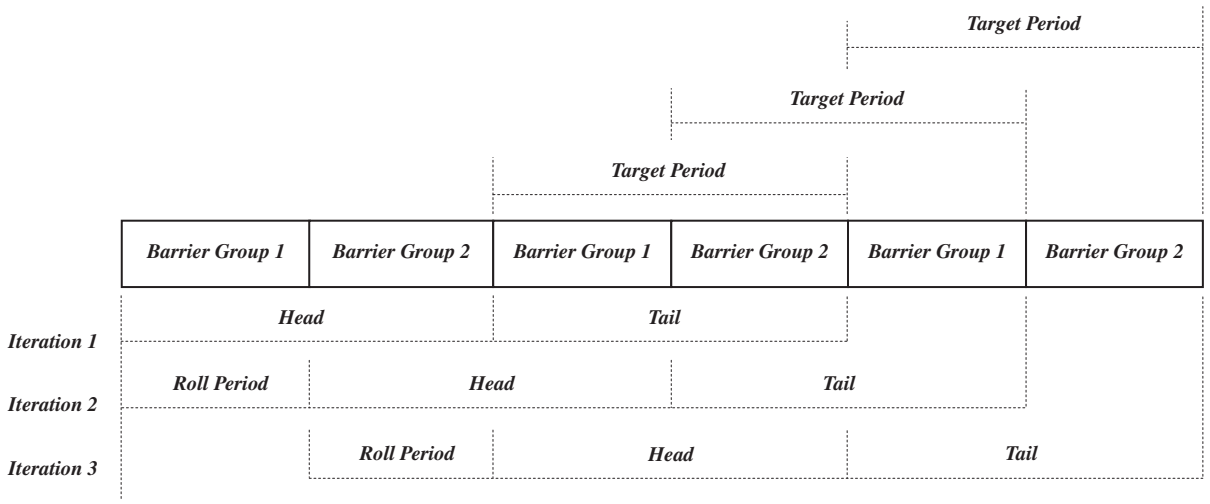

Fig. 5. Modified rolling horizon scheme.

Step 1: The head portion of the project horizon is set equal to the summed length of any two consecutive barrier groups (which can be viewed as a floating cycle and is also variable), during which the signal timing data that include phase splits and parameter settings (i.e., phase sequence, min green, max green and unit extension) are fully exploited to obtain the vehicle arrival flow for each actuated phase.

Step 2: Based on the flow data obtained from step 1, a flow prediction model is used to estimate the future vehicle arrival flow for each signal phase (and thus the total arrival flow for the entire intersection) over the tail portion. The tail portion is denoted the target period and equal to the floating cycle that follows the head portion.

Step 3: Based on the flow estimation obtained from step 2, a new signal timing plan is specified with an objective to minimize the total intersection control delay (as expressed in Section 2.2) over the tail portion. This new plan includes the optimal phase sequence for the tail portion, and the optimal min green, unit extension and max green for each signal phase.

Step 4: The project horizon is shifted into the future by a time period that is equal to the length of one barrier group (i.e., roll period), and then the optimization process repeats from step 1.

As can be realized, this modified rolling horizon scheme features the following three major differences:

(1) The head portion, tail portion (i.e., target period) and roll period are all time-variant since they are actually determined by the phase splits, which in essence are determined by the signal timing plan and real-time vehicle actuations.
(2) Optimal control parameters are all determined based on flow estimations, which may degrade the performance of the optimization procedure. However, these optimized control parameters that are used in the upcoming period will serve to be a feed-back mechanism (especially the unit extension and max green) that has the ability to adjust signal timing in response to real-time vehicle actuations.
(3) Since the roll period is equal to the length of one barrier group, the optimization frequency becomes once per floating cycle and the parameter specifications are only implemented in one (i.e., the upcoming) barrier group.

3. Methodology

3.1. Overview

Based on the discussion in Section 2, an on-line optimal control algorithm is proposed here that, together with signal control principles, constitutes a dynamically recursive optimization procedure, as shown in Fig. 6. This procedure consists of four major modules: Data Processing, Flow Prediction, Parameter Optimization and Signal Control. The first three modules correspond to the first three steps of the modified rolling horizon scheme, except that vehicle spillover (of each phase) is also obtained from step 1 and used as another input for step 3. The fourth module corresponds to the signal control system that employs the optimized control parameters output from step 3. The resulting phase splits, as well as those parameters that have been applied, will be used as inputs for step 1 in the next iteration, which begins a roll period later as stated in step 4.

Since this research focuses on maintaining the adaptive functionality of actuated controllers while improving the performance of traffic-actuated signal control system, which mainly applies to isolated intersections, it is assumed for simplicity that the vehicle arrival pattern associated with each vehicle movement conforms to Poisson process. (Strictly speaking, this assumption limits the application of this algorithm to non-congested periods of operation.) Under other conditions, such as

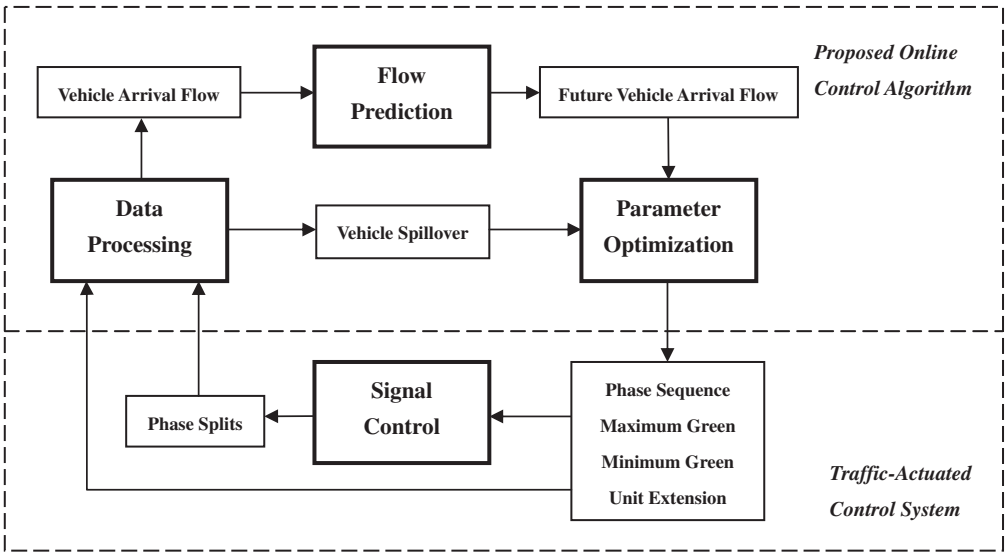

Fig. 6. Recursive optimization procedure.

the signalized network in urban areas, the concepts of control developed in this paper may still apply, for example, via replacing the Poisson-related arrival pattern by other appropriate distributions (e.g., the well-known Robertson's dispersion model). Besides, the proposed control algorithm is formulated based on the basic operation logic of full-actuated controllers. Other actuated control modes, such as semi-actuated and volume–density, are not considered. A few more assumptions and explications are addressed below:

1. The signal timing process for each actuated phase relies on the basic gap-seeking logic. Special operation parameters, such as recall, are not considered.
2. The signal timing of each cycle (or, floating cycle) relies on the basic dual-ring control process. Special operation functions, such as dual-entry, simultaneous gap-out and conditional service, are not considered.
3. The vehicle arrival flow rate for each signal phase is assumed to be constant within each cycle (or, floating cycle). The saturation flow rate for each signal phase is assumed to be constant during the whole timing process. The vehicle arrival flow rate is assumed to be strictly less than the saturation flow rate.
4. The effective green time for each signal phase is assumed to be equal to the actual displayed green interval. The total lost time for each signal phase is assumed to be constant and equal to the sum of yellow change and all clearance intervals (which are pre-set and constant).
5. Vehicles are assumed to be equivalently distributed on each lane of multi-lane approaches, and queue vertically at the stop line for both through and left-turn movement phases.
6. Right-turn movements are assumed to be serviced on exclusive right-turn lanes and have negligible effect on the signal operation. U-turn movements are not considered.
7. All timing-related variables take expected or average values.

3.2. Data processing

According to Highway Capacity Manual (TRB, 2000), effective green can be broken into queue service time and green extension period (refer to Fig. 7). During the queue service time, vehicles discharge at saturation flow rate until the queue dissipates. The total number of these vehicles is equal to the sum of initial queue, if any, plus those vehicles that arrive during the effective red and queue service time. Therefore,

$$S_i \times G_{qi} = Q_i + \lambda_i \times (R_i + L_i) + \lambda_i \times G_{qi} \tag{6}$$

Then, the queue service time can be expressed by

$$G_{qi} = [Q_i + \lambda_i \times (R_i + L_i)]/(S_i - \lambda_i) \tag{7}$$

During the green extension period, arriving vehicles travel through the intersection freely until the current phase terminates by gap-out control—the green interval terminates when no vehicles actuate the extension detector within a unit extension period, i.e., when the vehicle headway (in time) larger than the unit extension occurs. According to Poisson process, vehicle

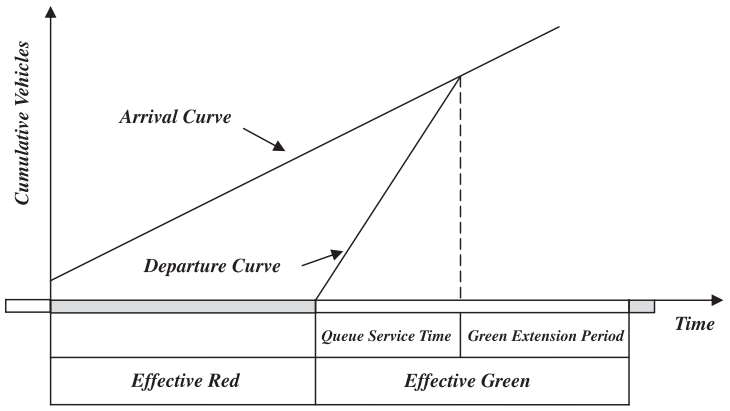

Fig. 7. Queue accumulation diagram.

headways have a negative exponential distribution and the number (n) of headways until the first one that invokes gap-out (i.e., larger than β) has a geometric distribution. Therefore,

$$P_{robability}(\text{number of headways until the first one larger than } \beta_i \text{ occurs}) = [1-\exp(-\lambda_i\beta_i)]^{n-1}\exp(-\lambda_i\beta_i)$$
$$M_{ean}(\text{number of headways until the first one larger than } \beta_i \text{ occurs}) = \exp(\lambda_i\beta_i)$$

During the $e^{\lambda\beta}$ vehicle headways, a total of $(e^{\lambda\beta}-1)$ vehicles are serviced. Thus, the green extension period can be expressed by

$$G_{ei} = [\exp(\lambda_i\beta_i)-1]/\lambda_i \tag{8}$$

It should be noted here that the green-time estimation model mentioned by HCM is formulated based on two indefensible assumptions: (1) the extension detectors are placed at the stop line and so the green extension period starts timing right after the queue service time expires, and (2) the phase definitely terminates by gap-out control. In real field situations, however, the extension detectors, especially those for through-movement phases, are set back a certain distance (e.g., 300 ft) from the stop line, and max-out control materializes frequently when the traffic demand reaches a relatively higher level. In addition, according to the timing logic of actuated control phases, the green extension period starts timing exactly when minimum green expires, so it is possible that the phase terminates while the vehicle queue has not yet dissipated. To avoid all these shortcomings, herein we take into account both the gap-out/max-out information and the signal timing data that includes phase split, min green, max green and unit extension to determine the effective green time for each actuated phase. Furthermore, the vehicle arrival flow rate and spillover are also determined. Specifically, it is assumed here that when minimum green expires, the vehicle queue length has been decreased to be shorter than the distance from the trailing edge of the extension detector to the stop line.

(1) Gap-out situation

When a signal phase terminates by gap-out control, the green extension period, i.e., the period from the end of minimum green to the end of green interval, can still be expressed by Eq. (8) unless the arrival pattern changes. Therefore, the effective green time is equal to the minimum green plus green extension period, i.e.,

$$g_{ei} = G_{\min i} + G_{ei} = G_{\min i} + [\exp(\lambda_i\beta_i)-1]/\lambda_i \tag{9}$$

Then, according to Eq. (2), the phase split can be expressed by

$$G_i = G_{\min i} + [\exp(\lambda_i\beta_i)-1]/\lambda_i + L_i \tag{10}$$

In Eq. (10), all variables except λ_i are known signal timing data obtained from phase i, and thus the vehicle arrival flow rate, λ_i, can be determined by solving the nonlinear inverse function $F^{-1}(\lambda_i)$, i.e.,

$$\lambda_i = F^{-1}(\lambda_i);\ \text{where } F(\lambda_i) = G_{\min i} + [\exp(\lambda_i\beta_i)-1]/\lambda_i + L_i - G_i = 0 \tag{11}$$

To determine the vehicle spillover, the following three cases that describe different gap-out situations need to be considered (refer to Fig. 8):

1. $G_{qi} \leqslant G_{\min i},\ \text{i.e.,}\ [Q_i + \lambda_i \times (R_i + L_i)]/(S_i - \lambda_i) \leqslant G_{\min i},\ \text{or}$
$\lambda_i \leqslant (S_i \times G_{\min i} - Q_i)/(R_i + L_i + G_{\min i})$ (12)

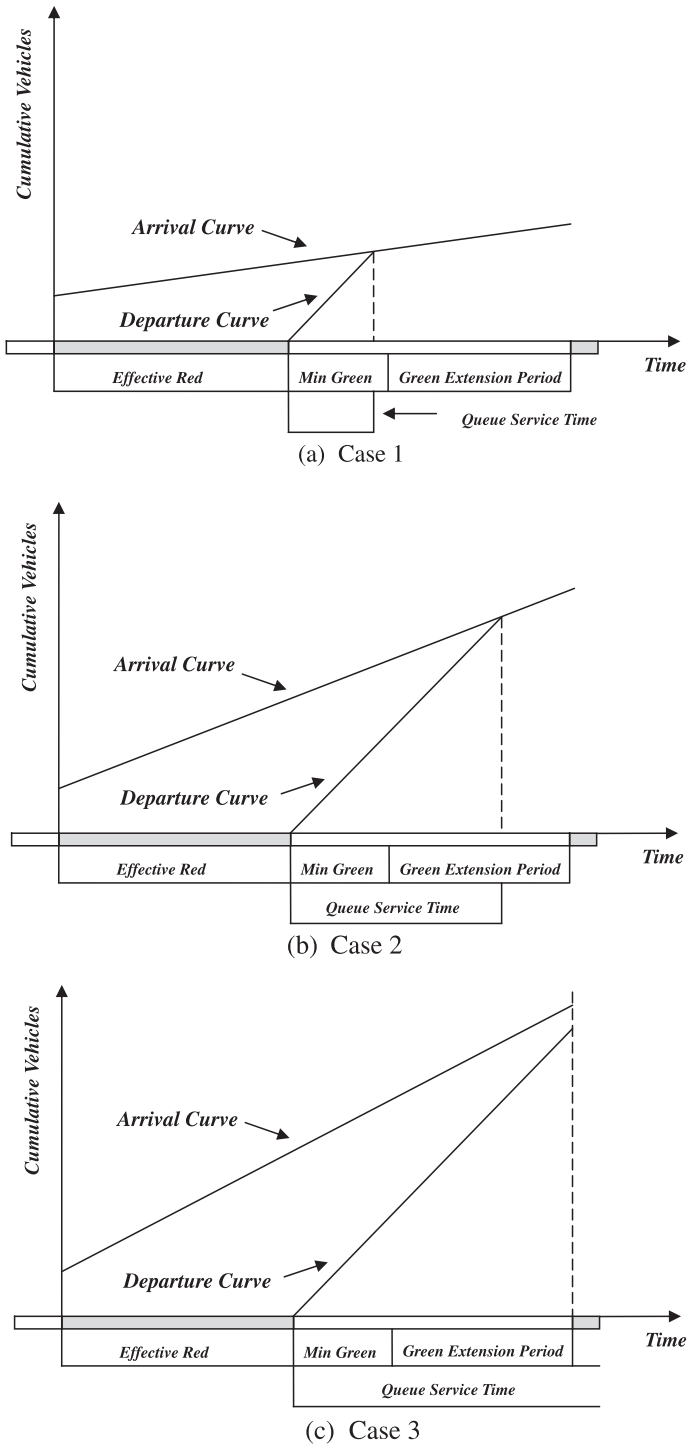

Fig. 8. Gap-out situations.

2. $G_{\min i} < G_{qi} \leqslant g_{ei}$, i.e., $G_{\min i} < [Q_i + \lambda_i \times (R_i + L_i)]/(S_i - \lambda_i) \leqslant G_i - L_i$, or
$(S_i \times G_{\min i} - Q_i)/(R_i + L_i + G_{\min i}) < \lambda_i \leqslant [S_i \times (G_i - L_i) - Q_i]/(R_i + G_i)$ (13)

$$3.\quad g_{ei} < G_{qi},\ \text{i.e.},\ G_i - L_i < [Q_i + \lambda_i \times (R_i + L_i)]/(S_i - \lambda_i),\ \text{or}\ \lambda_i > [S_i \times (G_i - L_i) - Q_i]/(R_i + G_i) \tag{14}$$

As can be seen, three consecutive, non-overlapped numerical intervals regarding the value of λ_i are illustrated by inequalities (12)–(14), which are also expressed in terms of known timing data. Based on the λ_i determined by Eq. (11), only one inequality (i.e., only one case) is "true". Therefore, the vehicle spillover can be determined by one of the following equations corresponding to the true case.

$$\begin{aligned}&\text{Case 1, 2}: {}^{spill}Q_i = 0;\\ &\text{Case 3}: {}^{spill}Q_i = Q_i + \lambda_i \times (R_i + G_i) - S_i \times (G_i - L_i)\end{aligned} \tag{15}$$

(2) Max-out situation

When a signal phase terminates by max-out control, the green extension period, i.e., the period from the end of minimum green to the end of green interval, is determined by

$$G_{ei} = G_{\max i} - G_{\min i} \tag{16}$$

And, the effective green is equal to maximum green, i.e.,

$$g_{ei} = G_{\max i} \tag{17}$$

Then, the phase split is expressed by

$$G_i = G_{\max i} + L_i \tag{18}$$

Unlike Eq. (10), although the variables in Eq. (18) are also known signal timing data obtained from phase i, the vehicle arrival flow rate, λ_i, cannot be determined. However, considering that, after the minimum green expires, arriving vehicles keep actuating the extension detector until the maximum green limit is reached, it is safe to presume that the expected time for the first headway larger than β_i to occur, as determined by Eq. (8), is greater than the green extension period as determined by Eq. (16), i.e.,

$$[\exp(\lambda_i \beta_i) - 1]/\lambda_i > G_{\max i} - G_{\min i} \tag{19}$$

Note that the term $(e^{\lambda\beta} - 1)/\lambda$ in inequality (19) is a monotone increasing function of λ, so the minimum value of λ_i, denoted by λ_{mini} here, can be achieved by solving another nonlinear inverse function $F^{-1}(\lambda_i)$, i.e,

$$\lambda_{\min i} = F^{-1}(\lambda_i);\ \text{where } F(\lambda_i) = [\exp(\lambda_i \beta_i) - 1]/\lambda_i - (G_{\max i} - G_{\min i}) = 0 \tag{20}$$

Specifically, since the vehicle arrival flow rate is strictly less than the saturation flow rate, we assume here that λ_i is approximately equal to the mean of λ_{mini} and S_i, i.e.,

$$\lambda_i = (\lambda_{\min i} + S_i)/2 \tag{21}$$

Again, to determine the vehicle spillover, the following three cases that describe different max-out situations need to be considered (refer to Fig. 9):

$$1.\quad G_{qi} \leqslant G_{\min i},\ \text{i.e.},\ [Q_i + \lambda_i \times (R_i + L_i)]/(S_i - \lambda_i) \leqslant G_{\min i},\ \text{or}\ \lambda_i \leqslant (S_i \times G_{\min i} - Q_i)/(R_i + L_i + G_{\min i}) \tag{22}$$

$$2.\quad G_{\min i} < G_{qi} \leqslant g_{ei},\ \text{i.e.},\ G_{\min i} < [Q_i + \lambda_i \times (R_i + L_i)]/(S_i - \lambda_i) \leqslant G_{\max i},\ \text{or}\ (S_i \times G_{\min i} - Q_i)/(R_i + L_i + G_{\min i}) < \lambda_i \leqslant (S_i \times G_{\max i} - Q_i)/(R_i + L_i + G_{\max i}) \tag{23}$$

$$3.\quad g_{ei} < G_{qi},\ \text{i.e.},\ G_{\max i} < [Q_i + \lambda_i \times (R_i + L_i)]/(S_i - \lambda_i),\ \text{or}\ \lambda_i > (S_i \times G_{\max i} - Q_i)/(R_i + L_i + G_{\max i}) \tag{24}$$

Similarly, three consecutive, non-overlapped numerical intervals regarding the value of λ_i are illustrated by inequalities (22)–(24), which are also expressed in terms of known timing data. Based on the λ_i determined by Eq. (21), only one inequal-

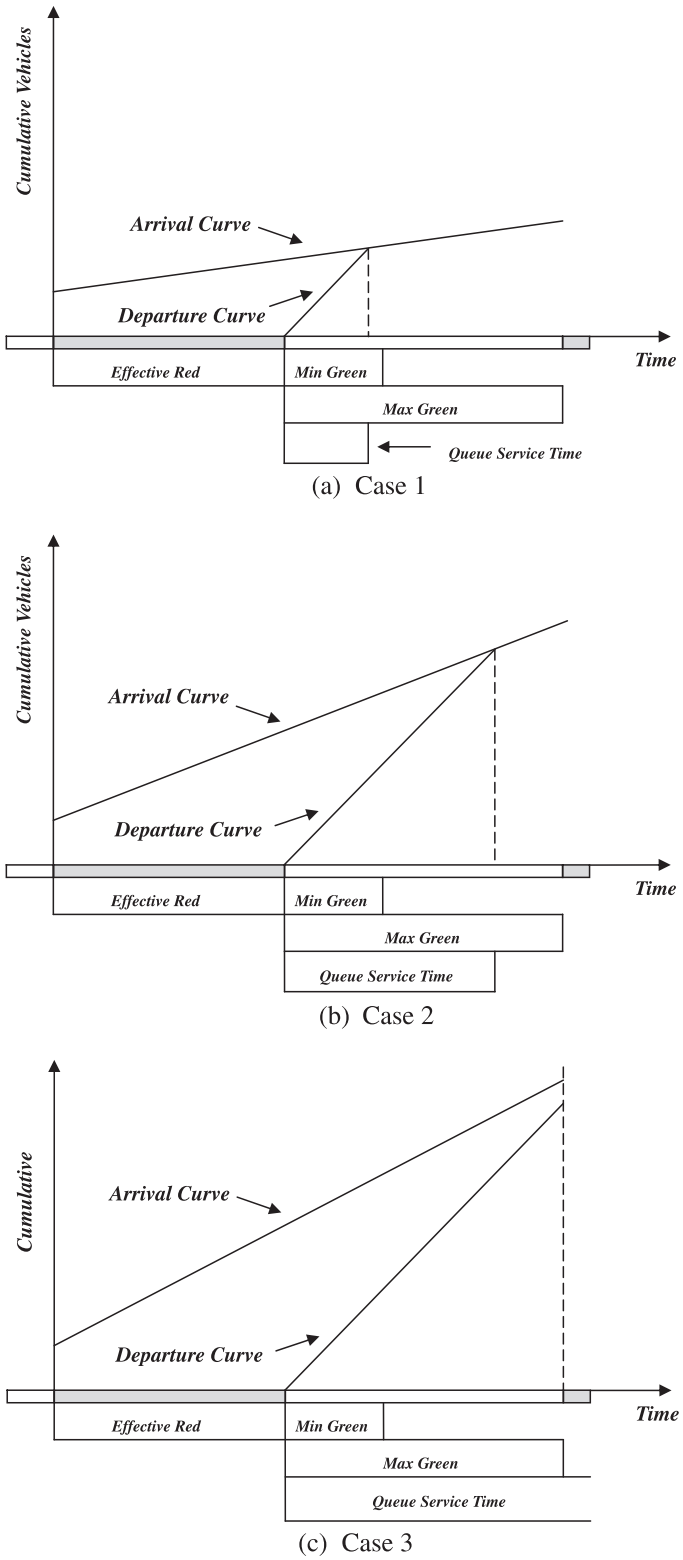

Fig. 9. Max-out situations.

ity is "true". Therefore, the vehicle spillover can be determined by one of the following equations corresponding to the true case.

$$\begin{aligned}&\text{Case 1, 2}: {}^{spill}Q_i = 0;\\ &\text{Case 3}: {}^{spill}Q_i = Q_i + \lambda_i \times (R_i + L_i + G_{\max i}) - S_i \times G_{\max i}\end{aligned} \tag{25}$$

3.3. Flow prediction

According to the modified rolling horizon scheme, we here employ a dynamic exponential smoothing model to estimate the future vehicle arrival flow for each signal phase:

Define

$\left(\lambda_i^{head}\right)^m$	the vehicle arrival flow rate associated with phase i over the head portion in iteration m, which is determined by Eq. (11) for gap-out cases or by Eq. (21) for max-out cases
$\left(\lambda_i^{tail}\right)^m$	the vehicle arrival flow rate associated with phase i over the tail portion in iteration m, which is predicted based on $\left(\lambda_i^{head}\right)^m$ and viewed as the estimated value for $\left(\lambda_i^{head}\right)^{m+1}$
a_i^m	the smoothing factor associated with phase i in iteration m

Suppose the target period is the tail portion in iteration $m+1$, then

$$\begin{aligned}&\left(\lambda_i^{tail}\right)^{m+1} = a_i^{m+1} \times \left(\lambda_i^{head}\right)^{m+1} + \left(1 - a_i^{m+1}\right) \times \left(\lambda_i^{tail}\right)^m,\ 0 < a_i^{m+1} \leqslant 1,\ \text{or}\\ &\left(\lambda_i^{tail}\right)^{m+1} = \left(\lambda_i^{tail}\right)^m + a_i^{m+1} \times \left[\left(\lambda_i^{head}\right)^{m+1} - \left(\lambda_i^{tail}\right)^m\right],\ 0 < a_i^{m+1} \leqslant 1\end{aligned} \tag{26}$$

The value of smoothing factor a^m can be any positive decimals ($\leqslant 1$), and is usually set small for slowly evolving conditions and large for significant transients. The properties of exponential smoothing model can be found elsewhere (Chou, 1970). Here, the initial value of a^m is arbitrarily determined (e.g., 0.3), and, if significant errors exist, it will be updated every floating cycle by replacing the estimated value with the true value, i.e.,

According to Eq. (26), we have

$$\left(\lambda_i^{tail}\right)^m = \left(\lambda_i^{tail}\right)^{m-1} + a_i^m \times \left[\left(\lambda_i^{head}\right)^m - \left(\lambda_i^{tail}\right)^{m-1}\right],\ 0 < a_i^m \leqslant 1$$

If

$$\left[\left(\lambda_i^{head}\right)^m - \left(\lambda_i^{tail}\right)^{m-1}\right] \Big/ \left(\lambda_i^{head}\right)^m > \delta, \quad \text{where } \delta \text{ is a predetermined tolerance value, say, } 0.1$$

Then, replace the first term on the right hand side $(\lambda^{tail})^{m-1}$ by $(\lambda^{head})^m$, and the updated smoothing factor, a^{m+1}, is determined by

$$a_i^{m+1} = \left[\left(\lambda_i^{tail}\right)^m - \left(\lambda_i^{head}\right)^m\right] \Big/ \left[\left(\lambda_i^{head}\right)^m - \left(\lambda_i^{tail}\right)^{m-1}\right] \tag{27}$$

Particularly noteworthy here is that, since the head portion (and thus the tail portion) is rolled into the future by a time period equal to one barrier group, in which only up to four phases are serviced, this prediction model can only apply to those (up to four) actuated phases in the barrier group that has just expired, as only the arrival flow rates of these phases can be obtained for the new head portion in the next iteration.

3.4. Parameter optimization

3.4.1. Determining optimal phase sequence

The optimal phase sequence is determined along with the specification of optimal phase splits, with the objective to be minimizing the total intersection control delay over the tail portion. (Note that these optimal phase splits are not parameter settings but will be used to determine the other three parameters.) Recalling Eq. (4) for the expression of total intersection control delay, we then have the total delay for the tail portion expressed by

$$\begin{aligned}&W_{total}^{tail} = \int_{t_1}^{t_2} \left[\sum_i A_i^{tail}(t) - \sum_i D_i^{tail}(t)\right] dt,\ \text{or}\\ &W_{total}^{tail} = \sum_i \int_{t_1}^{t_2} A_i^{tail}(t)dt - \sum_i \int_{t_1}^{t_2} D_i^{tail}(t)dt\end{aligned} \tag{28}$$

where t_1 is the start moment of the tail portion and t_2 is the end moment of the tail portion. Based on the future vehicle arrival flow rate determined by Eq. (26), the term $A_i^{tail}(t)$ in Eq. (28) can be expressed by

$$A_i^{tail}(t) = \lambda_i^{tail} \cdot t$$

Therefore,

$$\sum_i \int_{t_1}^{t_2} A_i^{tail}(t)dt = \sum_i \int_{t_1}^{t_2} \lambda_i^{tail} \cdot t \cdot dt = \sum_i \lambda_i^{tail} \cdot \int_{t_1}^{t_2} t \cdot dt = \frac{1}{2}(t_2^2 - t_1^2)\sum_i \lambda_i^{tail} \tag{29}$$

Since t_1 is actually the start moment of the target period, it can be set equal to zero for simplicity, and thus t_2 refers to the length of the target period and equals to the cycle length of the tail portion, i.e.,

$$t_2 = G_1^{tail} + G_2^{tail} + G_3^{tail} + G_4^{tail} = G_5^{tail} + G_6^{tail} + G_7^{tail} + G_8^{tail}, \quad \text{or } t_2 = \frac{1}{2}\sum_i G_i^{tail}$$

Then, Eq. (29) can be revised into

$$\sum_i \int_{t_1}^{t_2} A_i^{tail}(t)dt = \frac{1}{2} \cdot \left(\frac{1}{2}\sum_i G_i^{tail}\right)^2 \cdot \left(\sum_i \lambda_i^{tail}\right) = \left(\frac{1}{8}\sum_i \lambda_i^{tail}\right) \cdot \left(\sum_i G_i^{tail}\right)^2 \tag{30}$$

To express the term $D_i^{tail}(t)$ in Eq. (28), which is the cumulative number of departures at time t associated with phase i over the tail portion, we expect a clearance condition that, the phase green interval will be large enough to service the initial queue, if any, plus all the vehicles that arrive during the effective red and effective green, i.e., the phase is expected to be terminated without invoking any vehicle spillovers (refer to Case 1 and Case 2 in both gap-out and max-out situations). Therefore, $D_i^{tail}(t)$ can be expressed by

$$D_i^{tail}(t) = \begin{cases} S_i \cdot t, & t_{i1} \leqslant t < t_{i2} \\ \lambda_i^{tail} \cdot t, & t_{i2} \leqslant t < t_{i3}, t_1 < t_{i1} < t_{i2} < t_{i3} < t_2 \\ 0, & \text{otherwise} \end{cases}$$

where t_{i1} is the start moment of the queue service time (i.e., G_{qi}^{tail}) associated with phase i; t_{i2} is the end moment of the queue service time associated with phase i, equal to $t_{i1} + G_{qi}^{tail}$; and t_{i3} is the end moment of the effective green associated with phase i, equal to $t_{i1} + G_i^{tail} - L_i$. Therefore,

$$\begin{aligned} \sum_i \int_{t_1}^{t_2} D_i^{tail}(t)dt &= \sum_i \left(\int_{t_{i1}}^{t_{i2}} S_i \cdot t \cdot dt + \int_{t_{i2}}^{t_{i3}} \lambda_i^{tail} \cdot t \cdot dt\right) = \sum_i \left(\int_{t_{i1}}^{t_{i2}} S_i \cdot t \cdot dt + \int_{t_{i1}}^{t_{i3}} \lambda_i^{tail} \cdot t \cdot dt - \int_{t_{i1}}^{t_{i2}} \lambda_i^{tail} \cdot t \cdot dt\right) \\ &= \sum_i \left[\left(S_i - \lambda_i^{tail}\right) \cdot \int_{t_{i1}}^{t_{i2}} t \cdot dt + \lambda_i^{tail} \cdot \int_{t_{i1}}^{t_{i3}} t \cdot dt\right] = \frac{1}{2}\sum_i \left[\left(S_i - \lambda_i^{tail}\right)(t_{i2}^2 - t_{i1}^2) + \lambda_i^{tail}(t_{i3}^2 - t_{i1}^2)\right] \\ &= \frac{1}{2}\sum_i \left[\left(S_i - \lambda_i^{tail}\right)(t_{i2} - t_{i1})(t_{i2} + t_{i1}) + \lambda_i^{tail}(t_{i3} - t_{i1})(t_{i3} + t_{i1})\right] \end{aligned}$$

Recall $t_{i2} = t_{i1} + G_{qi}^{tail}$ and $t_{i3} = t_{i1} + G_i^{tail} - L_i$, then

$$\begin{aligned} \sum_i \int_{t_1}^{t_2} D_i^{tail}(t)dt &= \frac{1}{2}\sum_i \left[\left(S_i - \lambda_i^{tail}\right)G_{qi}^{tail}\left(2t_{i1} + G_{qi}^{tail}\right) \right. \\ &\quad \left. + \lambda_i^{tail}\left(G_i^{tail} - L_i\right)\left(2t_{i1} + G_i^{tail} - L_i\right)\right], \text{ or} \\ \sum_i \int_{t_1}^{t_2} D_i^{tail}(t)dt &= \frac{1}{2}\sum_i \left\{\left[\left(S_i - \lambda_i^{tail}\right)G_{qi}^{tail} + \lambda_i^{tail}\left(G_i^{tail} - L_i\right)\right]2t_{i1} \right. \\ &\quad \left. + \left(S_i - \lambda_i^{tail}\right)\left(G_{qi}^{tail}\right)^2 + \lambda_i^{tail}\left(G_i^{tail} - L_i\right)^2\right\} \end{aligned} \tag{31}$$

Recall Eq. (7) for the expression of queue service time, then

$$G_{qi}^{tail} = \left[Q_i^{tail} + \lambda_i^{tail} \times \left(R_i^{tail} + L_i\right)\right] \Big/ \left(S_i - \lambda_i^{tail}\right) \tag{7'}$$

In Eq. (7'), term Q_i^{tail} refers to the initial queue associated with phase i over the tail portion, and equals to the vehicle spillover over the head portion, which is determined by Eq. (15) for gap-out cases or Eq. (25) for max-out cases. Therefore

$$Q_i^{tail} = {}^{spill}Q_i^{head} \tag{32}$$

Then Eq. (7') can be revised into

$$G_{qi}^{tail} = \left[{}^{spill}Q_i^{head} + \lambda_i^{tail} \times \left(R_i^{tail} + L_i\right)\right] / \left(S_i - \lambda_i^{tail}\right) \tag{7''}$$

Substituting Eq. (7'') into Eq. (31) yields

$$\sum_i \int_{t_1}^{t_2} D_i^{tail}(t)dt = \frac{1}{2}\sum_i \left\{ \left[{}^{spill}Q_i^{head} + \lambda_i^{tail}\left(R_i^{tail} + G_i^{tail}\right)\right] 2t_{i1} + \frac{\left[{}^{spill}Q_i^{head} + \lambda_i^{tail}\left(R_i^{tail} + L_i\right)\right]^2}{S_i - \lambda_i^{tail}} + \lambda_i^{tail}\left(G_i^{tail} - L_i\right)^2 \right\}$$

$$= \sum_i \left[{}^{spill}Q_i^{head} + \lambda_i^{tail}\left(R_i^{tail} + G_i^{tail}\right)\right] t_{i1} + \frac{1}{2}\sum_i \left\{ \frac{\left[{}^{spill}Q_i^{head} + \lambda_i^{tail}\left(R_i^{tail} + L_i\right)\right]^2}{S_i - \lambda_i^{tail}} + \lambda_i^{tail}\left(G_i^{tail} - L_i\right)^2 \right\} \quad (33)$$

Thus, by substituting Eqs. (30) and (33) into Eq. (28), we have

$$W_{total}^{tail} = \left(\frac{1}{8}\sum_i \lambda_i^{tail}\right) \cdot \left(\sum_i G_i^{tail}\right)^2 - \sum_i \left[{}^{spill}Q_i^{head} + \lambda_i^{tail}\left(R_i^{tail} + G_i^{tail}\right)\right] \cdot t_{i1}$$

$$- \frac{1}{2}\sum_i \left\{ \frac{\left[{}^{spill}Q_i^{head} + \lambda_i^{tail}\left(R_i^{tail} + L_i\right)\right]^2}{S_i - \lambda_i^{tail}} + \lambda_i^{tail}\left(G_i^{tail} - L_i\right)^2 \right\} \quad (34)$$

Note that in Eq. (34), the expression of term R_i^{tail} depends on the two phase sequences over both the head and tail portions, either of which can be any one of the 32 possible phase orders as shown in Fig. 10.

As an example, here we suppose the phase sequence over the head portion is order #1 (which is known) and the phase sequence over the tail portion is going to be order #2; then, referring to Fig. 11, we have

$$R_1^{tail} = G_2^{head} + G_3^{head} + G_4^{head}$$
$$R_2^{tail} = G_3^{head} + G_4^{head} + G_1^{tail}$$
$$R_3^{tail} = G_4^{head} + G_1^{tail} + G_2^{tail}$$
$$R_4^{tail} = G_1^{tail} + G_2^{tail} + G_3^{tail}$$
$$R_5^{tail} = G_6^{head} + G_7^{head} + G_8^{head}$$
$$R_6^{tail} = G_7^{head} + G_8^{head} + G_5^{tail}$$
$$R_7^{tail} = G_5^{tail} + G_6^{tail} + G_8^{tail}$$
$$R_8^{tail} = G_5^{tail} + G_6^{tail}$$

The expression for term t_{i1} depends on the phase sequence over the tail portion. Again, referring to Fig. 11, for example, we have

$$t_{11} = 0$$
$$t_{21} = G_1^{tail}$$
$$t_{31} = G_1^{tail} + G_2^{tail}$$
$$t_{41} = G_1^{tail} + G_2^{tail} + G_3^{tail}$$
$$t_{51} = 0$$
$$t_{61} = G_5^{tail}$$
$$t_{71} = G_5^{tail} + G_6^{tail} + G_8^{tail}$$
$$t_{81} = G_5^{tail} + G_6^{tail}$$

Note here that the start moment of queue service time is assumed to be equivalent to the start moment of phase split. As can be realized, both R_i^{tail} and t_{i1} can also be expressed in terms of the phase split G_i^{tail}, and of course, there are a total of 32 possible expressions due to the 32 possible phase order numbers. Here we define

${}^kW_{total}^{tail}$	the total intersection control delay over the tail portion corresponding to phase order k, where k	1, 2, . . . , 32 as specified in Fig. 10
${}^kG_i^{tail}$	the phase split associated with phase i over the tail portion corresponding to phase order k	
${}^kR_i^{tail}$	the red duration associated with phase i over the tail portion corresponding to phase order k	
${}^k t_{i1}$	the start moment of phase i over the tail portion corresponding to phase order k	

Order #	Barrier Group 1
1	1 2 / 5 6
2	1 2 / 6 5
3	2 1 / 5 6
4	2 1 / 6 5

C_4^1 $C_2^1 \times C_2^1$

Barrier Group 1

P_2^2 $P_2^2 \cdot C_4^1 \cdot C_4^1$

Barrier Group 2

C_4^1 $C_2^1 \times C_2^1$

Order #	Barrier Group 2
1	3 4 / 7 8
2	3 4 / 8 7
3	4 3 / 7 8
4	4 3 / 8 7

Order #	Floating Cycle				Order #	Floating Cycle			
	Barrier Group 1		Barrier Group 2			Barrier Group 2		Barrier Group 1	
1	1 2 / 5 6	3 4 / 7 8			17	3 4 / 7 8	1 2 / 5 6		
2	1 2 / 5 6	3 4 / 8 7			18	3 4 / 7 8	1 2 / 6 5		
3	1 2 / 6 5	3 4 / 7 8			19	3 4 / 8 7	1 2 / 5 6		
4	1 2 / 6 5	3 4 / 8 7			20	3 4 / 8 7	1 2 / 6 5		
5	1 2 / 5 6	4 3 / 7 8			21	3 4 / 7 8	2 1 / 5 6		
6	1 2 / 5 6	4 3 / 8 7			22	3 4 / 7 8	2 1 / 6 5		
7	1 2 / 6 5	4 3 / 7 8			23	3 4 / 8 7	2 1 / 5 6		
8	1 2 / 6 5	4 3 / 8 7			24	3 4 / 8 7	2 1 / 6 5		
9	2 1 / 5 6	3 4 / 7 8			25	4 3 / 7 8	1 2 / 5 6		
10	2 1 / 5 6	3 4 / 8 7			26	4 3 / 7 8	1 2 / 6 5		
11	2 1 / 6 5	3 4 / 7 8			27	4 3 / 8 7	1 2 / 5 6		
12	2 1 / 6 5	3 4 / 8 7			28	4 3 / 8 7	1 2 / 6 5		
13	2 1 / 5 6	4 3 / 7 8			29	4 3 / 7 8	2 1 / 5 6		
14	2 1 / 5 6	4 3 / 8 7			30	4 3 / 7 8	2 1 / 6 5		
15	2 1 / 6 5	4 3 / 7 8			31	4 3 / 8 7	2 1 / 5 6		
16	2 1 / 6 5	4 3 / 8 7			32	4 3 / 8 7	2 1 / 6 5		

Fig. 10. Phase sequence.

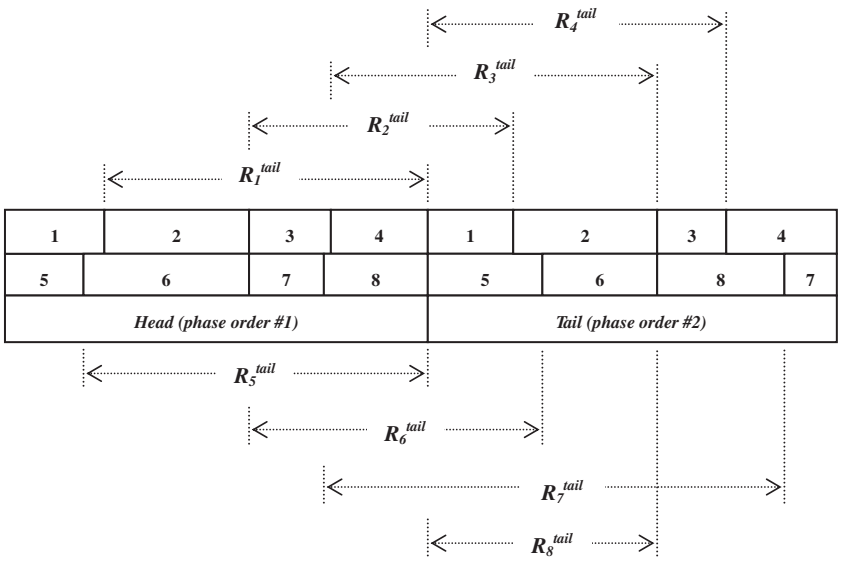

Fig. 11. Expression of red duration.

Then, Eq. (34) can be revised into

$$
\begin{aligned}
{}^{k}W_{total}^{tail} = & \left(\frac{1}{8}\sum_i \lambda_i^{tail}\right)\cdot\left(\sum_i {}^{k}G_i^{tail}\right)^2 - \sum_i\left[{}^{spill}Q_i^{head} + \lambda_i^{tail}\left({}^{k}R_i^{tail} + {}^{k}G_i^{tail}\right)\right]\cdot {}^{k}t_{i1} \\
& -\frac{1}{2}\sum_i\left\{\frac{\left[{}^{spill}Q_i^{head} + \lambda_i^{tail}\left({}^{k}R_i^{tail} + L_i\right)\right]^2}{S_i - \lambda_i^{tail}} + \lambda_i^{tail}\left({}^{k}G_i^{tail} - L_i\right)^2\right\}
\end{aligned} \tag{35}
$$

Since the objective is specified as minimization of total intersection control delay, the objective function can be expressed by

$$\min {}^{k}W_{total}^{tail}, \quad \text{where } k = 1,2,\ldots,32$$

Obviously, there is a specific value of minimized total delay that corresponds to the phase order k. Here, of the thirty-two possible values of minimized delay, we take the minimum as the optimal result, i.e.,

$$W_{optimal} = \min\left[\min {}^{k}W_{total}^{tail}\right], \text{ where } k = 1,2,\ldots,32$$

Additionally, three constraints are considered in formulating the optimization problem:

(1) Barrier condition. According to the concept of dual-ring signal control, the timing period in ring A should be equal to the timing period in ring B on either side of the barrier, therefore,

$$G_1^{tail} + G_2^{tail} = G_5^{tail} + G_6^{tail}$$
$$G_3^{tail} + G_4^{tail} = G_7^{tail} + G_8^{tail}$$

(2) Clearance condition. As mentioned previously, this constraint indicates that the phase green interval is expected to be large enough to service the initial queue, if any, plus all the vehicles that arrive during the effective red and effective green, and finally terminate with no vehicle spillover. Therefore,

$$G_i^{tail} - L_i \geqslant G_{qi}^{tail}$$

Substituting Eq. (7'') yields

$$G_i^{tail} - L_i \geqslant \left[{}^{spill}Q_i^{head} + \lambda_i^{tail}\times\left(R_i^{tail} + L_i\right)\right]/\left(S_i - \lambda_i^{tail}\right)$$

(3) Storage condition. It is assumed here that each link between signalized intersections and each left-turn bay has a finite storage capacity. The storage capacity can either be defined as maximum allowable queue length if vehicles are assumed to queue horizontally along the approach, or defined as maximum allowable queue size if vehicles are assumed to queue vertically at the stop line. In this research, the storage capacity is equivalent to maximum allowable queue size based on the assumption in Section 3.1, and it is expected that the maximum vehicle queue size of each phase, which forms at the start moment of effective green, cannot exceed the storage capacity, therefore,

$${}^{spill}Q_i^{head} + \lambda_i^{tail}\times\left(R_i^{tail} + L_i\right) < Q_{\max i}$$

Therefore, the complete optimization problem can now be expressed by

$$W_{optimal} = \min\left[\min {}^{k}W_{total}^{tail}\right], \quad \text{where } k = 1,2,\ldots,32 \tag{36}$$

subject to

$$
\begin{aligned}
{}^{k}W_{total}^{tail} = & \left(\frac{1}{8}\sum_i \lambda_i^{tail}\right)\cdot\left(\sum_i {}^{k}G_i^{tail}\right)^2 - \sum_i\left[{}^{spill}Q_i^{head} + \lambda_i^{tail}\left({}^{k}R_i^{tail} + {}^{k}G_i^{tail}\right)\right]\cdot {}^{k}t_{i1} \\
& -\frac{1}{2}\sum_i\left\{\frac{\left[{}^{spill}Q_i^{head} + \lambda_i^{tail}\left({}^{k}R_i^{tail} + L_i\right)\right]^2}{S_i - \lambda_i^{tail}} + \lambda_i^{tail}\left({}^{k}G_i^{tail} - L_i\right)^2\right\}, i = 1,2,\ldots,8
\end{aligned}
$$

$${}^{k}G_1^{tail} + {}^{k}G_2^{tail} = {}^{k}G_5^{tail} + {}^{k}G_6^{tail}$$
$${}^{k}G_3^{tail} + {}^{k}G_4^{tail} = {}^{k}G_7^{tail} + {}^{k}G_8^{tail}$$
$${}^{k}G_i^{tail} - L_i \geqslant \left[{}^{spill}Q_i^{head} + \lambda_i^{tail}\times\left({}^{k}R_i^{tail} + L_i\right)\right]/\left(S_i - \lambda_i^{tail}\right), \quad i = 1,2,\ldots,8$$
$${}^{spill}Q_i^{head} + \lambda_i^{tail}\times\left(R_i^{tail} + L_i\right) < Q_{\max i}, \quad i = 1,2,\ldots,8$$

The solution to this problem are the optimal phase sequence k, and a set of optimal phase splits ${}^{k}G_i^{tail}$ corresponding to the phase order k.

3.4.2. Determining optimal minimum green

Conventionally, the minimum green is set equal to an initial green time that allows all the vehicles potentially stored between the set-back detector (e.g., extension detector) and the stop line to enter the intersection (especially for through-movement phases). This setting assumes that the entire distance between the detector and the stop line is occupied by stored vehicles, an assumption that may be violated under light traffic conditions. Referring to Fig. 8a and Fig. 9a for example, the required queue service time is less than the minimum green, i.e., the queuing vehicles (stored vehicles plus the vehicles joining the queue) will enter the intersection before minimum green expires, and thus the minimum green will not be "fully" utilized and the phase may terminate by gap-out control later. This weakness is mitigated by volume–density control, in which the minimum green (or, added initial) is calculated based on the number of stored vehicles, and this computed initial green cannot exceed a pre-set maximum limit (i.e., maximum initial). A similar method is taken here to determine minimum green: set minimum green equal to queue service time if the queue service time is less than the pre-determined (i.e., conventional) minimum green, G_{mini}^{0}, otherwise, set it equal to the pre-determined minimum green, i.e.,

$$G_{\min}^{itail} = \begin{cases} G_{qi}^{tail} & \text{if } G_{qi}^{tail} \leqslant G_{\min}^{i0} \\ G_{\min}^{i0} & \text{if } G_{qi}^{tail} > G_{\min}^{i0} \end{cases}, \text{ or}$$

$$G_{\min}^{itail} = \min\left[G_{qi}^{tail}, G_{\min}^{i0}\right] \tag{37}$$

where

$$G_{qi}^{tail} = \left[{}^{spill}Q_i^{head} + \lambda_i^{tail} \times \left({}^{k}R_i^{tail} + L_i\right)\right] / \left(S_i - \lambda_i^{tail}\right)$$

3.4.3. Determining optimal unit extension

The optimal unit extension is expected to be a gap time that is large enough to invoke gap-out control at the end of the specified optimal phase green. Then, according to Eq. (10), the optimal phase split can be expressed by

$${}^{k}G_i^{tail} = G_{\min}^{itail} + \left[\exp\left(\lambda_i^{tail}\beta_i^{tail}\right) - 1\right] / \lambda_i^{tail} + L_i \tag{38}$$

Then, the optimal unit extension is determined by

$$\beta_i^{tail} = \ln\left[1 + \lambda_i^{tail} \times \left({}^{k}G_i^{tail} - L_i - G_{\min}^{itail}\right)\right] / \lambda_i^{tail} \tag{39}$$

Note that the natural logarithm in Eq. (39) requires that $\left[1 + \lambda_i^{tail} \times ({}^{k}G_i^{tail} - L_i - G_{mini}^{tail})\right]$ be greater than 0. According to the clearance condition, we have

$${}^{k}G_i^{tail} - L_i \geqslant G_{qi}^{tail}$$

Thus,

$${}^{k}G_i^{tail} - L_i - G_{\min}^{itail} \geqslant G_{qi}^{tail} - G_{\min}^{itail}$$

According to Eq. (37), we have

$$\begin{aligned} &{}^{k}G_i^{tail} - L_i - G_{\min}^{itail} \geqslant G_{qi}^{tail} - G_{qi}^{tail} \quad \text{if } G_{qi}^{tail} \leqslant G_{\min}^{i0} \\ &{}^{k}G_i^{tail} - L_i - G_{\min}^{itail} \geqslant G_{qi}^{tail} - G_{\min}^{i0} \quad \text{if } G_{qi}^{tail} > G_{\min}^{i0} \end{aligned}, \text{ or}$$

$$\begin{aligned} &{}^{k}G_i^{tail} - L_i - G_{\min}^{itail} \geqslant 0 \text{ if } G_{qi}^{tail} \leqslant G_{\min}^{i0} \\ &{}^{k}G_i^{tail} - L_i - G_{\min}^{itail} > 0 \text{ if } G_{qi}^{tail} > G_{\min}^{i0} \end{aligned}$$

Then,

$$\begin{aligned} &1 + \lambda_i^{tail} \times \left({}^{k}G_i^{tail} - L_i - G_{\min}^{itail}\right) \geqslant 1 \text{ if } G_{qi}^{tail} \leqslant G_{\min}^{i0} \\ &1 + \lambda_i^{tail} \times \left({}^{k}G_i^{tail} - L_i - G_{\min}^{itail}\right) > 1 \text{ if } G_{qi}^{tail} > G_{\min}^{i0} \end{aligned} \tag{40}$$

The requirement of natural logarithm is satisfied. Furthermore, substituting inequality (40) into Eq. (39), we have

$$\begin{aligned} &\beta_i^{tail} \geqslant 0 \quad \text{if } G_{qi}^{tail} \leqslant G_{\min}^{i0} \\ &\beta_i^{tail} > 0 \quad \text{if } G_{qi}^{tail} > G_{\min}^{i0} \end{aligned} \tag{41}$$

where

$$G_{qi}^{tail} = \left[{}^{spill}Q_i^{head} + \lambda_i^{tail} \times \left({}^{k}R_i^{tail} + L_i\right)\right] / \left(S_i - \lambda_i^{tail}\right)$$

3.4.4. Determining optimal maximum green

When the optimal unit extension fails to terminate the phase by gap-out control at the end of the optimized phase green interval, the maximum green limit is needed. Recalling Eq. (38) for the expression of the optimized phase split, we have

$$\begin{aligned} &{}^{k}G_i^{tail} - \boldsymbol{L}_i = G_{\min i}^{tail} + \left[\exp\left(\lambda_i^{tail}\beta_i^{tail}\right) - 1\right] \Big/ \lambda_i^{tail}, \quad \text{or} \\ &{}^{k}g_{ei}^{tail} = \boldsymbol{G}_{\min i}^{tail} + \left[\exp\left(\lambda_i^{tail}\beta_i^{tail}\right) - 1\right] \Big/ \lambda_i^{tail} \end{aligned} \tag{42}$$

As can be realized, given that G_{min}^{tail} is determined by Eq. (37) and β^{tail} determined by Eq. (39), the effective green over the tail portion, g_e^{tail}, is a monotone increasing function of λ^{tail}. Specifically, we assume here that g_e^{tail} reaches its maximum limit (i.e., maximum green time G_{max}^{tail}) as λ^{tail} increases to a value that is equal to the mean of λ^{tail} and the saturation flow rate S. Therefore,

$$\begin{aligned} &{}^{k}G_{\max}^{itail} = G_{\min}^{itail} + \left[\exp\left(\lambda_i'^{tail}\beta_i^{tail}\right) - 1\right] \Big/ \lambda_i'^{tail}, \text{ where } \lambda_i'^{tail} = \left(\lambda_i^{tail} + S_i\right) \Big/ 2, \text{ or} \\ &{}^{k}G_{\max}^{itail} = G_{\min}^{itail} + 2 \times \left[\exp\left(\beta_i^{tail} \times \frac{\lambda_i^{tail} + S_i}{2}\right) - 1\right] \Big/ \left(\lambda_i^{tail} + S_i\right) \end{aligned} \tag{43}$$

4. Testing and evaluation

The proposed adaptive control algorithm is tested and evaluated using the scalable, high-performance microscopic simulation package, PARAMICS (Cameron and Duncan, 1996). PARAMICS has been widely used in the testing and evaluation of various Intelligent Transportation System (ITS) strategies because of its powerful Application Programming Interfaces (API). Users can access the core functions provided in PARAMICS through API to customize and extend many features of the underlying simulation model without having to deal with the proprietary source codes. For the purpose of testing and evaluation in this paper, the proposed algorithm is developed as a plug-in through API programming and applied to a calibrated signalized network in PARAMICS. This network consists of 38 signals that are individually controlled by the free-mode traffic-actuated strategy and, for more detailed performance comparison, these signals are further coded under actuated-coordinated and volume–density control with properly tuned parameters (which are also programmed into PARAMICS plug-ins through API).

The study network is as shown in Fig. 12, which is so-called the "Irvine Triangle" located in southern California. This network includes a 6-mile section of freeway I-405, a 3-mile section of freeway I-5, a 3-mile section of freeway SR-133 and several adjacent surface streets, including two streets parallel to I-405 (i.e. Alton Parkway and Barranca Parkway), one street parallel to I-5 (i.e., Irvine Center Drive), and three crossing streets to I-405 (i.e. Culver Drive, Jeffery Road, and Sand Canyon Avenue). A total of thirty-eight signals under free-mode traffic-actuated control are included in the network. A previous study has calibrated this network for the morning peak period from 6 AM to 10 AM, and calibration results have properly reflected existing traffic conditions (Chu et al., 2004).

In the proposed control algorithm, the saturation flow rate, S, is assumed to be 1900 veh/h/lane for each through movement phase, and 1800 veh/h/lane for each left-turn movement phase. If the optimized min green is extremely short (e.g., <4 s), it is set to be 4 s in order to roughly cover the start-up lost time. If the optimized unit extension is not greater than $1/S$, which may cause "premature gap-out" right after the min green expires, it is set equal to $1/S + 0.1$ s. The unit extension needs to be further adjusted for those phases (especially through movement phases) that have relatively distant set-back extension detectors. Fig. 13 shows an example where the extension detectors for the through movement phase are placed beyond the left-turn bay. In this case, the unit extension for the through phase is adjusted by

$$\beta_i'^{tail} = \ln\left[1 + \lambda_i'^{tail} \times \left({}^{m}G_i^{tail} - L_i - G_{\min}^{itail}\right)\right] \Big/ \lambda_i'^{tail} \tag{44}$$

where $\lambda_i'^{tail}$ is the the approaching flow rate over the tail portion.

Three traffic demand scenarios are set up for the simulation:

1. Existing demand scenario: this scenario corresponds to the traffic condition for the morning peak period and the traffic demands are obtained directly from the calibrated simulation model;
2. Medium demand scenario: the traffic demands are equivalent to 75% of the existing demand scenario;
3. Low demand scenario: the traffic demands are equivalent to 50% of the existing demand scenario.

Simulations are conducted for a period of 4 h and 15 min for each scenario under each control mode. The first 15 min are considered to be the warm-up period for vehicles to fill in the network, in order to represent the typical (free-flow) traffic

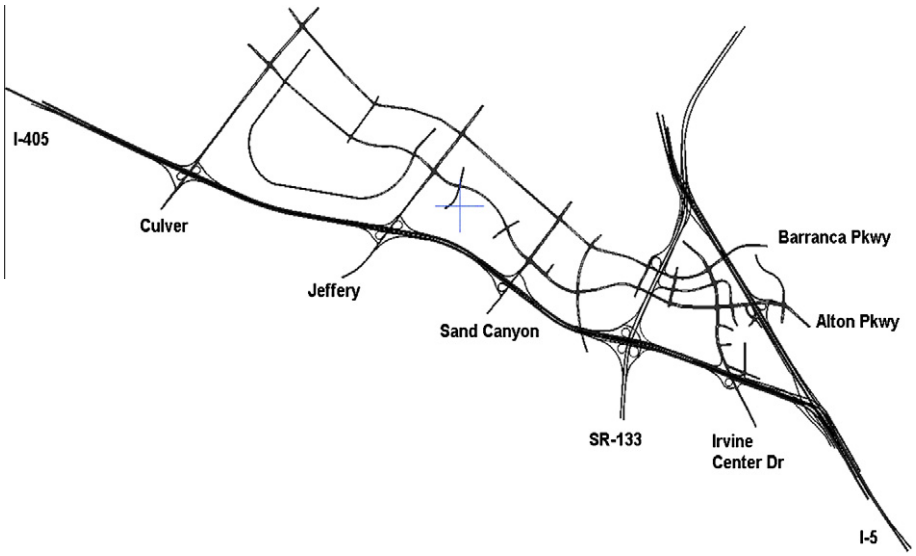

Fig. 12. Study network.

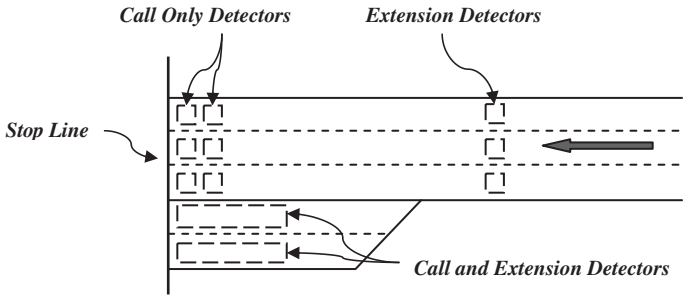

Fig. 13. Set-back extension detector configuration.

condition at the start time of simulation (i.e., 6 AM)—it takes less than 10 min for a vehicle to finish the longest trip in the network in the real world.

It should be noted here that the main objective of simulation is to evaluate if the proposed algorithm has the ability to improve the performance of the network under existing traffic conditions; therefore, the setup of only scenario 1 pertains directly to this end. However, scenario 2 and 3 are set up to further evaluate if this algorithm also has the ability to improve the performance under the conditions of lower demand levels, which may reflect non-congested or off-peak situations. And, since scenario 1 already reflects a relatively high demand level, the condition of an even higher demand level is not considered.

Two groups of measure of performance are specified as the major simulation results for analysis: (1) For the whole network: Average Travel Time and Average Vehicle Speed; and (2) for local intersections: Maximum Queue Length and Vehicle Travel Delay. Since this research focuses on the signal control at intersections, the measure of performance regarding the freeway are not included in the results.

These measures of performance are compared between these four control modes for each scenario as defined above. Since PARAMICS is a stochastic simulation model that introduces random effects in various processes during simulation, the results of several simulation runs using different seed numbers are needed to reflect the general traffic condition for a specific scenario. The method used to determine the number of required runs is explained as follows (Chu et al., 2004). First, five simulation runs are conducted; then, Eq. (45), which specifies the required number of runs, is used to determine if five runs satisfy the criterion. If not, one additional run is conducted and then the required number is calculated again. This process continues until the criterion is satisfied, at which point the results for the current number of runs are averaged for performance analysis.

$$N = \left(t_{\alpha/2} \times \frac{\sigma}{\mu \times \varepsilon} \right)^2 \tag{45}$$

where N is the required number of simulation runs; μ, σ is the mean and standard deviation of a measure of performance based on conducted simulation runs; ε is the allowable error specified as a fraction of the mean μ; and $t_{\alpha/2}$ is the critical value of the t-distribution at the significance level α.

In this study, $\varepsilon = 2\%$ and $\alpha = 0.05$.

The simulation results of the network are shown in Table 1. The corresponding improvements compared with the free-mode actuated control, which are represented as positive percentages, are also included. It is found that the network under either of the actuated-coordinated, volume–density and proposed control performs better than the free-mode actuated control in all scenarios—drivers spend less time in the network with improved traveling speed. Specifically, these better control modes have gained the most improvement in scenario 1 (i.e., 6.3%, 6.9%, 8.7% decrease in Average Travel Time, and 6.5%, 7.1%, 9.0% increase in Average Vehicle Speed), while the least improvement in scenario 3 (i.e., 0.2%, 0.6%, 0.7% decrease in Average Travel Time, 0.2%, 0.7%, 0.7% increase in Average Vehicle Speed). One possible reason underlying this result is that the low flow-level traffic may behave freely in the network without being affected by the change of control strategy. Furthermore, the network under the proposed algorithm has gained the best performance, which indicates that the adaptive functionality incorporated in this algorithm may provide more efficiency than the other actuated control strategies.

A T-intersection is selected to demonstrate the performance at the intersection level. This intersection corresponds to the junction of Irvine Center Drive and the Off Ramp from Southbound I-405, as shown in Fig. 14. Phases 2 and 6 are assigned to the through movements and operate as min-recall phases, and phase 4 is assigned to the left-turn movement (plus the right-turn movement) without recall. The extension detectors ($6' \times 6'$) for the through phases are placed 300 ft upstream from the stop line, and the call and extension detectors ($5' \times 50'$) for the left-turn phase are placed right behind the stop line. The operation parameters of this signal in the free-mode actuated, actuated-coordinated and volume–density control are shown respectively in Table 2.

For simplicity, only the simulation results from scenario 1 are taken for analysis, as shown in Table 3. The corresponding improvements, which are represented as positive percentages, are also included. Similarly, the network performs better under the actuated-coordinated, volume–density and proposed control, with the improvement being 11.4%, 16.8%, 17.3% decrease in Maximum Queue Length, and 10.7%, 16.0%, 16.2% decrease in Vehicle Travel Delay. Again, the network under the proposed control algorithm has achieved the best performance.

Table 1
Performance of the network.

	Average travel time (s)	Improvement (%)	Average vehicle speed (mile/h)	Improvement (%)
Scenario 1				
Free-mode actuated	297.0	n/a	36.6	n/a
Actuated-coordinated	278.3	6.3	39.0	6.5
Volume–density	276.5	6.9	39.2	7.1
Proposed adaptive	271.2	8.7	39.9	9.0
Scenario 2				
Free-mode actuated	213.3	n/a	43.8	n/a
Actuated-coordinated	212.7	0.3	44.0	0.4
Volume–density	210.7	1.2	44.4	1.4
Proposed adaptive	209.9	1.6	44.6	1.8
Scenario 3				
Free-mode actuated	197.1	n/a	45.9	n/a
Actuated-coordinated	196.7	0.2	46.0	0.2
Volume–density	195.9	0.6	46.2	0.7
Proposed adaptive	195.8	0.7	46.2	0.7

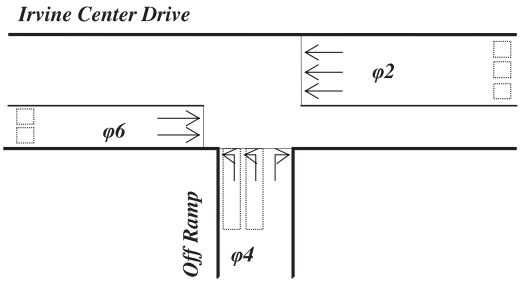

Fig. 14. Study intersection.

Table 2
Parameters for the study intersection.

	Phase	2	4	6
Free-mode actuated	Minimum green (s)	8	5	8
	Maximum green (s)	40	24	40
	Unit extension (s)	5	2	5
Actuated-coordinated	Coordinated	yes	no	yes
	Phase split (s)	40	20	40
	Coordination mode	Yield mode with cycle length = 60 s		
	Unit extension (s)	n/a	2	n/a
Volume–density	Added initial/actuation (s)	1	n/a	1
	Maximum initial (s)	20	n/a	20
	Minimum gap (s)	2	n/a	2
	Time before reduction (s)	15	n/a	15
	Time to reduce (s)	10	n/a	10

Note: the min initial, max green and unit extension in volume–density control are set equal to the min green, max green and unit extension in free-mode actuated control; while in actuated-coordinated control, only the min green of the non-coordinated phase is set equal to that in free-mode actuated control, and the max green of the non-coordinated phase, as well as the "min" and "max" green of the coordinated phases (which may refer to the yield and force-off points), is determined based on the setting of cycle length and phase splits.

Table 3
Simulation results of the intersection (Scenario 1).

	Maximum queue length (feet)	Improvement (%)	Vehicle travel delay (s)	Improvement (%)
Free-mode actuated	158.4	n/a	1440.5	n/a
Actuated-coordinated	140.3	11.4	1286.4	10.7
Volume–density	131.8	16.8	1210.0	16.0
Proposed adaptive	131.0	17.3	1206.6	16.2

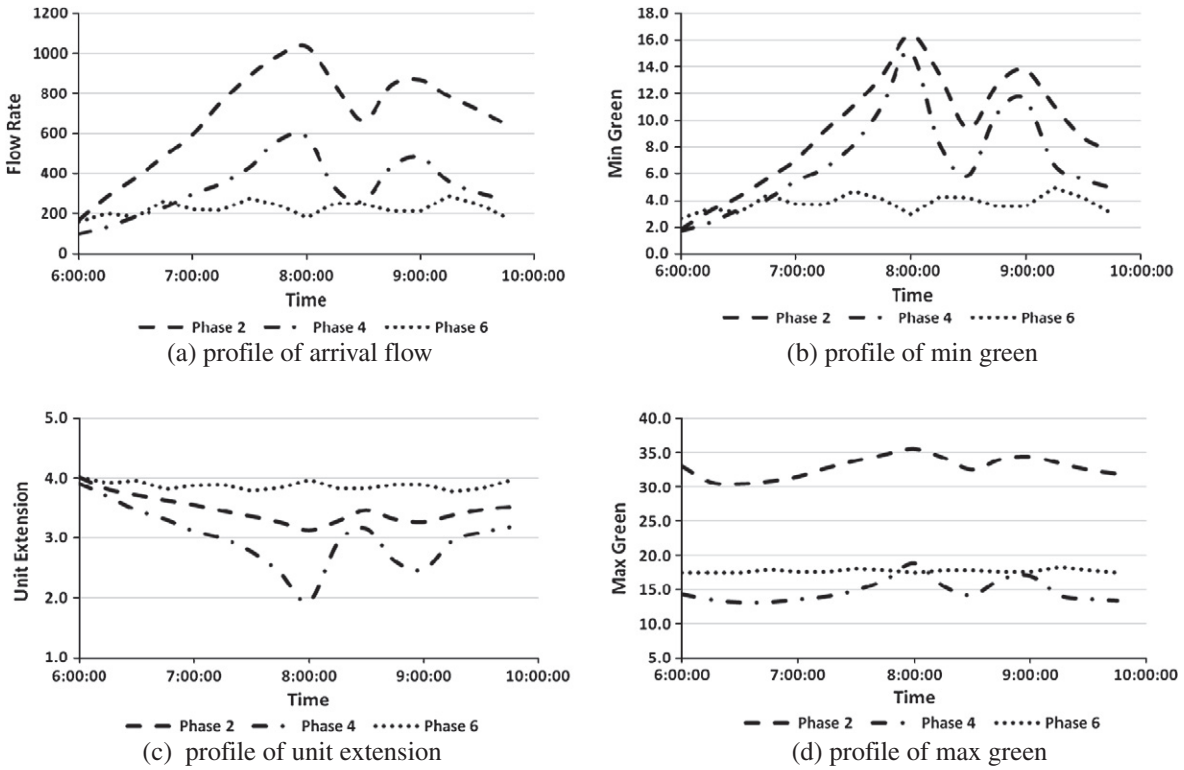

Fig. 15. Profiles of flow and parameters.

With the same example, the variation of the control parameters in the proposed algorithm is shown here to illustrate how these parameters are dynamically determined as the traffic condition changes. Fig. 15a shows the arrival flow profiles (identical under all control modes) for each signal phase over the analysis period. These profiles are plotted with smoothed lines based on the average flow data measured on 15-min intervals. As can be seen, phase 2 and phase 4 experience two "peak" periods around 8 AM and 9 AM, and phase 6 experiences relatively steady and low level of flow through the whole simulation period. Since there are actually only two signal phases that operate at this intersection (as phase 2 and phase 6 operate simultaneously), the phase sequence stays as "2/6–4–2/6–4...". Therefore, only the variations of min green, unit extension and max green are considered here.

Fig. 15b–d show the corresponding profiles of these three parameters that are real time determined. Again, these profiles are plotted with smoothed lines based on the mean value over 15-min intervals. As can be realized, the min green of phase 2 and phase 4 reaches two "peak" values around 8 AM and 9 AM respectively (i.e., 16.4 s and 13.8 s for phase 2; 15.2 s and 11.6 s for phase 4), while the value of the min green for phase 6 is relatively steady and low through the whole period (i.e., around 4.0 s). This pattern is consistent with the flow profile as shown in Fig. 15a, which infers that higher levels of traffic flow need more min green for vehicle queues to dissipate, and less min green is needed for relatively low levels of flow. The unit extension, on the contrary, takes two "trough" values around 8 AM and 9 AM for both phase 2 and phase 4 (i.e., 3.1 s and 3.3 s for phase 2; 1.9 s and 2.5 s for phase 4), while it is relatively steady and high for phase 6 (i.e., around 3.9 s). This pattern is reasonable because when the traffic flow increases, a smaller value is needed for the unit extension to terminate the phase by "gap-out" control, and thus no vehicle spillover may occur to cause additional intersection control delay. Last, the pattern of max green is also consistent with the flow profile (i.e., with two "peak" values being 35.5 s and 34.4 s for phase 2, 18.8 s and 17.0 s for phase 4, and a relatively steady value being around 17.7 s for phase 6), which means that a bigger value of max green is needed for higher levels of traffic flow, since the probability is increased that more vehicles enter the intersection.

The simulation results have shown that the proposed control algorithm has the potential to improve the performance of the signalized network. By applying this algorithm to different traffic flow scenarios, the improvement is found more promising for the situations of higher levels of traffic flow, even though the vehicle arrival pattern is assumed to be Poisson process in the algorithm formulation—an assumption that may be violated in urban areas where intersections are closely spaced and/or high-level traffic flows are loaded. Thus, better results would be expected were the real traffic flow pattern in urban networks to be modeled more accurately. Furthermore, better results would also be expected if the special operations in actuated controller, such as recall and dual-entry, had been considered in formulating the proposed control algorithm.

5. Conclusions and future work

This paper introduces an on-line control algorithm that aims to maintain the adaptive functionality of actuated controllers while improving the performance of traffic-actuated signal control system. In this algorithm, the real-time traffic condition of the target intersection, with all signal phases considered simultaneously, is viewed to be "over-saturated" (in the sense of a multi-server queuing system being continually occupied) throughout the timing process, and then the total intersection control delay, which serves as the primary performance indicator of level of service at signalized intersections, is calculated as the cumulative vehicle queue over a specific period. A modified rolling horizon scheme is devised to optimize those four basic control parameters that can be found in modern actuated controllers with the objective of minimizing total intersection control delay based on predicted future vehicle arrival flow. Microscopic simulation is employed to test and evaluate the proposed control algorithm in a calibrated network consisting of thirty-eight actuated signals. Simulation results indicate that the proposed control algorithm has the potential to improve the performance of the signalized network at different traffic demand levels.

A major advantage of the proposed control algorithm is that it is formulated based on the available traffic-actuated control protocol and signal timing features without any additional hardware investment, and its output is restricted to only those four basic control parameters that are employed in modern actuated controllers. Therefore, this algorithm can be implemented with minimal adaptation of existing field devices and the software that controls their operation. Unlike some other well-known real-time adaptive control systems that specify the optimal, fixed split or green duration for each signal phase, this algorithm determines those control parameters (especially the unit extension and max green) that have the ability to adjust phase split in response to real-time traffic conditions. Furthermore, this algorithm is programmed as a plug-in file in microscopic simulation, which enables the testing and evaluation of this algorithm in an existing, well-calibrated network.

A few efforts will be made in the future to expand and improve the proposed control algorithm, such as (1) incorporating special operation parameters and functions, (2) seeking a more sophisticated flow model that represent the vehicle arrival pattern in a signalized network, and (3) comparing the proposed control algorithm with other real-time traffic-responsive and traffic-adaptive control systems.

Appendix A. Notations

i	subscript, the index of actuated phase ($i = 1, 2, \ldots, 8$)
j	superscript, the index of floating cycle ($j, k = 1, 2, 3, \ldots$)
λ	average arrival flow rate (veh/s)
β	unit extension, i.e., passage time, gap or vehicle interval (s)
g_e	effective green time (s)
l_1	start-up lost time (s)
l_2	clearance lost time (s)
r_e	effective red time (s)
$A(t)$	vehicle arrival function of time t
$D(t)$	vehicle departure function of time t
G	phase split, equal to green interval plus change and clearance interval (s)
G_e	green extension period (s)
G_{max}	maximum green (s)
G_{min}	minimum green (s)
G_q	queue service time (s)
L	total lost time, equal to $(l_1 + l_2)$ (s)
Q	initial vehicle queue, i.e., vehicle spillover from the previous cycle (veh)
Q_{max}	storage capacity, i.e., maximum allowable queue size (veh)
$_{spill}$ Q	vehicle spillover, i.e., initial vehicle queue in the upcoming cycle (veh)
R	red duration time (s)
S	saturation flow rate (veh/s)
W_{total}	total intersection control delay (veh s)

References

Akcelik, R., 1981. Traffic Signals: Capacity and Timing Analysis. Research Report 123, Australian Road Research Board, Melbourne, Australia.

Akcelik, R., 1988. The highway capacity manual delay formula for signalized intersections. ITE Journal 58 (3), 23–27.

Akcelik, R., Rouphail, N.M., 1993. Estimation of delays at traffic signals for variable demand conditions. Transportation Research Part B 27 (2), 109–131.

Bang, K-L., 1976. Optimal control of isolated traffic signals. Traffic Engineering and Control 17 (7), 288–292.

Bell, M.G.H., 1990. A probabilistic approach to the optimization of traffic signal settings in discrete time. In: Proceedings of the 11th International Symposium on Transportation and Traffic Theory, Tokyo, pp. 619–631.

Brilon, W., Wu, N., 1990. Delays at fixed-time traffic signals under time-dependent traffic conditions. Traffic Engineering and Control 31 (12), 623–631.

Cameron, G.D.B., Duncan, G.I.D., 1996. Paramics – parallel microscopic simulation of road traffic. The Journal of Supercomputing 10 (1), 25–53.

Chou, Y., 1970. Statistical Analysis. Holt, Rinehart and Winston, New York.

Chu, L., Liu, X., Recker, W., 2004. Using microscopic simulation to evaluate potential intelligent transportation system strategies under nonrecurrent congestion. Transportation Research Record 1886, 76–84.

Cox, D.R., Smith, W.L., 1961. Queues. Methuen, London.

Dion, F., Rakha, H., Kang, Y., 2004. Comparison of delay estimates at under-saturated and over-saturated pre-timed signalized intersections. Transportation Research Part B 38 (2), 99–122.

Fambro, D.B., Rouphail, N.M., 1997. Generalized delay model for signalized intersections and arterial streets. Transportation Research Record 1572, 112–121.

Gartner, N.H., 1983. OPAC: a demand-responsive strategy for traffic signal control. Transportation Research Record 906, 75–81.

Gazis, D.C., Potts, R.B., 1963. The oversaturated intersection. In: Proceedings of the Second International Symposium on Traffic Theory, London, UK, pp. 221–237.

Head, K.L., Mirchandani, P.B., Sheppard, D.E., 1992. A hierarchical framework for real-time traffic control. Transportation Research Record 1360, 82–88.

Heidemann, D., 1994. Queue length and delay distributions at traffic signals. Transportation Research Part B 28 (5), 377–389.

Heydecker, B.G., 1990. A continuous time formulation for traffic-responsive signal control. In: Proceedings of the 11th International Symposium on Transportation and Traffic Theory, Tokyo, pp. 599–618.

Hunt, P.B., Robertson, D.I., Bretherton, R.D., Winton, R.I., 1981. SCOOT – A Traffic Responsive Method of Coordinating Signals. LR1014, Transport and Road Research Laboratory, Crowthorne, UK.

ITE, 2001. Traffic Control Devices Handbook. Institute of Transportation Engineers, Washington, DC.

Lin, F.B., Vijayakumar, S., 1988. Adaptive signal control at isolated intersections. Journal of Transportation Engineering 114 (5), 555–573.

Lin, F.B., Cooke, D.J., Vijayakumar, S., 1987. Use of predicted vehicle arrival information for adaptive signal control – an assessment. Transportation Research Record 1112, 89–98.

Lo, H.K., 2001. A cell-based traffic control formulation: strategies and benefits of dynamic timing plans. Transportation Science 35 (2), 148–164.

Lowrie, P.R., 1992. SCATS: A Traffic Responsive Method of Controlling Urban Traffic Control. Technical Report, Roads and Traffic Authority, NSW, Australia.

McNeil, D.R., 1968. A solution to the fixed-cycle traffic light problem for compound Poisson arrivals. Journal of Applied Probability 5, 624–635.

Michalopoulos, P.G., Stephanopoulos, G., Pisharody, V.B., 1980. Modeling of traffic flow at signalized links. Transportation Science 14 (1), 9–41.

Mirchandani, P.B., Head, K.L., 2001. A real-time traffic control system: architecture, algorithms, and analysis. Transportation Research Part C 9 (6), 415–432.

Miller, A.J., 1963. A computer control system for traffic networks. In: Proceedings of the 2nd International Symposium on the Theory of Traffic Flow, London, pp. 200–220.

Newell, G.F., 1960. Queues for a fixed-cycle traffic light. Annals of Mathematical Statistics 31 (3), 589–597.

Newell, G.F., 1965. Approximation methods for queues with application to the fixed-cycle traffic light. SIAM Review 7 (2), 223–240.

Pacey, G.M., 1956. The progress of a bunch of vehicles released from a traffic signal. RRL Note, No. RN/2665/GMP, Transport and Road Research Laboratory, Growthorne, UK.

Pavlis, Y., Recker, W.W., 2004. Inconsistencies in the problem of optimal signal control for surface street networks. In: Proceedings of the 7th International IEEE Conference on Intelligent Transportation Systems, pp. 361–366.

Robertson, D.I., 1969. TRANSYT: A Traffic Network Study Tool. RRL Report, No. LR253, Transport and Road Research Laboratory, Growthorne, UK.

Robertson, D.I., Bretherton, R.D., 1974. Optimal control of an intersection for any known sequences of vehicle arrivals. In: Proceedings of the 2nd IFAC-IFIP-IFORS Symposium of Traffic Control and Transport Systems, Monte Carlo.
Robertson, D.I., Bretherton, R.D., 1991. Optimizing networks of traffic signals in real time-the scoot method. IEEE Transactions on Vehicular Technology 40 (1), 11–15.
Rorbech, J., 1968. Determining the length of the approach lanes required at signal-controlled intersections on through highways. Transportation Research 2, 283–291.
Sims, A.G., Dobinson, K.W., 1979. SCATS – Sydney coordinated adaptive traffic system philosophy and benefits. International Symposium on Traffic Control Systems, 2B.
Stephanedes, Y.J., Chang, K.-K., 1993. Optimal control of freeway corridors. Journal of Transportation Engineering 119 (4), 504–514.
Stephanopoulos, G., Michalopoulos, P.G., 1979. Modeling and analysis of traffic queue dynamics at signalized intersections. Transportation Research Part A 13 (5), 295–307.
TRB, 2000. Highway Capacity Manual. Special Report 209. National Research Council, Washington, DC.
Vincent, R.A., Young, C.P., 1986. Self-optimizing traffic signal control using microprocessors – the TRRL "MOVA" strategy for isolated intersections. Traffic Engineering Control 27 (7/8), 385–387.
Webster, F.V., 1958. Traffic signal settings. Road Research Technical Paper No. 39, Road Research Laboratory, Her Majesty Stationary Office, London, UK.
Zheng, X., Recker, W., Chu, L., 2010. Optimization of control parameters for adaptive traffic-actuated signal control. Journal of Intelligent Transportation Systems 14 (2), 95–108.