Online offset optimization for urban traffic network with distributed model predictive control

Yunwen Xu * Na Wu * Dewei Li * Yugeng Xi *

* Department of Automation, Shanghai Jiao Tong University, and Key Laboratory of System Control and Information Processing, Ministry of Education of China, Shanghai, 200240, China. (e-mail: willing419@sjtu.edu.cn, wunu0419@sjtu.edu.cn, dwli@sjtu.edu.cn, ygxi@sjtu.edu.cn).

Abstract:
This paper proposes a distributed control framework to optimize the offset for a path in a traffic network with arbitrary topology. Each intersection along the target path applies the model predictive control to optimize their own phase sequence and green splits with the objective of minimizing the sum of queue lengths. The first intersection on this path is regarded as the main intersection and responsible for optimizing the start green time and duration of the first phase on this path with a weighted objective according to the real-time traffic information, while the other intersections take the constraints of offset imposed by intersections ahead into consideration. The signal cycles of these intersections are fixed but allowed to be different. For computation efficiency, the nonlinear optimization problem is approximately reformulated as a mixed-integer linear programming problem. Numerical experiments on a calibrated network of Caohejing District in Shanghai indicate that our proposed method can effectively decrease delay time and waiting time especially at medium and high traffic loads.

Keywords: Traffic signal control; Online offset; Model predictive control; Distributed control framework; Mixed-integer linear programming

1. INTRODUCTION

Traffic signals were first implemented in the urban traffic network to guarantee no conflict movements at the same time. With the increasing traffic demand and limited road resources, besides the safety consideration, they are expected to improve the traffic flow and alleviate congestion with appropriate signal settings, including the signal cycle length, phase sequences, green time and intersection-to-intersection offsets.

Great efforts have been made to study the settings of signalized intersections (Taale, 2002; Aboudolas et al., 2009; Zhou et al., 2014), which come to the conclusions that: 1) signal settings play a key role in shaping traffic streams and network efficiency in general; 2) the coordination between adjacent intersections with offsets shows advantage over decentralized signal control in terms of increasing traffic throughput. In the literature, Little et al. (1981) proposed a off-line MAXBAND program for setting arterial signals to achieve maximal bandwidth, that is, the length of the time window a vehicle can travel along without stop by red lights. Ying-Ying et al. (2008) established off-line offset optimization models by considering the relationship between the arriving and departing vehicles with the objective to minimize the delay of vehicles. However, due to the random nature of traffic system, the off-line signal settings with one or several fixed numbers of offsets can not express the volatility and random of traffic flow movement.

Some researchers turn to optimize the offsets with online traffic information. In the well-known practical traffic signal control systems SCOOT (Robertson and Bretherton, 1991) and SCATS (Sims, 1979), the online offsets are adjusted with small values such that the timing plans evolve to match the measured traffic data. Gong et al. (2009) adopted the nonlinear cointegration theory and model to optimize real time green light starting time based on a series of strict applicability tests with the practical data. Gomes (2015) proposed a new formulation of bandwidth maximization problem in which a linear program in the case of pulse arrival functions is developed to reduce the computational complexity. These approaches can provide effective online offsets for arterial roads but are incapable of generalizing to networks of arbitrary topology, while in urban traffic network, it is a common situation that several paths not along the arterial roads are with high traffic loads due to the recreational or social activities.

To accommodate networks with arbitrary topology, Coogan et al. (2017) formulated the offset optimization problem as a quadratically constrained quadratic program (QCQP) with the objective of minimizing the queues at all intersec-
tions in the network, where the cycle lengths of all intersections are assumed to be the same and the intersections are undersaturated. Ouyang et al. (2019) further presented a novel algorithm to solve the above QCQP problem to near-global optimality on a large-scale by using a tree decomposition reduction to relax the nonconvex problem and using randomized rounding to recover a near-global solution. This centralized structure of offset optimization by minimizing the queues of all links puts great burden on the online computing. In addition, the assumption of the same cycle length is unreasonable for the intersections with different link length and number of lanes.

In this paper, we propose a distributed framework to optimize the offset for a path with arbitrary topology based on the model predictive methodology, where the signal cycles for intersections along this path are fixed but allowed to be different. The target path can be with large demands or the major roads taken by vehicles in the network. During the online optimization, the first intersection along the target path is regarded as the main intersection and optimizes its phase sequence and green splits with the objective of minimizing the weighted sum of queues. The following intersections independently optimize the phase sequences and green splits in their own signal cycles by considering the offset constraints from the intersections ahead. For computation efficiency, the optimization problems are approximately reformulated as mixed integer programming problem. The performances of our proposed method are evaluated via simulation in Caoheng District of Shanghai.

The organization of this paper is as follows. Section 2 provides notations for network description and formulates the model predictive control for an isolated intersection. The distributed control framework and control process to optimize the offset for a path online is developed in Section 3. Numerical results are presented and discussed in Section 4. Finally, conclusions are provided in Section 5.

2. MODEL PREDICTIVE CONTROL FOR A ISOLATED SIGNALIZED INTERSECTION

2.1 Traffic network description

Consider a traffic network composed of a number of intersections and links(roads) whose sets are denoted as $N$ and $C$, respectively. Each intersection $n \in N$ consists of several input links. Each link $l \in C$ has a number of upstream links $C_{i+1}$ and downstream links $C_i$. There are certain phases associated with each intersection, where each phase is corresponding to a connection between one input link and one output link of this intersection. As shown in Fig. 1, a connection between input link $l$ to output link $m$ is denoted as phase $(l,m)$. One or more phases that can occur simultaneously without conflict compose a stage for a intersection. The sum of the durations of all stages equals to the signal cycle length.

In the traffic network, all the vehicles are assumed to follow their predetermined paths without rerouting until arriving at their destinations. Their path information can be obtained by the vehicle-to-infrastructure(V2I) technology in real-time using several platforms (Dey et al., 2016), such as Dedicated Short Range Communications (DSRC), 4G, Wi-Fi or Bluetooth. In this paper, we present a path $p$ by an ordered collection of links or phases, where $l \in p$ denotes path $p$ goes through link $l$. $(l,m) \in p$ expresses path $p$ goes through the phase $(l,m)$.

2.2 Model predictive control for a isolated intersection

Model predictive control(MPC) is a closed-loop control methodology by solving the optimization problem online in a rolling horizon way. At each control step, based on current system information, a optimal control sequence over a finite predictive horizon is obtained, but only the first control step of the optimal solution is implemented. At next control step, the optimization problem is resolved based on new initial condition. This framework has the advantage in handling the deviation between the predicted process and actual behavior due to system-model mismatch or disturbances.

Due to the above characteristics, model predictive control has been extensively studied for intersection control(Lin et al., 2011; Zhou et al., 2014) because the traffic system is essentially random. The optimization variables in the literature are mainly the green splits of ordered phases. In this subsection, for a isolated intersection with a fixed signal cycle, besides the green splits, we simultaneously optimize its phase sequence in a rolling way.

Take intersection $n$ with four input links as shown in Fig. 1 for example. It contains a number of phases (12 phases in Fig. 1) whose set is denoted as $P_n$. Among these phases, there are several conflict points as indicated in Fig. 1, each of which corresponds to a pair of phases that can not be activated simultaneously. Write the set of conflict points in intersection $n$ as $\Theta_n$ and $[(l,m),(l',m')] \in \Theta_n$. Denote $s_{l,m}(t)$ as the signal value for phase $(l,m)$ at time step $t$ and it is a binary variable, where 1 indicates phase $(l,m)$ is activated and 0 is not. Then at time step $t$, each conflict point in $\Theta_n$ is presented by the following constraint:

$$s_{l,m}(t) + s_{l',m'}(t) \leq 1 \tag{1}$$

Considering the frequent switching between phases tends to confuse the drivers and cause potential danger in reality, in every signal cycle with length $T_n$, the green time for each phase is constrained to be consecutive by

$$\sum_{t=0}^{\tau_n-2} |s_{l,m}(t_0 + t + 1) - s_{l,m}(t_0 + t)| \leq 2(1 - s_{l,m}(t_0) \cdot s_{l,m}(t_0 + \tau_n - 1)) \tag{2}$$

Fig. 1. A typical intersection with 12 available phases.
where $\tau_n = T_n / \Delta t$ and $\Delta t$ is the time interval. Note that $T_n$ is the integer multiples of time interval $\Delta t$.

With the signal value for phase $(l, m)$, according to the store-and-forward model, the traffic flow for phase $(l, m)$ is propagated by

$$x_{l,m}(t + 1) = x_{l,m}(t) + \sigma_{l,m}(t)\lambda_l(t)\Delta t - y_{l,m}(t)$$

$$y_{l,m}(t) = \min(x_{l,m}(t), c_{l,m}(t)s_{l,m}(t))$$

where $x_{l,m}(t)$ is the number of vehicles on link $l$ heading to link $m$ at time step $t$, $y_{l,m}(t)$ denotes the number of vehicles driving from link $l$ to $m$ within time step $t$, $c_{l,m}(t)$ represents the saturation traffic flow for movement $(l, m)$ within time step $t$, $\lambda_l(t)$ is the arrival rate of link $l$ within time step $t$. There have been many approaches developed to estimate short-term traffic demand for signalised links.

3. ONLINE OFFSET OPTIMIZATION WITH DISTRIBUTED MODEL PREDICTIVE CONTROL

In this section, we develop a distributed control framework to optimize the offset for a path online based on the model predictive control of isolated intersection and

3.1 Distributed control framework

A path is generally connected by a number of intersections with continuously ordered phases. Denote the intersections on path $p$ as $\{i_1, i_2, \ldots\}$ in order and the corresponding phases as $\{(l_i, l_{i+1}), i = 1, 2, \ldots\}$. We regard the first intersection $n_1$ as the main intersection. To be specific, intersection $n_1$ optimizes its signal settings according to the optimization problem I in a rolling way with the following objective:

$$z = \sum_{t = t_0 + 1}^{t_0 + \tau_f} \delta_{i_1,i_2}x_{i_1,i_2}(t) + \sum_{t = t_0 + 1}^{t_0 + \tau_f} \sum_{n \in \Gamma(n_1)} x_{l,m}(t)$$

where $(l_1, i_2)$ is the first phase on path $p$ and $\delta_{i_1,i_2} \geq 1$. Then its green time of phase $(l_1, i_2)$ at each signal cycle and the corresponding start time are obtained. Write the green time of phase $(l_1, i_2)$ at kth cycle as $g_{l_1,i_2}(k)$ and its start time as $b_{l_1,i_2}(k)$ relative to the reference time $T_0$. Note that the reference time is the common signal cycle start time of all intersections on path $p$.

Intersection $n_1$ will send the optimized setting of phase $(l_1, i_2)$ to intersection $n_2$, i.e., $b_{l_1,i_2}(k)$ and $g_{l_1,i_2}(k)$. Accordingly, intersection $n_i$, $i = 2, 3, \ldots$ will send these information to its adjacent intersection $n_{i+1}$ but along with the travel time of link $l_i$ which is set as the offset between intersection $n_i$ and $n_{i+1}$ in this paper. The real-time travel time of link $l_i$ can be estimated by the Bureau of Public Roads (BPR) function (Riemann et al., 2015) based on the current traffic states by

$$H_i(k) = L_i/[u_i(1 - a(Q_i(k)/Q_{1i})^b)]$$

where $a, b$ are the model parameters; $u_i$ is the free-flow speed of link $l_i$, $Q_{1i}(k)$ is the total number of vehicles in link $l_i$ at start of kth signal cycle and $Q_{1i}$ is the capacity of link $l_i$.

The above one-way transmission is conducted after the signal optimization of the first intersection $n_1$ is finished at its every control cycle. The shades of red in Fig. 2 illustrate the time for information sharing, where the control cycle is equal to the signal cycle, i.e., $T_c = T_n$. Besides the first intersection $n_1$, the other intersections need to consider the offset constraints from the main intersection while optimizing signal settings with their own signal cycle. To design the green wave for path $p$, the phase $(l_i, l_{i+1}), i = 2, \ldots$, needs to be activated by intersection $n_i$.
and the green time length is \( gl_{i,l_2}(k) \). Then there exists the following relationships along the time axis of intersection \( n_i \):

\[
\alpha T_{n_i} \leq t_{l_i,l_{i+1}}^n(k) \leq t_{l_i,l_{i+1}}^n(k) + g_{l_i,l_2}(k) \leq \beta T_{n_i} \tag{8}
\]

where \( \alpha, \beta (\alpha \leq \beta) \) are two positive integers and

\[
\beta T_{n_2} - t_{l_i,l_{i+1}}^n(k) + g_{l_i,l_2}(k) < T_{n_i}.
\]

Due to different values of \( t_{l_i,l_{i+1}}^n(k), g_{l_i,l_2}(k) \) and the signal cycles, there are several cases for the signal optimization of intersection \( n_i \):

- \( \beta - \alpha = 1 \): The green time of phase \((l_i,l_{i+1})\) is inside one signal cycle of intersection \( n_i \) as illustrated in Fig. 2. Then in the model predictive control for intersection \( n_i \), the signal settings of phase \((l_i,l_{i+1})\) are constrained by

\[
s_{l_i,l_{i+1}}(i) = 1, \quad t_1, \ldots , t_2 \tag{9}
\]

where

\[
t_1 = \text{floor}\left(\frac{t_{l_i,l_{i+1}}^n(k)}{\Delta t}\right), \quad t_2 = \text{floor}\left(\frac{t_{l_i,l_{i+1}}^n(k) + g_{l_i,l_2}(k)}{\Delta t}\right)
\]

- \( \beta - \alpha = 2 \): The green time of phase \((l_i,l_{i+1})\) is across two signal cycles of intersection \( n_i \) as illustrated in Fig. 2. If the signal settings of these two signal cycles are controlled in one optimization problem, the constraints of phase \((l_i,l_{i+1})\) are the same with Eq.(9). Otherwise, For the optimization of \( \alpha \)th signal cycle, \( s_{l_i,l_{i+1}}(t) = 1 \) with \( t \in [t_1,t_3] \) where \( t_3 = (\alpha + 1)T_{n_i}/\Delta t - 1 \); For the optimization of \((\alpha + 1)\)th signal cycle, \( s_{l_i,l_{i+1}}(t) = 1 \) in time interval \( t \in [t_3 + 1, t_2] \).

- \( \beta - \alpha > 2 \): The green time of phase \((l_i,l_{i+1})\) across more than two signal cycles of intersection \( n_i \), which tends to deteriorate the performance of other phases in intersection \( n_i \). In this paper, to avoid this case, we constrain the green time of phase \((l_i,l_2)\) in first intersection \( n_1 \) by

\[
\sum_{t = \tau_{n_1}}^{(j+1)\tau_{n_1}-1} s_{l_1,l_2}(t) \leq \gamma \min(\tau_{n_1}, \tau_{n_2}, \ldots ) \tag{10}
\]

where \( \tau_{n_1} = T_{n_1}/\Delta t, \quad j = 0,1, \ldots , T_{c,n}/T_{n_1} - 1 \) and \( \gamma \in (0,1) \). These constraints can be added to the optimization problem I for the signal optimization of main intersection \( n_1 \).

3.2 Distributed control process for online offset optimization

To sum up, at the beginning of each control cycle of main intersection \( n_1 \), the optimization problem I with the added linear constraints Eq.(10) is solved by intersection \( n_1 \) to get its signal settings. After getting the solution, intersection \( n_1 \) sends the settings of phase \((l_1,l_2)\) that is going to be implemented to intersection \( n_2 \). Then intersection \( n_i \), \( i = 2,3, \ldots \), continuously sends these information along with the estimated travel time of links \( l_2,l_3, \ldots , l_i \) to next intersection \( n_{i+1} \). Besides main intersection \( n_1 \), each intersection makes a record of the green time for exact phase on the path with offset optimization. At next control cycle of main intersection \( n_1 \), the whole process will repeat again.

From the perspective of intersection \( n_i \), \( i = 2,3, \ldots \), they optimize the signal settings with their own signal cycles by solving the optimization problem I plus with the linear constraints on phase \((l_i,l_{i+1})\). The exact formulations of these constraints are determined by the values of \( t_{l_i,l_{i+1}}^n(k) \), \( g_{l_i,l_2}(k) \) and the signal cycles as discussed above, such as Eq.(9).

**Remark 1.** Since each intersection is optimized at the beginning of their own signal cycle, to make sure the optimized setting of the first phase \((l_1,l_2)\) in the main intersection \( n_1 \) can be timely caught by following intersections, the implemented signal length (the integer multiples of signal cycle) of the main intersection at its every control cycle needs to be larger than other intersections’.

3.3 Linearization of optimization problem

The above path-based dynamic offset optimization process proposes a high online computing speed requirement. However, the optimization problem I that needs to be solved by each intersection is essentially a mixed-integer programming problem with nonlinear constraints: the minimum function Eq.(4) and Eq.(2) with the absolute term, which is difficult to be efficiently solved by existing optimizers, such as Cplex and Gurobi (Atamürt and Savelsbergh, 2005). Therefore, it is necessary to reformulate the original problem I to improve the computational efficiency. We first equivalently represent Eq.(2) with several linear constraints:

\[
s_{l,m}(t_0 + i) + s_{l,m}(t_0 + i + j) - s_{l,m}(t_0 + i + 1) \leq 1 \quad (11)
\]

with \( j = 2,3, \ldots , \tau_m - 1 \), \( i = 0, \ldots , \tau_m - j - 1 \). For clarity, the proof about the equivalence between Eq.(2) and Eq.(11) is omitted here. The minimum function Eq.(4) is relaxed by a set of linear inequalities together with the item to maximize \( y_{l,m}(t) \) in the objective. Then the original optimization problem I can be reformulated as a mixed-integer linear programming problem for computation efficiency as follows:

\[
\begin{align*}
\min \quad & z = \sum_{t = t_0}^{t_{0}+f} \sum_{\forall(l,m)} x_{l,m}(t) + \sum_{t = t_0}^{t_{0}+f} \sum_{\forall(l,m)} y_{l,m}(t) \\
\text{s.t.} \quad & \text{Relaxations of flow propagation constraints} : \\
& \forall(l,m) \in P_n, t = t_0, \ldots , t_0 + f - 1 : \\
& y_{l,m}(t) \leq x_{l,m}(t) \\
& y_{l,m}(t) \leq c_{l,m}(t) s_{l,m}(t) \\
& \text{Signal conflict points constraints} : \\
& \forall(l,m), (l',m') \in \Theta_n, t = t_0, \ldots , t_0 + f - 1 : \\
& \text{Eq.(1)}; \\
& \text{Consecutive green time constraints} : \\
& \forall(l,m) \in P_n, t = t_0 + i \tau_n, i = 0,1, \ldots , T_{c,n}/T_{n} : \\
& s_{l,m}(t + i_1) + s_{l,m}(t + i_1 + i_2) - s_{l,m}(t + i_1 + 1) \\
& \leq 1, \quad i_2 = 2, \ldots , \tau_n - 1, \quad i_1 = 0, \ldots , \tau_n - i_2 - 1 \\
& \text{Signal settings constraints from } T_{c,n} \text{ to } T_{p,n} : \\
& \forall(l,m) \in P_n, \\
& i = 0, \ldots , T_{p,n} - T_{c,n}, j = 0, \ldots , \tau_n : \\
& s_{l,m}(t'_j - \tau_n) = s_{l,m}(t'_j + i \tau_n + j)
\end{align*}
\]
4. NUMERICAL STUDY

4.1 Simulation settings

The simulation using the SUMO traffic simulator is performed by a road network from Caohejing District in Shanghai, which is downloaded from the OpenStreetMap. Its simplified network is shown in Fig. 3, containing 38 signalized intersections and 519 links. The red bold links in Fig. 3 denote the origins and destinations which are the entrances or exits of the parking lots, exhibition center or the residential areas in reality. With these origins and destinations, we consider 30 paths in the simulation. Each path is allocated with the low (400 veh/h), medium (700 veh/h) and high (1000 veh/h) traffic loads for performance study, respectively. According to the path distribution, Yishan Road and Caobao Road illustrated by the bold blue lines in Fig. 3 are with the relatively high traffic loads.

In the simulation, we choose fixed-time control (denoted as FTC) as the benchmark, where the possible phases of each signalized intersection are activated in a predetermined periodic way, similar to what is used in common practice. Their cycle length and green splits are determined based on Webster’s formula (Webster, 1958) using the average traffic flow. In the simulation their cycle lengths are the integer multiples of time interval $\Delta t$ (3s) and from 60 seconds to 120 seconds. To make fair comparison, we test our method by comparing the following three methods. Note that only the intersections on the Caobao Road and Yishan Road are controlled and other intersections are under FTC with the same signal settings in each method. All the simulations are performed for one hour.

- Actuated signal control (Taale, 2002) (denoted as ASC): ASC takes into account the actual traffic demand and the presence of vehicles to determine the changes of the traffic signals, where their stage settings and phase sequences are the same with FTC. The presence of vehicles on each lane is detected every 10 seconds by the detectors at the stoplines. The maximum green time for each phase is set to be 40 seconds.

- Distributed signal control with online offset optimization (denoted as DSCO): DSCO is our proposed method as described in Subsection 3.1 and 3.2. The signal cycle length $T_c$ of each controlled intersection is preset and the same with FTC. The control cycle $T_{c,n}$ is equal to the prediction horizon $T_{p,n}$. The weight $\delta_{1,2}$ in the optimization objective for the first intersections on above two controlled roads are set as 2. The parameter $\gamma$ in the green time length constraints is set as 0.6.

4.2 Performance study

In Fig. 4 and Fig. 5, we plot the average delay time and waiting time per vehicle under different control methods with low, medium and high traffic loads, respectively. Here the delay time essentially equals to the real travel time minus the ideal travel time which is the time a vehicle travels in free flow speed without signal control. The waiting time is the time of a vehicle whose speed is below 0.1 m/s. The average values in Fig. 4 and Fig. 5 are taken over both the simulation time and number of arriving vehicles.

At low loads, it can be easily seen from Fig. 4 and Fig. 5 that ASC provides the shortest delay time and waiting time, and FTC gets the longest. This is due to the fact that under ASC, the traffic states are detected every 10 seconds and the phases with no vehicles are allowed to be skipped, which is of common occurrence at low loads, while other methods make decisions of signal settings every signal cycle ($\geq 60s$). FTC actives the phases with preset sequence and green time, incapable of considering the random arrivals, which tends to allocate green time to the phase with empty queues. DSCO and DSC also allow the phases to be skipped in the optimization problem, but the deviation between the predicted states and actual behavior makes the signal settings not timely as ASC. Comparing DSCO and DSC, we can find that at low loads, DSC outperforms DSCO in terms of delay time but the performance of waiting time is inverse. This
can be explained by that the online offsets optimized in DSCO increase the throughput along Yishan Road and Caobao Road, and decrease the waiting time. However, the optimized offset is very likely not to be fully occupied at low traffic loads, which further affects the performance of other phases in the intersections behind on these two paths due to the imposed restrictions.

At medium traffic loads, DSCO yields the best performance in terms of both delay time and waiting time, followed by DSC, ASC and FTC. Compared with FTC and ASC which make the signal timing plans only based on current measured traffic states, DSC optimizes the signal settings using the model predictive control by taking the traffic loads into consideration. DSCO further coordinates the intersections along two main roads to dynamically generate the green wave and improve the network throughput.

At high traffic loads, the performance of different control methods have the same tendency with the medium traffic loads, except the comparison between FTC and ASC, where FTC provides better performance than ASC. This is because the advantage of ASC explained above is gradually weakening as traffic loads increase since the possibility of being idling will be greatly decreased at high traffic loads. Under ASC, the stages with large traffic loads tend to be continuously allocated the green time until the maximum green time, making the vehicles of other stages keep waiting. Another interesting observation from Fig. 4 and Fig. 5 is that the advantage of DSCO at medium traffic loads is more obvious than at high traffic loads in terms of delay and waiting time. This is possible because at high traffic loads, the vehicles are likely to be stuck due to traffic jam, making the designed offset between two adjacent intersections hard to be realized in reality, while these situations are of rare occurrences at medium loads.

5. CONCLUSION

This paper proposes a distributed framework to optimize the offset for a path with arbitrary topology based on the model predictive methodology, where the signal cycles for intersections along this path are fixed but allowed to be different. During the online control, each intersection separately optimizes the phase sequence and green splits in their own signal cycles with specific objective and constraints, where the first intersection along the target path is regarded as the main intersection with the weighted objective and other intersections with the offset constraints from the intersections ahead. The simulation results on two main roads in Caohéjing District of Shanghai show the advantage of our proposed method over other compared methods in terms of delay time and waiting time, especially at medium and high traffic loads.

REFERENCES