Control and Optimization Methods for Traffic Signal Control in Large-scale Congested Urban Road Networks

Konstantinos Aboudolas, Markos Papageorgiou, and Elias Kosmatopoulos

Dynamic Systems and Simulation Laboratory
Technical University of Crete, GR-731 00 Chania, Greece
E-mail: {aboud,markos,kosmatop}@dssl.tuc.gr

***Abstract*— The problem of designing real-time traffic signal control strategies for large-scale congested urban road networks via suitable application of control and optimization methods is considered. Three alternative methodologies are proposed, all based on the store-and-forward modeling (SFM) paradigm. The first methodology results in a linear multivariable feedback regulator derived through the formulation of the problem as a Linear-Quadratic (LQ) optimal control problem. The second methodology leads to an open-loop constrained quadratic optimal control problem whose numerical solution is achieved via quadratic-programming (QP). Finally, the third methodology leads to an open-loop constrained nonlinear optimal control problem whose numerical solution is effectuated by use of a feasible-direction algorithm. A simulation-based investigation of the signal control problem for a large-scale urban network using these methodologies is presented. Results demonstrate the efficiency and real-time feasibility of the developed generic control methods.**

I. Introduction

In view of the imminent traffic congestion and lack of possibilities for infrastructure expansion in urban road networks, the importance of efficient signal control strategies, particularly under saturated traffic conditions, can hardly be overemphasized. It is generally believed that real-time systems responding automatically to the prevailing traffic conditions, are potentially more efficient than clock-based fixed-time control settings, possibly extended via a simple traffic-actuated logic.

On the other hand, the development of network-wide real-time signal control strategies using elaborated network models is deemed infeasible due to the combinatorial nature of the related optimization problem [1]; as a consequence, the developed or implemented signal control strategies include many simplifications or heuristics which may render the strategies less efficient, particularly under saturated traffic conditions, unless a high effort is put in the fine-tuning of many parameters included in the signal control strategy.

The purpose of this paper is to demonstrate how control and optimization methods can be used for designing real-time traffic signal control strategies for large-scale congested urban networks. More specifically, a generic mathematical model, which is based on the SFM paradigm, for the traffic flow process in large-scale urban networks is developed first, upon which three alternative optimal control methodologies are applied for the design of signal control strategies that aim at minimizing and balancing the link queues so as to reduce the risk of queue spillback. Depending on the applied control methodology, signal control plans are computed in real time through a linear multivariable feedback regulator, an open-loop constrained quadratic optimal control problem (QPC), or an open-loop constrained nonlinear optimal control problem (NOC). In order to evaluate the efficiency of the proposed generic framework, we compare the closed-loop behaviour of the linear multivariable regulator with the open-loop behaviour of QPC and NOC methodologies.

II. Background

A variety of traffic signal control strategies for urban networks have been developed during the past few decades. Without attempting a survey of this vast research area we will address a few selected strategies (for an up-to-date account we refer the reader to [1]), some of which have been implemented in real-life conditions while others are still in the research and development stage. We may distinguish two principal classes of signal control strategies. In the first class, strategies are only applicable to (or efficient for) networks with undersaturated traffic conditions, whereby all queues at the signalized junctions are served during the next green phase. In the second class, we consider strategies applicable to networks with oversaturated traffic conditions, whereby queues may grow in some links with an imminent risk of spillback and eventually even of gridlock in network cycles.

Fixed-time strategies for isolated junction control (stage-based approaches SIGSET [2], SIGCAP [3] using the well-known Webster's delay formula) or coordinated control have been widely used due to their simplicity of implementation in networks with undersaturated traffic conditions. Arterial progression schemes that maximize the bandwidth of progression (MAXBAND [4], MULTIBAND [5]), and more general network optimization schemes that minimize delay, stops or other measures of disutility (TRANSYT-7F [6]) are also in use. The main drawback of these strategies is that their settings are based on historical rather than real-time data.

SCOOT [7] and SCATS [8] are two well-known and widely used coordinated traffic-responsive strategies. These well-designed strategies function effectively when the traffic conditions in the network are below saturation but their performance deteriorates when severe congestion persists during the rush period. Other model-based traffic-responsive strategies such as PRODYN [9] and RHODES [10] employ dynamic programming while OPAC [11] employs exhaustive enumeration. Due to the exponential complexity of these solution algorithms, the basic optimization kernel is not real-time feasible for more than one junction.

Store-and-forward modeling of traffic networks was first suggested by Gazis and Potts [12] and has since been used in various works notably for road traffic control. This modeling philosophy offers a major advantage: it allows highly efficient optimization and control methods to be used for large-scale congested urban networks. A recently developed strategy of this type is the signal control strategy TUC [13] that will be outlined later.

More recently, a number of strategies have been proposed employing various computationally expensive numerical solution algorithms, including genetic algorithms [14], [15], multi-extended linear complementary programming [16], and mixed-integer linear programming [17], [18]. In [17], [15], and [18] the traffic flow conditions are modeled using the cell transmission model [19], a convergent numerical approximation to the first-order hydrodynamic model of traffic flow. However, these approaches are in a relatively premature stage and their implementation and feasibility in real-life and real-time conditions are still questionable.

III. The Signal Control Design Problem

We start with a brief description of the problem of designing signal control strategies via the SFM philosophy and a definition of the main control objective.

A. Problem Formulation

The urban road network is represented as a directed graph with links $z \in Z$ and junctions $j \in J$. For each signalized junction j, we define the sets of incoming I_j and outgoing O_j links. It is assumed that the offset, the cycle time C_j, and the lost time L_j of junction j are fixed. In addition, to enable network offset coordination, we assume that $C_j = C$ for all junctions $j \in J$. Furthermore, the signal control plan of junction j is based on a fixed number of stages that belong to the set F_j, while v_z denotes the set of stages where link z has right of way (r.o.w.). Finally, the saturation flow S_z of link $z \in Z$, and the turning movement rates $t_{w,z}$, where $w \in I_j$ and $z \in O_j$, are assumed to be known and fixed.

By definition, the constraint

$$\sum_{i \in F_j} g_{j,i} + L_j = C \tag{1}$$

holds at junction j, where $g_{j,i}$, is the green time of stage i at junction j. In addition, the constraint

$$g_{j,i} \geq g_{j,i,\min}, \quad i \in F_j \tag{2}$$

where $g_{j,i,\min}$ is the minimum permissible green time for stage i at junction $j \in J$, is introduced to guarantee allocation of sufficient green time to pedestrian phases.

Consider a link z connecting two junctions M and N such that $z \in O_M$ and $z \in I_N$ (Fig. 1). The dynamics of link z

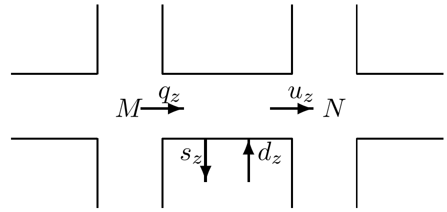


Fig. 1. An urban road link.

are given by the continuity equation

$$x_z(k+1) = x_z(k) + T\big[q_z(k) - s_z(k) + d_z(k) - u_z(k)\big] \tag{3}$$

where $x_z(k)$ is the number of vehicles within link z at time kT, $q_z(k)$ and $u_z(k)$ are the inflow and outflow, respectively, of link z in the sample period $[kT, (k+1)T]$; with T the discrete time step and $k = 0, 1, \ldots$ the discrete-time index. In addition, d_z and s_z, are the demand and the exit flow within the link, respectively. For the exit flow we set $s_z(k) = t_{z,0} q_z(k)$, where the exit rates $t_{z,0}$ are assumed to be known.

Queues are subject to the constraints

$$0 \leq x_z(k) \leq x_{z,\max}, \quad \forall\, z \in Z \tag{4}$$

where $x_{z,\max}$ is the maximum admissible queue length. This constraint may automatically lead to a suitable upstream gating in order to protect downstream areas from oversaturation during periods of high demand.

The inflow to the link z is given by $q_z(k) = \sum_{w \in I_M} t_{w,z} u_w(k)$, where $t_{w,z}$ with $w \in I_M$ are the turning movement rates towards link z from the links that enter junction M.

We now introduce a critical simplification for the outflow u_z that characterizes the suggested modeling approach. Assuming that space is available in the downstream links and that x_z is sufficiently high, the outflow (real flow) u_z of link z is equal to the saturation flow S_z if the link has r.o.w., and equal to zero otherwise. However, if the discrete time step T is equal to C, an average value for each period (modeled flow) is obtained by

$$u_z(k) = G_z(k) S_z / C \tag{5}$$

where G_z, is the green time of link z, calculated as $G_z(k) = \sum_{i \in v_z} g_{j,i}(k)$.

B. Control Objective

As already mentioned, the main control objective is to minimize the risk of oversaturation and spillback of link queues. To this end, one may attempt to minimize and balance the links' relative occupancies $x_z / x_{z,\max}$. This criterion is physically reasonable as well as convenient from the numerical solution point of view, as we will see later. Alternatively, one may minimize the total time spent but this may increase the risk of link queue spillback.

C. Linear-Quadratic Optimal Control

Replacing (5) in (3) leads to a linear state-space model for road networks of arbitrary size, topology, and characteristics which is given by

$$\mathbf{x}(k+1) = \mathbf{x}(k) + \mathbf{B}\Delta\mathbf{g}(k) + T\Delta\mathbf{d}(k) \tag{6}$$

where $\mathbf{x}(k)$ is the state vector (with elements the number of vehicles x_z of each link z); $\Delta\mathbf{g}(k) = \mathbf{g}(k) - \mathbf{g}^N$ and $\Delta\mathbf{d}(k) = \mathbf{d}(k) - \mathbf{d}^N$ are the control and demand deviations, respectively; $\mathbf{g}(k)$ is the control vector (with elements all the green times $g_{j,i}$ of stage i at junction j); $\mathbf{g}^N$ is a nominal control vector (with elements the nominal green times $g_{j,i}^N$ of stage i at junction j) which may correspond to a pre-specified fixed signal plan; $\mathbf{d}(k)$ is the disturbance vector (with elements the demand flows d_z of each link z); $\mathbf{d}^N$ is a nominal disturbance vector whereby $\mathbf{B}\mathbf{g}^N + T\mathbf{d}^N = \mathbf{0}$ holds for the nominal (e.g. steady-state) values. Finally $\mathbf{B}$ is a constant matrix of appropriate dimensions containing the network characteristics (saturation flows, turning rates).

A quadratic criterion that addresses the control objective has the general form

$$\mathcal{J} = \frac{1}{2}\sum_{k=0}^{\infty}\left(\|\mathbf{x}(k)\|_{\mathbf{Q}}^2 + \|\Delta\mathbf{g}(k)\|_{\mathbf{R}}^2\right) \qquad (7)$$

where $\mathbf{Q}$ and $\mathbf{R}$ are nonnegative definite, diagonal weighting matrices. The diagonal elements of $\mathbf{Q}$ are set equal to $1/x_{z,\max}$ in order to minimize and balance the relative occupancies of the network links. Furthermore, the magnitude of the control reactions can be influenced by the choice of the weighting matrix $\mathbf{R} = r\mathbf{I}$. To this end, the choice of r may be performed via a trial-and-error procedure so as to achieve a satisfactory control behaviour for a given application network.

Minimization of the cost criterion (7) subject to (6) (assuming $\Delta\mathbf{d}(k) = \mathbf{0}$) leads to a linear multivariable feedback regulator given by

$$\mathbf{g}(k) = \mathbf{g}^N - \mathbf{L}\mathbf{x}(k) \qquad (8)$$

where the feedback gain matrix $\mathbf{L}$ results as a straightforward solution of the corresponding algebraic Riccati equation. This is the multivariable regulator approach taken by the signal control strategy TUC [13] to calculate in real time the network splits, while cycle time and offset are calculated by other parallel algorithms [20].

Note that the LQ control theory does not allow for direct consideration of the constraints (1) and (2). For this reason, a suitable knapsack algorithm is used after the application of (8) that modifies the calculated $g_{j,i}$ green times of each junction so as to satisfy the constraints (1) and (2). More specifically, the knapsack algorithm for each junction j reads: for given $g_{j,i}$ (resulting from (8)), find the modified green times $\tilde{g}_{j,i}$ that minimize

$$\sum_{i\in F_j}\left(\tilde{g}_{j,i} - g_{j,i}\right)^2/g_{j,i} \qquad (9)$$

subject to (1), (2). It may be readily shown that the minimization of (9) subject to (1) leads to a solution that satisfies $\tilde{g}_{j,i}/g_{j,i} = \tilde{g}_{j,l}/g_{j,l}\ \forall(i,l)$, i.e. the modified $\tilde{g}_{j,i}$ preserve the same splits as $g_{j,l}$ along with satisfying (1). The above knapsack problem approximates this solution to the extent allowed by the additional constraint (2). The numerical solution of a knapsack problem is known to call for at most as many iterations as the number of involved variables, which, in our case, hardly exceeds 3 or 4 stages at each junction.

D. Open-loop Quadratic-Programming Control

In contrast to other SFM based approaches (see for instance [21]), we will now introduce the green times G_z of each link z as additional independent variables. The reason behind this modification is that we want to preserve model validity also under nonsaturated traffic conditions [22]. The introduced link green times G_z are constrained as follows:

$$0 \le G_z(k) \le \sum_{i\in v_z} g_{j,i}(k), \quad \forall j \in J. \qquad (10)$$

The main reason for introducing independent G_z in the problem formulation lies in the following observation: if the queue x_z is not sufficiently long or even zero; or if the downstream link queue is too long to accommodate a high inflow; then the constraints (4) will become active and will reduce the corresponding stage greens accordingly. As an illustrative example, assume that at a certain cycle there are two links z and w having r.o.w. simultaneously during a stage (M,i), and that $x_z \approx 0$ while $x_w \gg 0$. If G_z and G_w are not independently introduced, we have by definition $G_z = G_w = g_{M,i}$. Then, the stage green $g_{M,i}$ will be strictly limited by the constraint $x_z \ge 0$ although link w may need a longer green phase for dissolving x_w. In contrast, by introducing G_z and G_w independently, the algorithm can guarantee $x_z \ge 0$ by choosing G_z accordingly short without constraining G_w and the stage green. Similarly, if the link r downstream of link z is close to spillback, the constraint $x_r \le x_{r,\max}$ can be guaranteed by choosing G_z accordingly short without constraining the green time of other links that are having r.o.w. during the same stage.

In view of the above modification, replacing (5) in (3) leads to a linear state-space model for road networks of arbitrary size, topology, and characteristics

$$\mathbf{x}(k+1) = \mathbf{x}(k) + \overline{\mathbf{B}}(k)\mathbf{G}(k) + T\mathbf{d}(k) \qquad (11)$$

where $\mathbf{G}(k)$ is the link control vector with elements the green times G_z of each link z; $\overline{\mathbf{B}}$ is a matrix of appropriate dimensions containing the network characteristics. Note that in this approach $\overline{\mathbf{B}}$ may be time-variant, if the involved turning rates are time-variant.

In this case, a quadratic criterion that addresses the control objective has the form

$$\mathcal{J} = \frac{1}{2}\sum_{k=0}^{K}\sum_{z\in Z}\frac{x_z^2(k)}{x_{z,\max}}. \qquad (12)$$

On the basis of the linear model and the constraints presented in Section III-A plus the constraint (10) and the quadratic cost criterion (12), a (dynamic) optimal control problem may be formulated over a time-horizon K, starting with the known initial state $\mathbf{x}(0)$ in the state equation (11). More precisely, the resulting QP problem reads: minimization of the cost criterion (12) subject to (1), (2), (4), (10), (11). In summary, the optimization problem has three types of time-dependent decision variables, namely the stage green times $g_{j,i}(k)$, the state variables $x_z(k)$, and the link green times $G_z(k)$. This QP problem (with very sparse matrices) may be readily solved by use of broadly available codes or commercial software within few CPU-seconds even for large-scale networks and long time-horizons.

E. Open-loop Nonlinear Optimal Control

In this design approach, we re-introduce the outflow $u_z(k)$ into our problem and recall that the outflow is given by (5) only under the assumption that $x_z(k)$ satisfies the constraints (4). Instead of (5), we may now define a nonlinear outflow function that models the real traffic flow process more accurately. More precisely, assuming that $T \ll C$, the outflow $u_z(k)$ is given by

$$u_z(k) = \begin{cases} 0 & \text{if } x_{\mathrm{d},z}(k) \geq c x_{z,\max}(k) \\ \min\left\{\frac{x_z(k)}{T}, \frac{G_z(k) S_z}{C}\right\} & \text{else} \end{cases} \tag{13}$$

where $x_{\mathrm{d},z}(k)$ is a downstream link of link z, and $c \in (0,1]$. By introducing (13), the state variables are allowed to change their value more frequently than the control variables. More precisely, typical discrete-time model steps T for the traffic flow model (3) using (5) may be in the order of 5 s while the control variables change their value in discrete-time control steps T_c, e.g. at each cycle or more. Note that, when using (13), the queue constraints (4) are considered indirectly and may hence be dropped; the link outflow in (13) becomes zero if there is no vehicle in the link or the downstream link is full. Note also that the basic simplification of SFM, i.e. a continuous link outflow (rather than zero flow during red and free flow during green), is still maintained in this approach.

Replacing (13) in (3) we obtain a nonlinear state-space model for road networks of arbitrary size, topology, and characteristics [23]

$$\mathbf{x}(k+1) = \mathbf{f}\big[\mathbf{x}(k), \mathbf{g}(\kappa), \mathbf{d}(k)\big], \quad \kappa = [k/\tau] \tag{14}$$

where $\mathbf{f}$ is a nonlinear vector function, κ is a discrete-time index, and $T_\mathrm{c} = \tau T$.

The cost criterion in a nonlinear optimal control problem may have an arbitrary nonlinear form. In the particular case the cost criterion to be minimized has the general form

$$\mathcal{J} = \sum_{z\in Z} \frac{x_z^2(K)}{x_{z,\max}} + \sum_{k=0}^{K-1} \Bigg\{ \sum_{z\in Z} \frac{x_z^2(k)}{x_{z,\max}} + a_\mathrm{u} \sum_{j=1}^{|J|} \Phi_j\big[\mathbf{g}(\kappa)\big]^2 + a_\mathrm{f} \sum_{j=1}^{|J|} \sum_{i=1}^{|F_j|} \big[g_{j,i}(\kappa) - g_{j,i}(\kappa-1)\big]^2 \Bigg\} \tag{15}$$

where a_u, a_f are positive weighting factors, and $\Phi_j\big[\mathbf{g}(\kappa)\big] = \sum_{i\in F_j} g_{j,i}(\kappa) + L_j - C$, $\forall j \in J$. This criterion, excluding the last two penalty terms, attempts the minimization and balance of the links' relative occupancies $x_z/x_{z,\max}$. The first penalty term in the cost criterion allows the indirect consideration of the constraints (1), while the second penalty term is included in the cost criterion so that high-frequency oscillations of the control trajectories be suppressed. The weights a_f and a_u were adjusted via trial-and-error, striking a balance between acceptable time-variations in the optimal control trajectories and violations of constraint (1) on one hand and efficiency and fast convergence to the optimum on the other.

On the basis of the nonlinear traffic flow model (14), the constraint (2) and the cost criterion (15), a (dynamic) NOC problem is formulated over a time-horizon K, starting with the known initial state $\mathbf{x}(0)$ in the state equation (14). More specifically, the resulting NOC problem reads: minimization of the cost criterion (15) subject to (14), (2). This NOC problem may be readily solved by use of a feasible direction algorithm within few CPU-mins even for large-scale networks and long time-horizons [23]. We omit more details on the NOC numerical solution method because of space limitation.

F. Discussion

We conclude this section with some remarks pertaining to the consequences of the simplification (5) and to the application of the open-loop QPC and NOC methodologies in real time.

Let us first discuss the consequences of simplification (5). First, the updating of the control decisions cannot be effectuated more frequently than at every cycle which, however is deemed sufficient for fast network-wide real-time control reactions; on the other hand, this feature limits the real-time communication requirements between junction controllers and the central computer to one message exchange per cycle, in contrast to the second-by-second communication requirements of other signal control systems such as SCOOT [7]. Second, the model is not aware of short-term queue oscillations due to green-red switchings within a cycle, because it models a continuous (uninterrupted) average outflow from each network link (as long as there is sufficient demand). Finally, offset and cycle time have no impact within the SFM and must be either fixed or updated in real time independently [20]. These consequences of simplification (5) is the price to pay for avoiding the explicite modeling of red-green switching which would renders the resulting optimization problem discrete (combinatorial) and lead to exponential increase of computational complexity as in [9-11, 14-18].

For the application of the open-loop QPC and NOC methodologies in real time, the corresponding algorithms may be embedded in a rolling horizon (model-predictive) scheme. More precisely, the optimal control problem may be solved on-line once per cycle using the current state (current estimates of the number of vehicles in each link) of the traffic system as the initial state and predicted demand flows; the optimization yields an optimal control sequence for K cycles whereby only the first control (signal control plan) in this sequence is actually applied to the signalized junctions of the traffic network. Note that the saturation flows S_z and the turning-movement rates $t_{w,z}$, may be assumed to be time-variant and may be estimated or predicted in real time by well-known recursive estimation schemes [24]; in addition, the predicted demand flows $\mathbf{d}(k)$ may be calculated by use of historical information or suitable extrapolation methods (e.g., time series or neural networks). This rolling-horizon procedure avoids myopic control actions while embedding a dynamic optimization problem in a traffic-responsive environment.

Moreover, it should be stressed that, in contrast to LQ, in QPC and NOC methodologies the control decisions are based on the explicit minimization of a suitable cost criterion subject to all control and state constraints. Therefore, the aforementioned methodologies could be also utilized as off-line network optimization tools for calculating optimum signal control plans, since their traffic flow models (11) and (14) are incorporating all network characteristics.

Finally, the employed linear SFM (8) or (11) is simpler than the cell transmission model (CTM) employed in some previous works [15], [17], [18]; the latter calls for subdivision of network links into sorter segments (cells) and correspondingly shorter time steps T of 5 s or less. Although, the CTM may describe the (inhomogeneous) link-internal traffic state more accurately than the SFM, there is hardly any modeling difference at the junctions which are most relevant for signal control. Note that real-time application of a CTM based control strategy would call for specific measurements (or estimates) for each cell which are usually not available in current network control infrastructures. Thus, we believe that the additional complexity and measurement requirements of the CTM are probably not paying-off via more efficient resulting control decisions in a real-life setting.

IV. Application Results

To illustrate the efficiency and the real-time feasibility of the developed generic framework to the problem of urban signal control, the urban network of the city centre of Chania, Greece, is considered. For this network, we compare the closed-loop behaviour of the linear multivariable regulator with the open-loop behaviour of QPC and NOC methodologies. To ensure fair and comparable results all methodologies are evaluated by use of the same simulation model, namely the nonlinear traffic flow model (14).

A. Network and Scenario Description

The urban network of the city centre of Chania consists of 16 signalized junctions and 71 links (Fig. 2). According to the notation of Section III-A, the following sets are defined: $J = \{1, \dots, 16\}$, $Z = \{1, \dots, 71\}$. The cycle time in the network is $C = 90$ s, and $T = C$ is taken as a control interval for all strategies. For the NOC methodology we consider $T = 5$ s and $c = 0.85$ (i.e., overloaded links in (13) are considered the links z for which $x_z \geq 0.85 x_{z,\max}$).

Several tests have been conducted in order to investigate the behaviour of the three alternative methodologies for different scenarios. The scenarios were created by assuming more or less high initial queues $x_z(0)$ in the origin links of the networks while the demand flows d_z were kept equal to zero. The optimization horizon for each scenario is 450 s (5 cycles).

TABLE I

Comparison of Assessment Criteria

Strategy / Scenario	LQ TTS	LQ RQB	QPC TTS	QPC RQB	NOC TTS	NOC RQB
1	31.1	532	30.4	445	29.9	461
2	15.2	223	13.8	183	13.5	184
3	9.3	79	8.9	63	8.8	65
Average	18.6	278	17.7	230	17.4	237
Improvement	—	—	4.5%	17.2%	6.1%	14.9%

B. Comparison of Objective Functions

For each of three distinct scenarios of initial states $\mathbf{x}(0)$ and for each control approach, two evaluation criteria were calculated for comparison. The total time spent

$$\mathrm{TTS} = T \sum_{k=0}^{K} \sum_{z \in Z} x_z(k) \qquad \text{(in veh} \cdot \text{h)} \tag{16}$$

and the relative queue balance

$$\mathrm{RQB} = \sum_{k=0}^{K} \sum_{z \in Z} \frac{x_z^2(k)}{x_{z,\max}} \qquad \text{(in veh).} \tag{17}$$

Note that, as mentioned earlier, the control results of each strategy are applied to the nonlinear model (14). Eventually $x_z(k)$ over a whole cycle was calculated first as the average of the corresponding 5-s values resulting from (14), before applying the above criteria on the basis of $T = C = 90$ s.

Table I displays the obtained results. As can be seen QPC and NOC lead to a reduction of both evaluation criteria compared to LQ. More specifically, when QPC is applied, the TTS and RQB are improved by 4.5% and 17.2%, respectively; when NOC is applied, the TTS and RQB are improved 6.1% and 14.9%, respectively, compared to LQ.

NOC is seen to be superior to all other strategies in terms of the TTS. This is because the nonlinear traffic flow model used by NOC is more accurate than the linear model used by LQ or QPC (and is therefore used as a common simulator for the comparison).

Regarding the RQB, it can be seen that QPC is superior to all other strategies. On close examination, this is quite comprehensible as the RQB is the exact cost criterion considered by QPC, while, in the cost criteria considered by LQ and NOC there are partially competitive subgoals.

The average computational time per scenario for QPC and NOC is 10 s and 8 min, respectively.

C. Detailed Results

In the sequel we report on some more detailed illustrative results focussing on the particular junctions 12 and 13. These two junctions carry heavy loads, since they represent a major entrance to and exit from the city centre (see Fig. 2).

For the aforementioned scenarios, the calculated optimal state and control trajectories demonstrate the efficiency of the three alternative methods to solve the urban signal control problem. Figures 3 and 4 depict the optimal trajectories

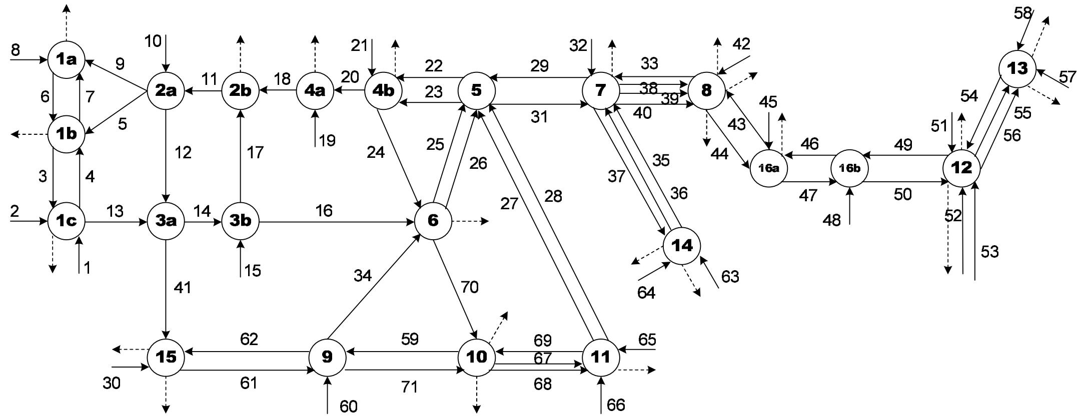


Fig. 2. The Chania urban road network.

for a particular scenario for the three methods. The main observations are summarized in the following remarks:

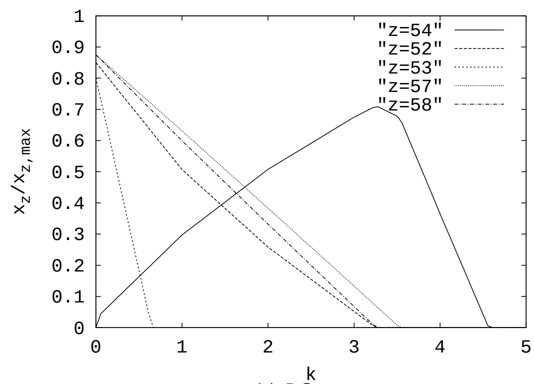
- Both QPC and NOC manage to dissolve the queues in a quite balanced way (see Figs. 3(b) and 3(c)) and thus, the desired control objective of queue balancing is achieved. Note that, these two strategies with different utilized traffic flow models accomplish the desired goal in a very similar way.
- The outflows of the origin links 57 and 58 enter the internal link 54 (solid line in Figs. 3(a)–3(c)) according to the green times of the corresponding junctions. It may be seen that QPC and NOC exhibit similar behaviour while managing particularly the queue of link 54 (see Figs. 4(b) and 4(c)).
- In contrast, the LQ strategy first allows the high initial queues to flow into the internal link 54 and then, in order to manage the developed long queue therein, it gradually increases the green time of stage 1 (see Fig. 4(a)) where link 54 has r.o.w. This somewhat slower behaviour is due to the reactive nature of the linear feedback regulator.

Both NOC and QPC deliver satisfactory results with similarly efficient control behaviour for different scenarios. Thus, taking into account that QPC needs less computational effort than NOC [23], QPC may be considered as a quite satisfactory method for the solution of the urban signal control problem and a strong competitor of LQ in terms of efficiency and real-time feasibility.

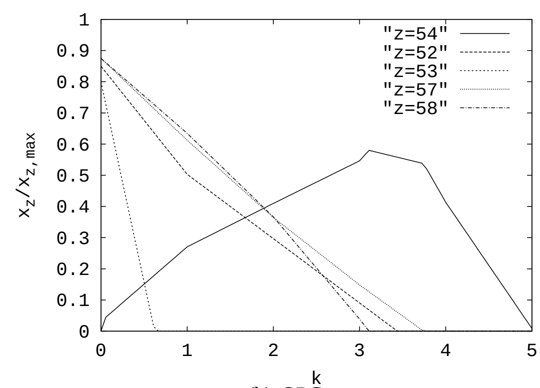
V. Conclusions And Future Work

Planning new transit routes, introducing tolls in city centres, or imposing traffic restrictions are important ingredients for combating traffic congestion in urban road networks. However, it is important to supplement these policies, with signal control techniques that contribute to the improvement of the traffic conditions via real-time decisions, particularly under saturation.

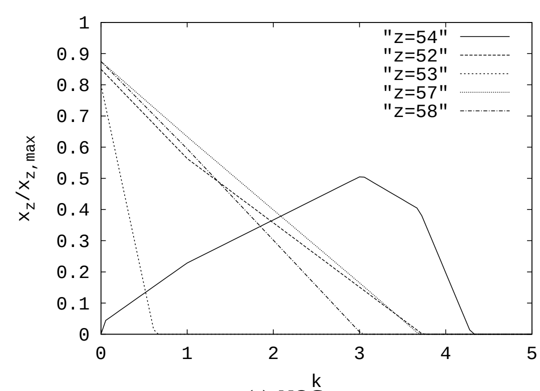
The presented generic framework develops a methodological foundation for a rational approach to traffic signal control problems, which combines store-and-forward traffic flow modeling, mathematical optimization, and optimal control. Clearly, the presented three alternative strategies, each



(a) LQ

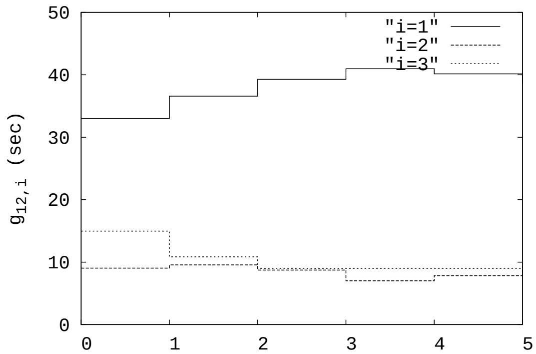


(b) QPC

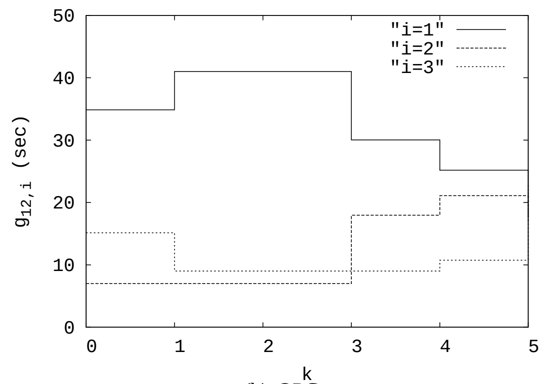


(c) NOC

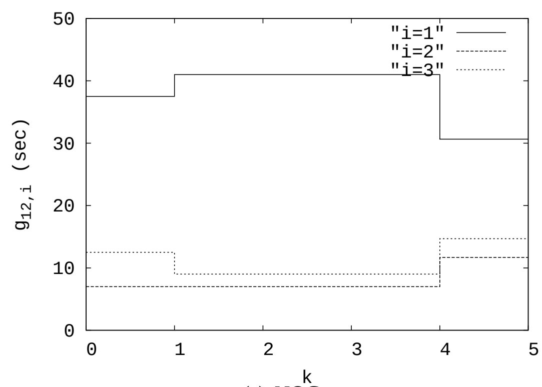
Fig. 3. Relative occupancies within the links at junctions 12, 13.



(a) LQ



(b) QPC



(c) NOC

Fig. 4. Optimal control trajectories of junction 12.

with its advantages and shortcomings, can be understood as optimal queue management tools.

Future work will deal with: (a) the comparison of the proposed open-loop QPC and NOC methodologies embedded in a rolling horizon scheme with other strategies (e.g. TUC) in more elaborated simulation involving external and internal demands and saturated traffic conditions as well as in real-life conditions and (b) improvements of the NOC strategy to cope with hard constraints (1) (rather than the use of penalty terms) through null space methods.

ACKNOWLEDGMENT

This work was partially supported by the Greek General Secretariat for Research and Technology of the Ministry of Development under the PENED-2003 project No. 03ED-898. The authors would like to thank the reviewers for their constructive comments and suggestions.

REFERENCES

[1] M. Papageorgiou, C. Diakaki, V. Dinopoulou, A. Kotsialos, and Y. Wang, "Review of road traffic control strategies," *Proc. IEEE*, vol. 91, no. 12, pp. 2043–2067, 2003.

[2] R. B. Allsop, "SIGSET: A computer program for calculating traffic capacity of signal-controlled road junctions," *Traffic Eng. Control*, vol. 12, pp. 58–60, 1971.

[3] ——, "SIGCAP: A computer program for assessing the traffic capacity of signal-controlled road junctions," *Traffic Eng. Control*, vol. 17, pp. 338–341, 1976.

[4] J. D. C. Little, M. D. Kelson, and N. H. Gartner, "MAXBAND: A program for setting signals on arteries and triangular networks," *Transport. Res. Rec.*, no. 795, pp. 40–46, 1981.

[5] N. H. Gartner, S. F. Assmann, F. Lasaga, and D. L. Hou, "A multiband approach to arterial traffic signal optimization," *Transport. Res.*, vol. 25B, pp. 55–74, 1991.

[6] K. Courage and C. E. Wallace, *TRANSYT-7F User's Guide*, Federal Highway Administration, Washigton, D.C., 1991.

[7] P. B. Hunt, D. I. Robertson, R. D. Bretherton, and M. C. Royle, "The SCOOT on-line traffic signal optimization technique," *Traffic Eng. Control*, vol. 23, pp. 190–192, 1982.

[8] P. R. Lowrie, "SCATS: The Sydney co-ordinated adaptive traffic system—Principles, methodology, algorithms," in *Proc. IEE Int. Conf. Road Traffic Signalling*, London, UK, 1982, pp. 67–70.

[9] J. L. Farges, J. J. Henry, and J. Tufal, "The PRODYN real-time traffic algorithm," in *Proc. 4th IFAC Symp. Trasport. Systems*, Baden-Baden, Germany, 1983, pp. 307–312.

[10] P. Mirchandani and L. Head, "RHODES—A real–time traffic signal control system: Architecture, algorithms, and analysis," in *TRISTAN III (Triennial Symp. Transport. Analysis)*, vol. 2, San Juan, Puerto Rico, Jun. 17–23 1998.

[11] N. H. Gartner, "OPAC: A demand-responsive strategy for traffic signal control," *Transport. Res. Rec.*, no. 906, pp. 75–84, 1983.

[12] D. C. Gazis and R. B. Potts, "The oversaturated intersection," in *Proc. 2nd Int. Symp. Traffic Theory*, London, UK, 1963, pp. 221–237.

[13] C. Diakaki, M. Papageorgiou, and K. Aboudolas, "A multivariable regulator approach to traffic-responsive network-wide signal control," *Control Eng. Prac.*, vol. 10, pp. 183–195, 2002.

[14] G. Abu-Lebdeh and R. F. Benekohal, "Development of traffic control and queue management procedures for oversaturated arterials," *Transport. Res. Rec.*, no. 1603, pp. 119–127, 1997.

[15] H. K. Lo, E. Chang, and Y. C. Chan, "Dynamic network traffic control," *Transport. Res.*, vol. 35A, no. 8, pp. 721–744, 2001.

[16] B. D. Schutter and B. D. Moor, "Optimal traffic light control for a single intersection," *Eur. J. Control*, vol. 4, no. 3, pp. 260–276, 1998.

[17] H. K. Lo, "A novel traffic signal control formulation," *Transport. Res.*, vol. 33A, no. 6, pp. 433–448, 1999.

[18] C. Beard and A. Ziliaskopoulos, "A system optimal signal optimization formulation," in *Proc. 85th TRB Annual Meeting*, Washigton, D.C., Jan. 22–26 2006.

[19] C. F. Daganzo, "The cell transmission model: A simple dynamic representation of highway traffic consistent with hydrodynamic theory," *Transport. Res.*, vol. 28B, no. 4, pp. 269–287, 1994.

[20] C. Diakaki, V. Dinopoulou, K. Aboudolas, M. Papageorgiou, E. Ben-Shabat, E. Seider, and A. Leibov, "Extensions and new applications of the traffic-responsive urban control strategy: Coordinated signal control for urban networks," *Transport. Res. Rec.*, no. 1856, pp. 202–211, 2003.

[21] M. G. Singh and H. Tamura, "Modeling and hierarchical optimization of oversaturated urban traffic networks," *Int. J. Control*, vol. 20, no. 6, pp. 913–934, 1974.

[22] M. Papageorgiou, "An integrated control approach for traffic corridors," *Transport. Res.*, vol. 3C, no. 1, pp. 19–30, 1995.

[23] K. Aboudolas, "Optimal control of traffic signals in urban road networks," (in greek), Dept. of Production Engineering and Management, Tech. Univ. of Crete, 2003.

[24] M. Cremer, "Origin-destination matrix: Dynamic estimation," in *Concise Encyclopedia of Traffic and Transportation Systems*, M. Papageorgiou, Ed. Oxford: Pergamon Press, 1991, pp. 310–315.