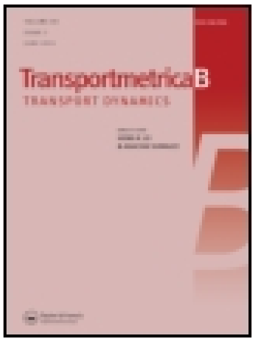

Transportmetrica B: Transport Dynamics

ISSN: 2168-0566 (Print) 2168-0582 (Online) Journal homepage: http://www.tandfonline.com/loi/ttrb20

A link partitioning approach for real-time control of queue spillbacks on congested arterials

Mohsen Ramezani, Nicolas de Lamberterie, Alexander Skabardonis & Nikolas Geroliminis

To cite this article: Mohsen Ramezani, Nicolas de Lamberterie, Alexander Skabardonis & Nikolas Geroliminis (2016): A link partitioning approach for real-time control of queue spillbacks on congested arterials, Transportmetrica B: Transport Dynamics

To link to this article: http://dx.doi.org/10.1080/21680566.2016.1142399

Published online: 16 Feb 2016.

TRANSPORTMETRICA B: TRANSPORT DYNAMICS, 2016
http://dx.doi.org/10.1080/21680566.2016.1142399

Taylor & Francis
Taylor & Francis Group

A link partitioning approach for real-time control of queue spillbacks on congested arterials

Mohsen Ramezani[a,b], Nicolas de Lamberterie[a], Alexander Skabardonis[c] and Nikolas Geroliminis[a]

[a]École Polytechnique Fédérale de Lausanne (EPFL), School of Architecture, Civil and Environmental Engineering (ENAC), Urban Transport Systems Laboratory (LUTS), Lausanne, Switzerland; [b]Department of Civil Engineering, Institute of Transport Studies, Monash University, Melbourne, Australia; [c]Department of Civil and Environmental Engineering, Institute of Transportation Studies, University of California, Berkeley, CA, USA

ABSTRACT
In oversaturated urban traffic conditions when traffic demand exceeds capacity at signalised intersections, queues fail to clear during the allocated green times. Once a queue reaches the upstream intersection in an arterial, a queue spillback occurs that reduces the upstream link capacity. To mitigate the negative impacts of spillbacks, this article introduces a real-time adaptive traffic signal control method for global management of spillbacks along signalised arterials. The key idea of the proposed method is to implement a real-time partitioning of the arterial to detect critical cluster(s) of consecutive links with oversaturated traffic conditions. The partitioning approach enables to develop locally smaller-sized decentralised signal control strategies operating on the most upstream and downstream intersections of each cluster. Micro-simulation investigations on a real-world arterial site demonstrate the benefits of the proposed approach compared to an existing pre-timed signal control strategy and a classical decentralised green extension strategy. Utilising an advanced queue length detection method and specific focus on queue spillbacks prevention, the control strategy leads to significant reduction of congestion, and arterial total delay.

ARTICLE HISTORY
Received 29 August 2014
Accepted 11 January 2016

KEYWORDS
Feedback signal control; spillover; partitioning; real-time urban arterials management

1. Introduction

In the last decades, drastically increased traffic demand in cities has resulted in severe congestion problems. For sustainable development, constructing new infrastructures is less effective nowadays because of limited availability of urban space, financial matters, and environmental issues. In addition, as traffic is not equally distributed over the day with alternation of low and peak demand periods, such investments would consequently lead to over-capacity situations during most of the day.

Alternatively, the development of intelligent transport system technologies with new monitoring paradigms and computational tools enables us to capture traffic states in real-time and to implement traffic-responsive signal control schemes instead of fixed ones, which are particularly not adapted to oversaturated conditions. Although these new technologies can ameliorate traffic conditions in local small-scale congested areas, an optimal traffic-responsive signal control in scale of an arterial or a network is computationally intractable (see, e.g. Quinn 1992; Abu-Lebdeh and

CONTACT Mohsen Ramezani mohsen.ramezani@monash.edu

Benekohal 1997; Papageorgiou et al. 2003; Asthana et al. 2012) for heuristic signal control of multiple intersections.

In urban networks with signalised intersections, queues fail to fully clear during the allocated green times when demand exceeds the link capacity. This creates growing queues that eventually spill over to the immediate upstream link and impede the arrival flows such that vehicles cannot discharge at the capacity rate, though the upstream signal is green. Spillbacks may also occur when turning vehicles fill up the available storage length of turn bays and block the through movement. This results in a significant reduction of the intersection capacity both for the through traffic movement and the cross-street traffic entering the arterial that leads to excessive delays. Moreover, urban topologies with numerous and closely spaced signalised intersections contribute to the extension and possibility of spillbacks and gridlock conditions.

This paper tackles the traffic signal control problem in congested urban arterials by integrating a partitioning approach into a control scheme that reduces the risk of spillback occurrence through a feedback strategy. First, the arterial partitioning method detects in real-time the links with long queues, clusters them together if they are consecutive, and then identifies the entrance and exit intersections of each cluster. These intersections are regarded as critical junctions (considering the analogy of an active bottleneck in a freeway) that the proper adjustment of the associated signal timing settings could improve traffic conditions in the whole arterial. The partitioning method yields a small number of time-varying clusters of congested links that enables to implement locally smaller-sized (compared to the number of intersections in the arterial) decentralised signal control strategies while ensuring at the same time the global coherence of the control strategy along the arterial. In this manner, we seek to improve the arterial traffic condition at a very low cost by acting merely on critical intersections, as opposed to an arterial-wide optimisation process, that would require measurements of turning movement rates. The benefits of the arterial partitioning approach are evaluated through comparisons with the decentralised green extension strategy without partitioning and a state-of-the-practice optimal coordinated strategy using the Synchro optimisation software (TafficWare 2004).

The proposed control approach has the following main features: (i) it is proactive, real-time, and traffic-responsive, (ii) it is equitable between cross-streets and the arterial, and (iii) it does not require information about turning movements, which is difficult to be estimated in real-time. However, it is assumed that queue length for each link is available at the end of every cycle (see, e.g. Ramezani and Geroliminis 2015). The proposed adaptive signal control strategy aims at preventing or delaying spillback occurrence. To this end, it triggers once the queue cannot be completely cleared within the fixed (initial) signal timing, that is, early detection, and considers the remaining available space in the arterial link beyond the back of the queue as the measure of the queue spillback risk. The early detection of oversaturated conditions results in a proactive control strategy. The main contribution of the article is to introduce a feedback signal control strategy based on the arterial partitioning method. The proposed partitioning considers both the temporal and spatial extents of congestion in the arterial. With partitioning, the arterial signal control problem breaks down to several smaller-sized local signal control problems that can be handled more efficiently. A partitioning approach is developed in Ramezani and Geroliminis (2012) to identify traffic states for travel time distribution estimation. A field implementation of the proposed method seems readily applicable to various arterial sites to detect and control congested arterials with recent improvements in real-time detection of queue length with probe vehicles and loop detectors.

Identification of long queues and queue spillbacks has been addressed based on flow and occupancy measurements by fixed-location detectors (Skabardonis and Geroliminis 2008; Wu, Liu, and Gettman 2010; Geroliminis and Skabardonis 2011) and probe data (Ramezani and Geroliminis 2015). It has been found in Skabardonis and Geroliminis (2008) that for short links, the delay can increase by 50–100% under the occurrence of queue spillovers. Recently, probe data have been used to estimate queue length in real-time (Comert and Cetin 2009; Cheng et al. 2012; Christofa, Argote, and Skabardonis 2013; Ramezani and Geroliminis 2015) enabling to detect spillbacks. For instance, a variable threshold based on probe vehicles penetration rates and traffic demand has been identified to detect potential spillbacks with decent accuracy (Christofa, Argote, and Skabardonis 2013). Also, a method has been developed to reconstruct queue shockwave profiles and probabilistically infer the occurrence of spillbacks based on position and velocity of probe vehicles (Ramezani and Geroliminis 2015). Other methods for queue length estimation based on loop detectors or probe vehicles can be found in Liu et al. (2009), Ban, Hao, and Sun (2011), and others.

Several signal control strategies have been proposed to cope with oversaturated traffic conditions. The seminal work (Gazis 1964) deals only with two closely spaced intersections. Afterwards, some models introduce queue length constraints in the optimisation process in order to limit the occurrence of spillbacks (e.g. Michalopoulos and Stephanopoulos 1978; Haddad et al. 2010; Haddad and Mahalel 2014). Other strategies have suggested the implementation of metering approaches (Gordon 1969; Rathi 1988) or queue location management methods (Gal-Tzur, Mahalel, and Prashker 1993) in order to prevent or delay queue spillbacks occurrence particularly upstream of critical intersections. Reverse offsets are also used to account for downstream congested links and to prevent from inefficient green phase for the arterial through movement (Pignataro, McShane, and Crowley 1978; May, Montgomery, and Quinn 1988; Rathi 1988). Methods based on the optimisation procedure have been introduced to adjust splits (Diakaki, Papageorgiou, and Aboudolas 2002) and offsets and cycle lengths (Abu-Lebdeh and Benekohal 1997) for network control by considering local interactions. For instance, a traffic-responsive urban control model has been developed based on store-and-forward modelling of the urban network traffic in Diakaki, Papageorgiou, and Aboudolas (2002), in which the optimal green phase durations are dynamically determined with an objective to equalise queues in all links. Further, recent advances in the aggregated network traffic modelling have offered large-scale urban network signal control strategies (see, e.g. Keyvan-Ekbatani et al. 2012; Aboudolas and Geroliminis 2013; Geroliminis, Haddad, and Ramezani 2013; Keyvan-Ekbatani et al. 2015). Other state-of-the-art traffic signal strategies can be found in Abbas, Bullock, and Head (2001), Head and Mirchandani (2001), Shelby et al. (2008), and Varaiya (2013).

The remainder of paper is organised as follows. Section 2 describes the partitioning methodology to identify clusters of successive congested links in the arterial. Section 3 introduces the signal control strategy based on the arterial partitioning approach. Investigations of the signal control strategy on a real-world arterial site with the quantitative analyses are presented in Section 4. Finally, Section 5 concludes the findings and gives suggestions for future work.

2. Arterial partitioning methodology

The partitioning method comprises detection of arterial links with long queues, clustering the links with long queues if they are adjacent, and consequently identifying the entrance and exit intersections of each cluster. Afterwards by controlling signal settings only at those intersections, we aim at improving the traffic situation of the whole arterial. The arterial partitioning method first differentiates congested links from the uncongested ones based on the link queue length at the end of the red phase. A cluster of congested links consists of consecutive oversaturated or close-to-spillback links in the arterial delimited with uncongested links. It is assumed that queue lengths can be estimated with some error from loop detectors and/or probe data. Section 5 investigates the effect of measurement errors on the performance of the overall control strategy.

We choose a binary representation of the link traffic state, congested and uncongested states. Two strategies are considered to characterise the traffic state of link i at cycle k: early and late detection strategies. *Early detection* is a preventive and proactive strategy such that the link traffic state is considered as congested when queue length, q_k^i, is still moderate but exceeds the maximum number of vehicles per lane, q_{cr}^i, that can be cleared during the green phase of the arterial through movement with the initial fixed-time control, g_{init}^i. Let c denotes the saturation flow (capacity) per lane on the arterial [veh/s/lane], then $q_{\text{cr}}^i = g_{\text{init}}^i \cdot c$. *Late detection* is a strategy such that the link traffic state is considered congested when queue spillback occurrence becomes imminent, that is, $s_k^i < s_{\text{cr}}$, Algorithm 1 Arterial partitioning pseudo code

1. *If* cycle k of intersection i has been completed
    2. *Loop* $j = 1,2 \ldots, i_{\max}$ (link 1 is the most upstream in the arterial, $i_{\max}$ is the most downstream link)
        3. *If* $q_k^j > q_{\mathrm{cr}}^j$
            If ($j = 1$) *or* ($q_k^{j-1} \leq q_{\mathrm{cr}}^{j-1}$)
                Append intersection $j-1$ to the set of entrance intersections $\mathcal{L}$
            End If
            If ($j = i_{\max}$) or ($q_k^{j+1} \leq q_{\mathrm{cr}}^{j+1}$)
                Append intersection j to the set of exit intersections $\mathcal{F}$
            End If
        End If
    End Loop
4. *If needed modify the signal timing of intersection i based on Equations* (1)–(2)

where s_k^i denotes the empty space [veh/lane] available beyond the back of the queue in link i at cycle k and s_{cr} is a predefined threshold designating the critical number of empty spaces available beyond the back of the queue below which the risk of queue spillover occurrence is considered significant. Note that this parameter is not link-dependent. Figure 1(a) illustrates the arterial partitioning based on early detection of queues. It is worth pointing out that in Figure 1(a), two clusters of congested links are identified with the early detection strategy, whereas only one congested link is detected with the late detection strategy. While sensitivity analysis of s_{cr} has been performed in the micro-simulation, a more careful consideration is needed before such a strategy is applied in the field.

The early detection strategy avoids the increase of delay time due to the queue build-up. Moreover, it pre-empts the increase of congestion that could lead to significant risk of queue spillback, whereas the late detection strategy is only triggered once spillbacks are imminent. Hence, we adopt early detection strategy as it seems more promising and have been found more efficient through extensive simulations (30% more reduction of delay times). Further, the arterial partitioning updates at the end of every cycle of each intersection because (i) the control strategy is traffic-responsive, that is, traffic states (queue length) are monitored frequently and are fed back to the controller, and (ii) signal settings are changed at the end of each cycle. Accordingly, the update procedure of the control strategy is as follows. Once the cycle of intersection i is finished, a new queue length measurement of link i (i.e. the upstream link of intersection i) is available, and based on the previous estimations of the other links' queue lengths, the control strategy updates the arterial partitions. Then, if needed, the control strategy modifies the signal timing of intersection i (see Section 3). Note that there is no restriction on cycle length of signal timings. The pseudo code for arterial partitioning approach is presented in Algorithm 1.

Given a cluster of congested links, the entrance intersection designates the signalised intersection located at the upstream of the cluster. The intersection is denoted by $l \in \mathcal{L}$ and the link just downstream is denoted by $l+1$. Correspondingly, the exit intersection is located at the downstream of the cluster and is denoted by $f \in \mathcal{F}$ as well as the link just upstream (see Figure 1(a)). Figure 1(b) depicts the schematic of the case study arterial, while Figure 1(c) illustrates the queue length of links in the arterial with the initial signal plan normalised by the links lengths. Queue spillback from 'University' intersection on several links up to 'Dwight' link is apparent. Hence, 'University' intersection is the exit intersection of the cluster; see Figure 1(c), whereas 'Addison', 'Allston', 'Dwight', and 'Grayson' intersections are successively the entrance intersections during the simulation. Traffic Congestion is consistently present from cycle 25 until the end of the simulation when very low demand enables all long queues to clear. Arterial partitioning with the early detection method and the initial signal plan in Figure 1(d) reveals that besides the queue spillbacks, 'Cedar' and 'Gilman' links also face oversaturated conditions for about half of the simulation. Thus, the arterial partitioning method appears to be a valuable performance measurement tool and can be integrated within coordinated or actuated control strategies, with the objective of decreasing the risk of spillbacks.

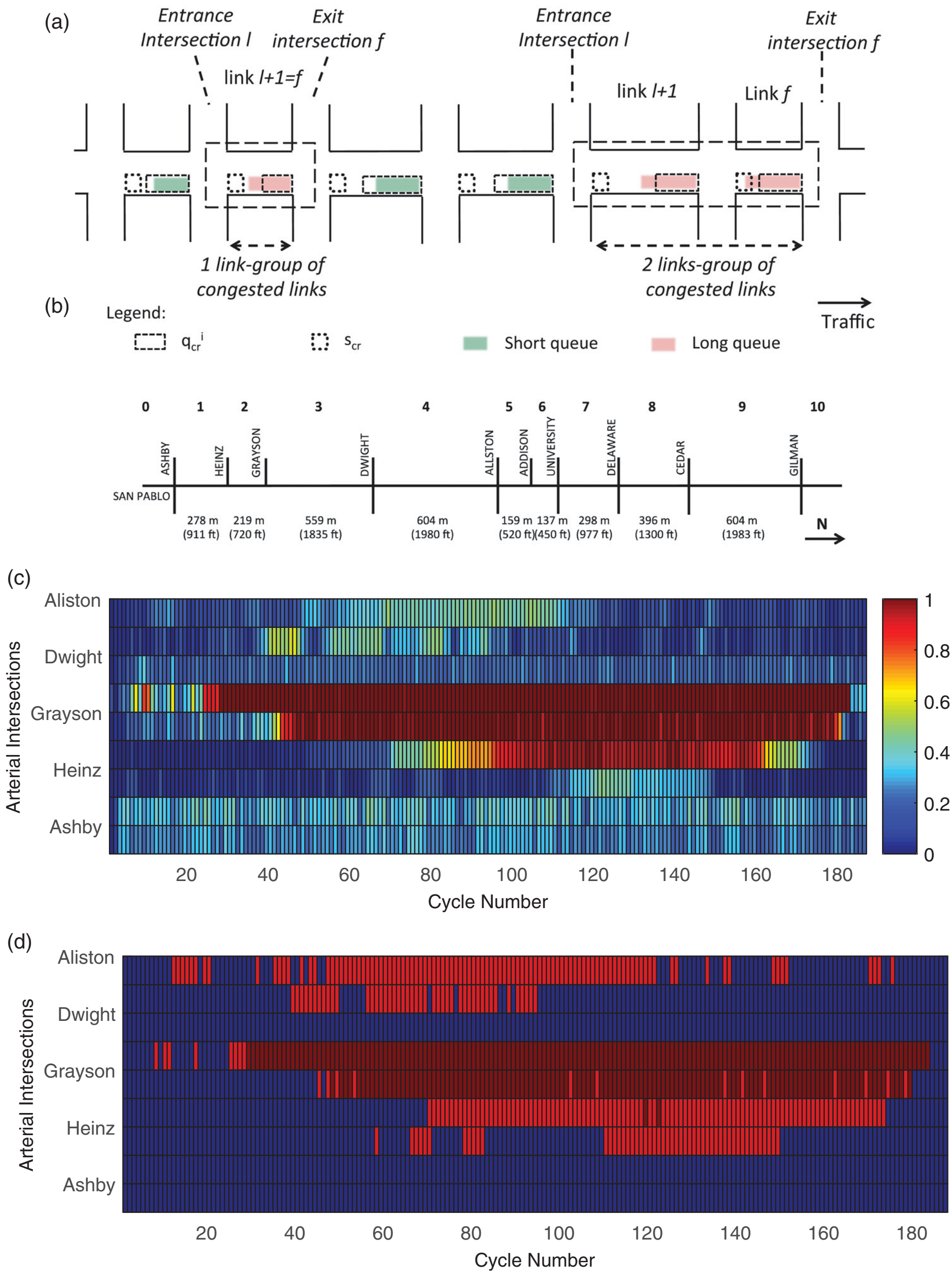

Figure 1. Arterial partitioning of San Pablo Avenue (Berkeley, CA) the northbound approach. (a) Schematic example of arterial partitioning with early detection model. (b) Study section of San Pablo arterial (The significant directional traffic is northbound). (c) Evolution of normalised links queue lengths over time with the initial control plan (links queue occupancy from 0 to 1). (d) Arterial partitioning with early detection of long queues and the initial control plan (blue: no detection of long queue; red: long queue detection; dark red: queue spillover detection).

3. Signal control methodology

Arterial partitioning at each cycle identifies cluster(s) of congested links with corresponding exit and entrance intersections that are delimited with uncongested links. The key intuition of the proposed signal control is to act only on exit and entrance intersections. If one considers a cluster of congested links as a whole, that is, a reservoir system, to control its accumulation, which in general is desirable to decrease to the uncongested state, it is needed to manipulate the inflow to and the outflow from the cluster. Therefore, the key idea is to increase the green phase duration of the arterial through movement at the exit intersection (increase the outflow) and also, if practical, to decrease the green phase duration of the arterial through movement at the entrance intersection and in certain cases, at more upstream intersections than the entrance intersection (decrease the inflow). The underlying assumption is that the traffic demand on the arterial is higher than the traffic demand on cross-streets. (In case otherwise, decreasing the inflow is associated with the green phase reduction of the cross-street turning movement.) The additional time needed for green extension at the exit intersection is deducted from the through movement phase of the corresponding cross-street and similarly, the extra time resulting from the green time reduction at the entrance intersection is added to the through movement phase of the corresponding cross-street. The green time extension at the exit intersection enables us to increase the discharging traffic flow and thus clear more quickly long queues, while a careful management of green time reduction at the entrance intersection mitigates the effect of queue spillbacks. The corollary of such control strategy is that the interior intersections of a cluster of congested links are still controlled with the initial signal plan. Note that, with the proposed control strategy the cycle length of intersections and offsets remain unchanged, which keeps the condition of coordination among the signals efficient.

In the sequel, first the signal control strategy for exit intersections is introduced. Then, the entrance intersection control strategy is presented. Finally, constraints to guarantee a satisfactory level of service for the traffic on the cross-street at exit intersections are described.

3.1. Exit intersection signal control

At the beginning of cycle $k+1$ for exit intersection f, the green phase extension is calculated based on the difference between the queue length, q_k^f, and $q_{\rm cr}^f$. In addition, when the risk of a potential queue spillback in link f becomes significant, that is, $s_k^f < s_{\rm cr}$, another compensatory term is added to the green extension. Hence as long as $q_k^f > q_{\rm cr}^f$, the feedback control strategy is implemented based on the attributes of the two last cycles:

$$g_{k+1}^f = g_k^f + K[q_k^f - q_{k-1}^f] + K_{s1}[\min(s_{k-1}^f, s_{\rm cr}) - \min(s_k^f, s_{\rm cr})], \tag{1}$$

where g_k^f is the cycle k green time for the through arterial movement at intersection f, and K and K_{s1} represent time increments (in seconds) for each additional vehicle queuing, respectively, beyond $q_{\rm cr}^f$ and passed $s_{\rm cr}$ threshold.[1]

Note that controller (1) is a feedback control law that during oversaturated traffic states and in case of no potential spillover ($q_k^f > q_{\rm cr}^f$ and $s_k^f > s_{\rm cr}$), alters the green duration to maintain and avoid the increase in the queue size at link f, until the control trigger condition, that is, $q_k^f > q_{\rm cr}^f$, is overruled. In case of potential spillover, controller (1) follows the same logic though with a higher gain that results to a faster response. With this configuration, the controller acts moderately in oversaturated conditions and when the chance of spillback becomes imminent, it acts more abruptly to avoid spillback occurrence. Note that, feedback controller parameters, K and K_{s1}, have to be chosen according to proper control engineering methods or manual fine-tuning to maximise the reduction of the network delay time; however, feedback controllers are shown to be robust to a moderate range of parameter values (Keyvan-Ekbatani et al. 2012).

3.2. Entrance intersection signal control

The objective of the entrance intersection control is to prevent or delay the occurrence of queue spillovers by reducing the amount of traffic inflow to the cluster of congested links. Queue spillovers reduce the intersection's capacity, which speeds up the extent of upstream link congestion and deteriorates the traffic flow rate in cross-streets. Therefore, reduction of the arterial through movement green time at the entrance intersection of the cluster is profitable particularly when there is a substantial risk of queue spillback, that is, when $s_k^{l+1} < s_{\mathrm{cr}}$. Hence, we propose the following feedback control:

$$g_{k+1}^l = g_k^l - K_{s2}[\min(s_{k-1}^{l+1}, s_{\mathrm{cr}}) - \min(s_k^{l+1}, s_{\mathrm{cr}})], \tag{2}$$

where K_{s2} is a parameter to be calibrated to maximise the reduction of the arterial delay time.

Some precautions should be applied to the entrance intersection control as too small green time reduction would not decrease queue spillback occurrence sufficiently, whereas a too large reduction would cause long queues in the upstream link leading to an additional queue spillback. When the link upstream of the entrance intersection is not long enough to accommodate extra vehicles, both above precautions are incompatible. Therefore, a multiple entrance control is proposed to account for short upstream links such that reduction of green phase duration for the arterial through movement is applied to all upstream intersections starting from the entrance intersection as long as the corresponding upstream links are too short (e.g. less than 250 [m]).

3.3. Signal control considerations

Green extension for the arterial through movement at the exit intersection should remain moderate so that a sufficient amount of time is reserved for the cross-street. Moreover, such priority given to arterial through movement has to be stopped once the queue in the cross-street becomes too long, because in this case, the benefits of the green extension for the arterial through movement would not compensate the additional delay time in the corresponding cross-street.

To avoid complicated requirements for signal control implementation, an all-or-nothing constraint is considered for the queues in cross-streets. Let $\hat{q}_{\mathrm{cr}}^f$ denotes the maximum cross-street queue that can be cleared during the initial green phase duration for the cross-street through movement. Let $\hat{q}_k^f$ denotes the queue in the cross-street of intersection f at cycle k and α be the parameter that designates the priority ratio of the arterial traffic flow over the cross-street traffic flow, thus $\alpha > 1$. Considering next cycle $k+1$, we propose that if $\hat{q}_k^f < \alpha \hat{q}_{\mathrm{cr}}^f$ then green extension at the exit intersection f is permitted, otherwise the green extension is prohibited. Note that the condition is checked for both directions of the cross-street. Moreover, pedestrian minimum green time requirements can also be integrated in the control strategy to guarantee a minimum green phase for the through movements of both the cross-street at the exit intersections and the arterial at the entrance intersections.

In addition, the reasoning behind the entrance intersection signal control is to restrict the inflow to the congested cluster. On the same note, because of hierarchical structure of urban networks, it is expected that arterials serve more traffic flow than other secondary roads. Hence, the most intuitive source of inflow to the congested cluster is the arterial through movement. However, in case of hypothetical situations such as a cross-street with considerable traffic and significant turning ratio into the arterial, the most effective approach to apply the green reduction is to decrease the inflow form the cross-street. In practical cases, a pre-study to determine aggregated ratio of traffic demand in arterial over the cross-streets along with consideration of turning ratios, can shed more light on practical aspects of green time reduction. Moreover, adding the extra green time to cross-streets with the green time reduction strategy might waste the intersection capacity in case there is not enough demand in the cross-street. To avoid such unnecessary waste of intersection capacity, the extra green time can be added to the phase that contributes less as an inflow to the congested cluster while there is the minimum risk of intersection capacity waste, for example, in case the opposite streams of the arterial belong to different phases, adding the extra green time to the other through movement of arterial seems more reasonable. Note that these practical modifications can be readily integrated to the control strategy for filed implementations.

4. Application on a real-world arterial site

We investigate the characteristics of the proposed signal control strategy based on the arterial partitioning method through extensive micro-simulation studies of San Pablo arterial in Berkeley, California. The arterial consists of 10 signalised intersections with two lanes per direction (without turning bays) from Ashby Avenue to Gilman Street as it is illustrated in Figure 1(b). The length of links ranges between 140 and 600 [m] for the total length of 3.2 [km]. The speed limit is 50 [km/h] and all the initial traffic plans are fixed-time coordinated with a common cycle length of 80 seconds. The study site with above properties and time-varying turning ratios and arterial and cross-streets demands (at all approaches) is implemented in Aimsun micro-simulation.

The simulation duration is 4 h (5–9 pm) and we are interested to implement the proposed dynamic signal control on the northbound approach, which has the heaviest directional traffic flow. Results with the fixed-time control and a medium level of demand are presented in Figure 1(c) and 1(d). For further investigation, we also consider a higher demand, in which the traffic demand in cross-streets are increased by 5%. The so-called high demand entails input demand flows close to the saturation flows for some cross-streets approaches.

We compare the initial fixed-time signal control with three adaptive signal control methods: (i) a signal control based on the arterial partitioning method and only exit intersection control strategy (named as 'Exit only'), (ii) a signal control with the arterial partitioning method and exit and entrance control strategies (named as 'Exit and Entrance'), and (iii) a non-coordinated dynamic signal control (named as 'Non-Coordinated'), in which a green extension similar to the one for exit intersection is implemented in downstream intersection of a link once the link is congested. The latter control strategy does not rely on the arterial partitioning procedure, thus comparisons will shed light on benefits of the arterial partitioning approach. Moreover, 'Exit only' strategy does not include the green reduction at the entrance intersection; hence comparisons demonstrate the green reduction strategy effect on the overall performance of the proposed control strategy. Results are listed in Table 1 representing the network, including all the approaches of the arterial and the cross-streets, total delay and number of stops per vehicle averaged over 10 runs. The values in parentheses show the improvement percentage of the adaptive control methods over the initial fixed-time signal control. Note that fine tuning of the control parameters has been investigated carefully. The obtained values are,

$$K = 0.2\left[\frac{(\mathrm{s/veh})}{\mathrm{lane}}\right], K_{s1} = 0.25\left[\frac{(\mathrm{s/veh})}{\mathrm{lane}}\right], K_{s2} = 0.4\left[\frac{(\mathrm{s/veh})}{\mathrm{lane}}\right], \quad \text{and} \quad s_{\mathrm{cr}} = 15\left[\frac{\mathrm{veh}}{\mathrm{lane}}\right].$$

Also, parameter α that designates the priority ratio of arterial traffic flow over cross-street traffic flow is fixed to 1.5. Sensitivity analyses reveal that with α between 1 and 2, average delay time is reduced. Nevertheless, low α values (i.e. close to 1) and very large values are sub-optimal, since a too restrictive constraint would not efficiently prevent queue spillback occurrences along the arterial, and on the contrary, too high values would durably deteriorate the traffic state in cross-streets. Regarding multiple entrance control, a threshold value of 250 [m] has been retained to differentiate short and long links.

Figure 2(a) and 2(b) display results for one single simulation with 'Exit and Entrance' control strategy. Comparison of Figures 1(c) and 2(a) demonstrates the efficiency of the signal control based on the arterial partitioning method since queues are contained between the critical intersection of 'University'

Table 1. Evaluation of different signal control strategies.

	Network total delay [s/km/veh]		Number of stops [per vehicle]	
Signal control strategy	Medium demand	High demand	Medium demand	High demand
Fixed-time	66.6 (—)	75.6 (—)	1.58 (—)	1.85 (—)
Non-coordinated	63.8 (4.2%)	74.4 (1.4%)	1.55 (1.9%)	1.84 (0.5%)
Exit only	58.5 (12.2%)	71.2 (5.7%)	1.48 (6.3%)	1.80 (2.7%)
Exit and entrance	57.9 (13.1%)	68.6 (9.3%)	1.47 (6.9%)	1.77 (4.3%)

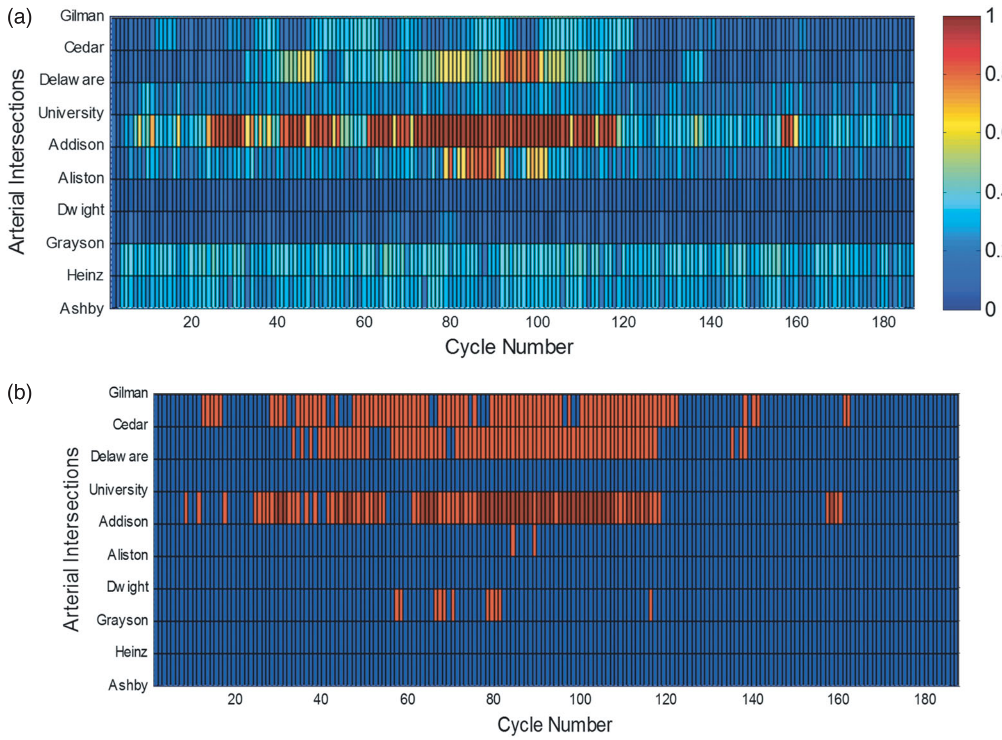

Figure 2. Simulation results with the signal control strategy based on the arterial partitioning method. (a) Evolution of normalised link queue lengths over time with the proposed signal control based on the arterial partitioning method. (b) Arterial partitioning with early detection of long queues and the signal control based on the arterial partitioning method (blue: no detection of long queue; red: long queue detection; dark red: queue spillover detection).

and Allston Street and are dissipated within 70 cycles (i.e. 1.5 h) before the end of the simulation. Moreover, arterial partitioning results in Figure 2(b) indicate that the link downstream of Allston intersection faces oversaturated conditions only during two cycles.

For both demand levels, signal control methods based on the arterial partitioning approach outperform the non-coordinated dynamic signal control, which shows very low performance specifically for the high demand case. Utilising arterial partitioning, 'Exit and Entrance' control strategy outperforms 'Exit only' strategy, specifically in case of the higher demand the amelioration is significant (+59% of network total delay time reduction with addition of entrance control). Indeed, the higher the level of demand is, the more crucial is to limit the upstream traffic flow to mitigate congestion spread and queue spillback occurrence. Besides, non-coordinated dynamic signal control leads to higher standard deviations of distribution of delay time (5.4 [s/km/veh]) over 10 replications than the initial fixed-time control (4.1 [s/km/veh]), whereas arterial partitioning strategies reduce the variability of the arterial delay time (3.8 [s/km/veh]).

Further comparison between signal control strategies is conducted through analysis of green phase extension or reduction for the arterial through movement. The cumulative green extensions during the whole simulation with 'non-coordinated' signal control is twice the amount associated with the 'Exit and Entrance' control strategy that is based on the arterial partitioning approach (e.g. for one simulation, 854 [s] versus 441 [s]). Moreover, Table 1 shows that 'non-coordinated' signal control is not as satisfactory as the other ones. Therefore, green time extensions in 'non-coordinated' strategy are largely spoiled, because it strongly penalises the cross-streets and the opposite direction of the through movement. On the contrary, the arterial partitioning method enables a smart allocation of a reduced amount of additional green times. Moreover, when a specific control is activated at the

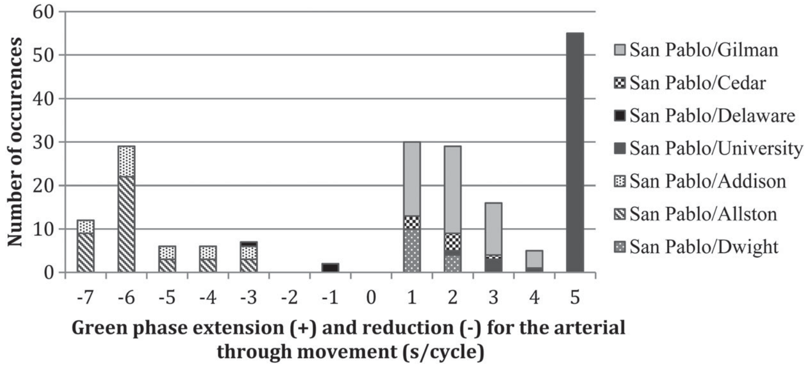

Figure 3. Distribution of green phase extension and reduction for one scenario with 'Exit and Entrance' control strategy.

Table 2. Comparison of average reduction of delay time with medium demand between arterial approaches and cross-streets.

	Average reduction of delay time (s/veh) and improvements (%) compared to the initial signal plan		
Approach	Non-coordinated	Exit only	Exit and entrance
Network (s/km/veh)	2.8 (4.2%)	8.1 (12.2%)	8.7 (13.1%)
Arterial NB	313 (52%)	309.5 (51.3%)	309.7 (51.1%)
Arterial SB	2.6 (2.1%)	1.7 (1.3%)	1.0 (0.8%)
Dwight EB	−2.8 (−7.6%)	−0.3 (−1.0%)	0.3 (0.4%)
Dwight WB	−1.6 (−5%)	0.5 (1.4%)	−0.1 (−0.2%)
Allston EB	2.0 (3.7%)	9.6 (21.2%)	14 (31.3%)
Allston WB	12 (2.3%)	32.3 (30.4%)	44.8 (43.9%)
Addison EB	−4.9 (−16.5%)	−0.1 (−0.3%)	2.7 (8.9%)
University EB	−19.8 (−56.1%)	−7.8 (−21.8%)	−8.5 (−24%)
University WB	−31.5 (−62.2%)	−7.6 (−15.5%)	−9.0 (−17.6%)
Delaware EB	0.8 (2.1%)	1.9 (5.3%)	1.9 (5.3%)
Delaware WB	2.1 (3.7%)	2.0 (2.5%)	2.7 (4.6%)
Cedar EB	−8.9 (−20.8%)	−0.6 (−2.8%)	2.2 (3.4%)
Cedar WB	−9.1 (−18.2%)	0.1 (−2.7%)	−1.4 (−5.0%)
Gilman EB	−6.2 (−16.3%)	−7.4 (−18.9%)	−5.6 (−14.5%)
Gilman WB	−20.5 (−36.2%)	−12 (−21.7%)	−10.4 (−17.9%)

Notes: NB: northbound, SB: southbound, EB: eastbound, and WB: westbound approach.

entrance intersections, green time reduction for the arterial through movement induces an increase of the green time for the corresponding cross-streets. Hence, the entrance intersection control redistributes the green time of some deprived cross-streets to others, as it is illustrated in Figure 3, being thus more equitable for the whole network.

Figure 3 shows that green extensions and reductions for the arterial through movement are limited respectively to 5 and 7 [s] per cycle. Therefore, full arterial partitioning strategy succeeds in reducing significantly the network total delay time with only small adjustments of phase durations. Note that the constraints introduced with pedestrian minimum crossing time requirements are eventually not binding. Also, it is clear that two separated groups of intersections are concerned by green phase extensions (i.e. University, Gilman, Cedar and Dwight intersections) and reductions (i.e. Allston, Addison and Delaware intersections). Comparative analysis of the benefits of different control strategies for cross-streets and both directions of the arterial is performed in Table 2. First of all, the three control strategies results are almost identical for the northbound approach of the arterial. Time saving exceeds 5 min for this approach, which represents more than 50% of delay time. Besides, dynamic signal control strategies that are implemented for arterial northbound approach do not impact the southbound approach. In particular, the entrance control does not deteriorate the arterial southbound traffic.

Considering the cross-streets, signal control strategies based on the arterial partitioning method strongly outperform 'non-coordinated' signal control. This can be highlighted with 'Allston' westbound approach, which is upstream to San Pablo/University critical intersection. Using a control strategy, congestion in the arterial is less than with the initial control and thus, queue spillbacks are limited so that traffic coming from Allston Street can enter into the arterial more easily. However, the delay time reduction with 'non-coordinated' signal control is significantly less than the other ones. This happens because green phase duration for the arterial through movement at San Pablo/Allston intersection, which is in the middle of a group of congested links, is inefficiently increased with non-coordinated strategy since no enough space is available in the downstream link to dissipate the queue. On the contrary, the arterial partitioning approach prohibits any green extension on interior intersections of a group of congested links. Finally, the constraint on cross-street traffic avoids huge cross-streets delays (e.g. University, Cedar or Gilman) that cross the arterial at intersections that are often exit intersections, by prohibiting (if necessary) the green phase extension for the arterial through movement.

Figure 4 presents time-space trajectory diagrams for the four signal control scenarios with the high demand during three cycles (i.e. 240 s). Although congestion is widely spread for the initial signal control with spillbacks in San Pablo/Addison, Allston, and Dwight intersections (see Figure 4(a)), signal

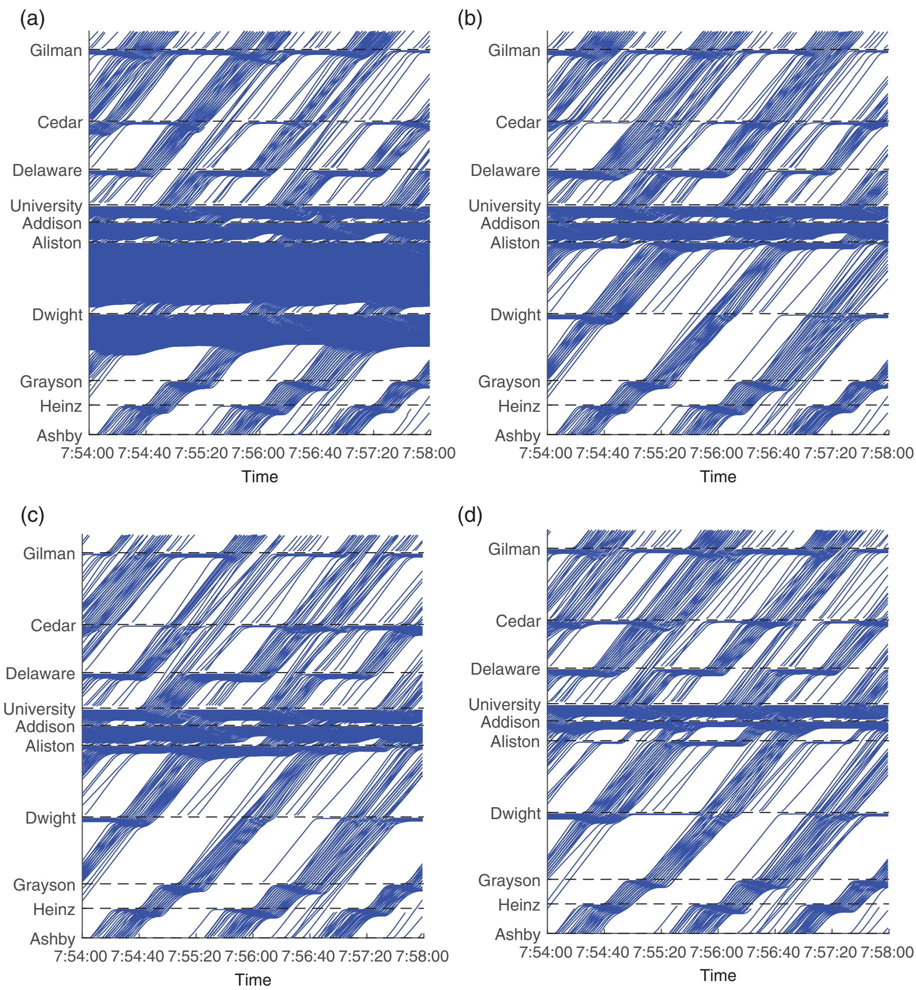

Figure 4. Time-space diagrams depicting vehicles trajectories for the high demand scenario. (a) Initial signal control. (b) Arterial partitioning with exit signal control only. (c) 'Non-coordinated' dynamic signal control. (d) Arterial partitioning with exit and entrance signal control.

control strategies ameliorate the arterial congestion. Nevertheless, queue spillbacks still occur in San Pablo intersections with Addison and Allston without the arterial partitioning (Figure 4(c)). Finally, congestion is more efficiently managed with arterial partitioning and 'Exit and Entrance' control strategy (Figure 4(d)) where green time reduction in San Pablo/Allston intersection, which is the intersection upstream to the entrance intersection, limits queue build-up in the short link downstream. Similarly, green time reduction in San Pablo/Allston entrance intersection creates a negative offset between it and San Pablo/University critical intersection, which reduces the inefficient green time duration for the arterial through movement.

Further, we investigate how queue length measurement errors influence the effectiveness of the control strategy. To this end, we add an additive white noise (a random variable from a Gaussian distribution with zero mean and different standard deviations) to queue measurements, which affects both the clustering and control processes. The results demonstrate that the control strategy is robust to measurement errors, that is, the control performance changes insignificantly with errors in the range of 5 vehicles while the control performance with errors in range of 10 vehicles outperforms the initial signal plan (in case of medium demand and 'Exit and Entrance' control strategy the network total delay equals to 64.06 [s/km/veh]). There is a vast literature that provides queue lengths estimates utilising data from loop detectors and/or probe vehicles with an accuracy which is sufficient for our approach (e.g. Liu et al. 2009; Ban, Hao, and Sun 2011; Ramezani and Geroliminis 2015).

Finally, the performance of the signal control strategy based on the arterial partitioning method has been compared to static signal settings resulting from Synchro optimisation software with a 15 min update window. To this end, arterial signal splits, cycle length and offsets were determined with Synchro based on a static traffic input. Then, we test both the proposed dynamic signal control strategy based on the arterial partitioning method and Synchro pre-timed signal control. Average results for 10 runs with high demand show that while Synchro significantly improves the traffic on the arterial (56% reduction of delay times with Synchro for the northbound approach versus 31% for arterial partitioning approach compared to the initial signal plan), it does not guarantee an acceptable level of service for the cross-streets, which develops very long queues. Hence, Synchro signal settings lead to a drastic increase of network total delay in comparison with the initial fixed-time signal plan. In addition, the zone separation tool in Synchro, which enables multiple cycle length along the arterial, does not provide improvements. On the contrary, the proposed signal control improves the traffic condition on the arterial and simultaneously accounts for cross-street traffic because of the explicit threshold on cross-streets queues.

5. Conclusions

This paper has presented an elegant traffic-responsive signal control strategy based on an arterial partitioning approach. Clusters of successive congested links in the arterial are first detected, followed by identifying entrance and exit intersections of these clusters of congested links. Given this segmentation, locally smaller-sized decentralised dynamic signal control strategies have been triggered to mitigate long queues and to prevent spillbacks. The traffic control strategy extends the green phase duration of the arterial through movement at the exit intersections and restricts the inflow at the entrance intersections and upstream – in case of closely spaced intersections – when a potential risk of queue spillback exists.

The major contributions of the proposed method are the following: First, the signal control is proactive and get activated once oversaturated conditions appear. Therefore, the control pre-empts any increase of congestion that could lead to queue spillbacks instead of reacting to already critical traffic states. Second, considering the empty space available beyond the back of queue, it explicitly integrates the potential risks of queue spillbacks. Finally, arterial partitioning approach enables to easily address locally oversaturated conditions ensuring coherence of the dynamic signal control at the scale of the arterial.

Simulations on a real-world arterial site demonstrate the proficiency of the proposed signal control strategy, which can reduce the delay time on the arterial even with a high level of traffic in cross-streets. In addition, the strategy outperforms 'non-coordinated' dynamic signal control, which allocates inefficient additional green time for the arterial through movement that strongly penalises cross-streets. On the contrary, entrance intersection control redistributes the green time of some deprived cross-streets due to exit intersections signal control to others. Therefore, the arterial partitioning approach has been found more equitable for the cross-streets. Moreover, from small green phase extensions and reductions for the arterial through movement significant improvements are achieved, which also avoids traffic deterioration of the arterial in the opposite direction.

With recent developments of queue length estimation based on probe vehicles, the proposed signal control approach can be readily applicable in real-world arterials. However, additional simulation tests on other networks would be of great interest to scrutinise the effect of network characteristics such as the critical link length and the impacts of non-protected left-turns for both arterial and cross-street approaches on the performance of the control strategy. In addition, the static constraint on queue lengths in cross-streets could be adjusted based on the level of congestion in the arterial and level of service of cross-streets. Further research could compare between the proposed method and the Synchro with adaptive coordinated signal timing along with the investigation of bi-directional signal control based on the arterial partitioning approach. A future research direction is to extend the control strategy to be applicable to networks, which include several arterials with competing traffic flow directions.

Another important application of such a controller is its integration in a hierarchical framework at the network level. Recent work in perimeter control for large-scale urban networks (e.g. Keyvan-Ekbatani et al. 2012; Geroliminis, Haddad, and Ramezani 2013; Ramezani, Haddad, and Geroliminis 2015) has developed efficient control of transfer flows between different regions of a city at the network scale. Such approaches might develop local spillbacks in the perimeter of the applied control. The strategies of this paper can significantly contribute to avoid the spillbacks and to distribute traffic in the proximity of perimeter signal control locations in a smooth way. Such direction is under development.

Note

1. The control strategy for the first cycle of activation reads: $g^f_{k+1} = g^f_{\text{init}} + K[q^f_k - q^f_{\text{cr}}] + K_{s1}[s_{\text{cr}} - \min(s^f_k, s_{\text{cr}})]$.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding information

This research was financially supported by ERC Starting Grant 'METAFERW: Modelling and controlling traffic congestion and propagation in large-scale urban multimodal networks'.

ORCID

Mohsen Ramezani http://orcid.org/0000-0001-6839-6839

References

Abbas, M., D. Bullock, and L. Head. 2001. "Real-time Offset Transitioning Algorithm for Coordinating Traffic Signals." *Transportation Research Record* 1748: 26–39.

Aboudolas, K., and N. Geroliminis. 2013. "Perimeter and Boundary Flow Control in Multi-reservoir Heterogeneous Networks." *Transportation Research Part B: Methodological* 55: 265–281.

Abu-Lebdeh, G., and R. Benekohal. 1997. "Development of Traffic Control and Queue Management Procedures for Oversaturated Arterials." *Transportation Research Record* 1603: 119–127.

Asthana, R., N. J. Ahuja, M. Darbari, and P. K. Shukla. 2012. “A Critical Review on the Development of Urban Traffic Models & Control Systems.” *International Journal of Scientific & Engineering Research* 3: 1–5.
Ban, X., P. Hao, and Z. Sun. 2011. “Real Time Queue Length Estimation for Signalized Intersections using Travel Times from Mobile Sensors.” *Transportation Research Part C: Emerging Technologies* 19 (6): 1133–1156.
Cheng, Y., X. Qin, J. Jin, and B. Ran. 2012. “An Exploratory Shockwave Approach to Estimating Queue Length Using Probe Trajectories.” *Journal of Intelligent Transportation Systems* 16 (1): 12–23.
Christofa, E., J. Argote, and A. Skabardonis. 2013. “Arterial Queue Spillback Detection and Signal Control Based on Connected Vehicle Technology.” *Transportation Research Record,* 2356 (1): 61–70.
Comert, G., and M. Cetin. 2009. “Queue Length Estimation from Probe Vehicle Location and the Impacts of Sample Size.” *European Journal of Operational Research* 197 (1): 196–202.
Diakaki, C., M. Papageorgiou, and K. Aboudolas. 2002. “A Multivariable Regulator Approach to Traffic-responsive Network Wide Signal Control.” *Control Engineering Practice* 10: 183–195.
Gal-Tzur, A., D. Mahalel, and J. Prashker. 1993. “Signal Design for Congested Networks Based on Metering.” *Transportation Research Record* 1398: 111–118.
Gazis, D. 1964. “Optimum Control of a System of Oversaturated Intersections.” *Operations Research* 12: 815–831.
Geroliminis, N., J. Haddad, and M. Ramezani. 2013. “Optimal Perimeter Control for Two Urban Regions with Macroscopic Fundamental Diagrams: A Model Predictive Approach.” *IEEE Transactions on Intelligent Transportation Systems* 14 (1): 348–359.
Geroliminis, N., and A. Skabardonis. 2011. “Identification and Analysis of Queue Spillovers in City Street Networks.” *IEEE Transactions on Intelligent Transportation Systems* 12 (4): 1107–1115.
Gordon, R. A. 1969. “Technique for Control of Traffic at Critical Intersections.” *Transportation Science* 3 (4): 279–288.
Haddad, J., B. De Schutter, D. Mahalel, I. Ioslovich, and P. O. Gutman. 2010. “Optimal Steady-state Control for Isolated Traffic Intersections.” *IEEE Transactions on Automatic Control* 55: 2612–2617.
Haddad, J., and D. Mahalel. 2014. “Offset Effects on the Capacity of Paired Signalised Intersections during Oversaturated Conditions.” *Transportmetrica A: Transport Science* 10 (8): 740–758.
Head, L., and P. Mirchandani. 2001. “A Real-time Traffic Signal Control System: Architecture, Algorithms, and Analysis.” *Transportation Research Part C: Emerging Technologies* 9 (6): 415–432.
Keyvan-Ekbatani, M., A. Kouvelas, I. Papamichail, and M. Papageorgiou. 2012. “Exploiting the Fundamental Diagram of Urban Networks for Feedback-based Gating.” *Transportation Research Part B: Methodological* 46 (10): 1393–1403.
Keyvan-Ekbatani, M., M. Yildirimoglu, N. Geroliminis, and M. Papageorgiou. 2015. “Multiple Concentric Gating Traffic Control in Large-scale Urban Networks.” *IEEE Transactions on Intelligent Transportation Systems* 16 (4): 2141–2154.
Liu, H., X. Wu, W. Ma, and H. Hu. 2009. “Real-time Queue Length Estimation for Congested Signalized Intersections.” *Transportation Research Part C: Emerging Technologies* 17 (4): 412–427.
May, A., F. Montgomery, and D. Quinn. 1988. “Control of Congestion in Highly Congested Networks.” CODTU IV conference, Jakarta, Indonesia.
Michalopoulos, P., and G. Stephanopoulos. 1978. “Optimal Control of Oversaturated Intersections: Theoretical and Practical Considerations.” *Traffic Engineering and Control* 19 (5): 216–221.
Papageorgiou, M., C. Diakaki, V. Dinopoulou, A. Kotsialos, and Y. Wang. 2003. “Review of Road Traffic Control Strategies.” *Proceedings of the IEEE* 91 (12): 2043–2067.
Pignataro, L., W. McShane, and K. Crowley. 1978. “Traffic Control in Oversaturated Street Networks.” National cooperative highway research program report 194, Transportation Research Board, National Research Council.
Quinn, D. 1992. “A Review of Queue Management Strategies.” *Transportation Research Record* 33: 600–605.
Ramezani, M., and N. Geroliminis. 2012. “On the Estimation of Arterial Route Travel Time Distribution with Markov Chains.” *Transportation Research Part B: Methodological* 46: 1576–1590.
Ramezani, M., and N. Geroliminis. 2015. “Queue Profile Estimation in Congested Urban Networks with Probe Data.” *Computer-Aided Civil and Infrastructure Engineering* 30 (6): 414–432.
Ramezani, M., J. Haddad, and N. Geroliminis. 2015. “Dynamics of Heterogeneity in Urban Networks: Aggregated Traffic Modelling and Hierarchical Control.” *Transportation Research Part B: Methodological* 74: 1–19.
Rathi, A. 1988. “A Control Scheme for High Traffic Density Sectors.” *Transportation Research Part B: Methodological* 22 (2): 81–101.
Shelby, S., D. Bullock, Z. Sabra, R. Ghaman, N. Soyke, and D. Gettman. 2008. “Overview and Performance Evaluation of ACS Lite: Low-cost Adaptive Signal Control System.” Pre-print CD-ROM. 87th TRB annual meeting, Washington, DC.
Skabardonis, A., and N. Geroliminis. 2008. “Real-time Monitoring and Control on Signalized Arterials.” *Journal of Intelligent Transportation Systems* 12 (2): 64–74.
TrafficWare. 2004. *SYNCHRO 6 User Guide*. Albany, NY: Trafficware.
Varaiya, P. 2013. “Max Pressure Control of a Network of Signalized Intersections.” *Transportation Research Part C: Emerging Technologies* 36: 177–195.
Wu, X., H. X. Liu, and D. Gettman. 2010. “Identification of Oversaturated Intersections using High-resolution Traffic Signal Data.” *Transportation Research Part C: Emerging Technologies* 18 (4): 626–638.