Heuristic Algorithms for Online Traffic Signal Control with Cell Transmission Models

Qing He
IBM T.J. Watson Research Center
P.O. Box 218, Yorktown Heights, NY 10598
Email: qhe@us.ibm.com
Tel: +1 (914) 945-1142

Wei-Hua Lin
Department of Systems and Industrial Engineering
University of Arizona, Tucson, AZ 85721
Email: weilin@sie.arizona.edu
Tel: +1(520)621-6553

Hongchao Liu
Department of Civil Engineering
Texas Tech University, Lubbock, TX 79409
Email: hongchao.liu@ttu.edu
Tel: +1(806)742-3523 ext. 229

K. Larry Head
Department of Systems and Industrial Engineering
University of Arizona, Tucson, AZ 85721
Email: larry@sie.arizona.edu
Tel: +1(520)621-2264

Submitted to TRB 91th Annual Meeting, January 2010, Washington, D.C.

August 1, 2011

Word Count: 6160
Abstract and Manuscript Text: 4410
Number of Tables and Figures: 7 (= 1750 words)

1 Corresponding author, fax: +1(914)-945-3434
Abstract
Traffic signal control with an exact traffic macroscopic model, such as cell transmission model, can capture traffic queue dynamics and easily tackle oversaturated condition in the real world. However, these control problems usually are very expensive to obtain exact optimal solutions for online signal control. In this paper, three heuristic solution algorithms, (the Dive-and-fix method, the Ratio-cluster method, and the Cumulative-departure method) are specially designed to solve the traffic signal control problem formulated as a 0-1 mixed-integer linear programming problem with cell transmission model. These three solution algorithms are based on two fundamental approaches. First, the 0-1 mixed-integer linear program is solved via linear relaxation (LR). Second, the non-integer solutions obtained from the LR are converted into the integer solutions by taking advantage of the underlying physical mechanism embedded in the LR solutions that lead to the optimal signal control. In particular, proportional capacities for different approaches and the cumulative exit flow at each intersection obtained from the LR solutions are utilized to determine green time allocation for each approach. It is demonstrated that the near-optimal solutions obtained with these algorithms are very close to the optimal solutions under both uncongested and congested traffic conditions.

1. Introduction
The optimal traffic signal control problems can be solved either by analytical models or by heuristics approaches. With the availability of a vast amount of high resolution and high quality traffic data (e.g. collected every few seconds), the use of optimal control formulations with exact models has become increasingly important compared with heuristics approaches [1]. Most traditional analytical formulations incorporate oversimplified traffic flow models that do not consider physical queues. As a result, these formulations do not properly address traffic flow characteristics, such as shock waves, queue spillback, and other traffic flow dynamics. The literature usually deals with unsaturated ([2]-[6]) and oversaturated ([7]-[14]) traffic conditions with significantly different formulations or strategies since the fundamental traffic characteristics change dramatically from unsaturated to oversaturated conditions. Traditional models generally include some simplifying assumptions that would limit their applicability to oversaturated conditions.

Recently a number of papers have developed dynamic traffic signal control formulations based on the cell transmission model (CTM). The CTM provides a convergent approximation to the Lighthill, Whitham and Richards (LWR) model ([15][16]) and accommodates the entire range of flow-density relationships ([17][18]). The CTM-based signal control formulation can address both unsaturated and oversaturated conditions considering shockwaves and physical queues. Lo ([19][20]) formulated the network signal optimization problem as a mixed-integer linear programming problem using CTM, assuming that the cycle lengths are fixed. Lin and Wang ([21]) formulated a more computationally efficient version of the mixed integer linear program for the signal optimization problem with CTM. Beard and Ziliaskopoulos ([22]) proposed a CTM-based system optimal signal optimization formulation combined with system optimal traffic assignment which provides several improvements over existing mixed-integer linear program formulations including turning movements for exclusive turn lanes. Most recently, Zhang et al ([23]) examined the design of robust traffic signal control with the CTM. A scenario-based stochastic programming model was proposed to optimize the timing of pre-timed signals along arterials under day-to-day demand variations.
For the CTM-based formulations with the mixed-integer linear programming approach, the problem size can grow very quickly with the size of the network and the time horizon. The “curse of dimensionality” makes it impossible to solve these formulations directly using commercially available packages such as CPLEX, LINDO, etc. Recently ([9][11][12][20][23]), Genetic Algorithm (GA) has been the most common algorithm to achieve near optimal solutions. However, there is no guarantee that GA will converge to local optimal. Without knowing the lower bounds through GA, it is not possible to obtain estimates of the optimality gaps or evaluate the quality of the solutions. The purpose of this paper is to develop heuristic algorithms that solve the CTM-based traffic signal control models. The algorithms developed use a linear relaxation method and then convert the solutions into feasible integer solutions. Computational issues are addressed for solving these problems for real-time applications.

Three approaches are discussed to obtain optimal solutions from traffic signal models formulated with the mixed integer linear programming approach. The first approach, the “Dive-and-fix” method is based on a traditional heuristic algorithm for binary formulations. The second approach makes use of average green time split ratio in each cluster identified by the linear relaxation solutions. The fractional green time split ratios at any given time step are mapped into signal phases to obtain mixed-integer solutions. The third approach is based on cumulative vehicle departures. It strives to ensure that the cumulative departure at each intersection for the optimal integer solutions match well with the one generated from the linear relaxation solutions.

2. Overview of 0-1 MIP signal control formulation with CTM

The 0-1 mixed integer model for the signal control problem considered in this paper was developed by Lin and Wang ([21]). The model is based on the cell transmission model (CTM) [17][18], which is a discrete version of the first-order hydrodynamic theory of traffic flow, i.e., the LWR model [15][16], expressed as follows,

$$ \frac{\partial k}{\partial t} + \frac{\partial q}{\partial x} = 0 \quad \text{and} \quad q = F(k,x,t) \quad (1) $$

where $k$ and $q$ denote traffic density and flow, respectively, which may vary across location $x$ and time $t$. The first equation in (1) is the flow conservation equation and the second equation is the function representing the relationship between flow and density. In the cell-transmission model, the flow density relationship is assumed to be of the form:

$$ q = \min \{ q_{\max} , W(k_{\text{jam}} - k) \} \quad (2) $$

where $k_{\text{jam}}$ is the jam density; $q_{\max}$ is the flow capacity; $V$ is the free flow speed and $W$ is the backward wave speed. By discretizing the network into cells with the cell length equal to the distance traveled by vehicles traveling at free flow speed in a single time step, the LWR model can then be approximated by a set of state-space equations,

$$ n_i(t+1) = n_i(t) + y_{i+1}(t+1) - y_i(t) \quad (3) $$

$$ y_i(t) = \min \{ n_i(t), Q \frac{W}{V} \lfloor N - n_{i+1}(t) \rfloor \} \quad (4) $$

where $i$ refers to cell $i$, and $i+1$ ($i-1$) represents the cell downstream (upstream) of cell $i$. The variables $n_i(t)$, $y_i(t)$, $Q$ and $N$ denote the number of vehicles, the actual outflow, the maximal flow and the maximum occupancy of a cell, respectively.

The reason for choosing the Lin and Wang ([21]) model is implementation simplicity and accuracy compared with other models in the same category. An example of a two-intersection one-way roadway segment represented by the model is given in Fig. 1(a). The detail of the 0-1 mixed integer model is shown below ([21]),

Sets
O = set of origin cells with only output links.
D = set of destination cells with only input links.
E = set of ordinary cells with both input and output links.
I = set of intersection cells with input and constrained output links.

Exogenous variables:
n_{i, t} = number of vehicles in cell i at time t.
y_{i, t} = number of vehicles leaving cell i at time t.
q_{i, t} = actual capacity in vehicles for intersection cell i in time interval \([t, t+1]\).
w^{(i,\cdot)}_t = 0-1 variables for traffic signal at intersection \((i, j) \in I\) in time interval \([t, t+1]\). If \(w^{(i,\cdot)}_t = 1\), approach i is green, otherwise red.
v^{(i,\cdot)}_t = dummy variables for maintaining minimum green time

Fig. 1 Traffic signal control with the CTM model
Endogenous variables:

\( D_{i,t} \) = number of vehicles entering the origin \( i \) in \([t, t+1]\).

\( Q_i \) = capacity, maximum number of vehicles that can enter or exit cell \( i \) in \([t, t+1]\).

\( N_i \) = jam density, maximum number of vehicles that can reside in cell \( i \) in \([t, t+1]\).

\( C \) = cycle length in number of time steps.

\( T \) = the total number of time steps,

\( \omega \) = wave coefficient (a dimensionless unit representing the ratio of wave speed over free-flow speed).

\( GMin \) = minimum green (in time steps, a constant number assumed).

\( GMax \) = maximum green (in time steps, a constant number assumed).

\( \alpha \) = coefficient for delay.

\( \beta \) = coefficient for early arrival flow to avoid vehicle holding at a cell even though there is capacity available downstream.

The objective function can be stated as the minimization of a weighted linear combination of delay and early arrival flow.

\[
\text{Min } \alpha \sum_{i \in D} \sum_{t} ty_{i,t} + \beta \sum_{i \in D} \sum_{t} ty_{i,t}
\]

The constraints for Ordinary Cells are

\[
y_{i,t} \leq n_{i,t} \quad \forall t
\]

\[
y_{i,t} \leq \omega (N_i - n_{i,t-1}) \quad \forall t
\]

\[
y_{i,t} \leq Q_i \quad \forall t
\]

\[
n_{i,t+1} = n_{i,t} + y_{i,t-1} - y_{i,t} \quad \forall t
\]

The constraints for Origin Cells are

\[
y_{i,t} \leq n_{i,t} \quad \forall t
\]

\[
y_{i,t} \leq \omega (N_i - n_{i,t-1}) \quad \forall t
\]

\[
y_{i,t} \leq Q_i \quad \forall t
\]

\[
n_{i,t+1} = n_{i,t} + D_{i,t} - y_{i,t} \quad \forall t
\]

The constraints for Destination Cells are

\[
y_{i,t} = n_{i,t} \quad \forall t
\]

\[
n_{i,t+1} = n_{i,t} + y_{i,t-1} - y_{i,t} \quad \forall t
\]

The constraints for Intersection Cell \((i,j)\) are

\[
y_{i,t} \leq n_{i,t} \quad \forall t
\]

\[
y_{j,t} \leq \omega (N_j - n_{j,t-1}) \quad \forall t
\]

\[
y_{j,t} \leq q_{i,t} \quad \forall t
\]

\[
n_{i,t+1} = n_{i,t} + y_{i,t-1} - y_{i,t} \quad \forall t
\]

\[
y_{j,t} \leq n_{j,t} \quad \forall t
\]

\[
y_{j,t} \leq \omega (N_j - n_{j,t+1}) \quad \forall t
\]

\[
y_{j,t} \leq q_{j,t} \quad \forall t
\]

\[
n_{j,t+1} = n_{j,t} + y_{j,t-1} - y_{j,t} \quad \forall t
\]
The intersection capacities for approach \( q_{i,t} = w_t^{(i,j)} Q_i \) \( \forall t \) and \( q_{j,t} = (1 - w_t^{(i,j)}) Q_i \) \( \forall t \). The constraints to maintain the minimum green time

\[
\sum_{t=1}^{GMin} |q_{i,t} - q_{i,t-1}| \leq Q_i \quad \forall t, i \in I
\]

(25)

Which can be transformed into a set of linear constraints using a dummy variable \( v_t^{(i,j)} \)

\[
q_{i,t} - q_{i,t-1} \leq v_t^{(i,j)} \quad \forall t, k \in I
\]

(26)

\[
q_{i,t} + q_{i,t-1} \leq v_t^{(i,j)} \quad \forall t, k \in I
\]

(27)

\[
\sum_{t=1}^{GMin} v_t^{(i,j)} \leq Q_i \quad \forall t \in [0, T - GMin], (i, j) \in I
\]

(28)

The constraints to maintain the maximum green time are

\[
\sum_{t=1}^{GMax} q_{i,t} \leq Q_i GMax \quad \forall t \in [0, T - GMin], (i, j) \in I
\]

(29)

\[
\sum_{t=1}^{GMax} q_{j,t} \leq Q_j GMax \quad \forall t \in [0, T - GMin], (i, j) \in I
\]

(30)

The constraints to ensure flow conservation at the network level-clear the vehicles through all the cells in time \( T \)

\[
\sum_{i} \sum_{j} y_{i,j} = \sum_{j} \sum_{i} D_{j,i} \quad \forall t, \forall i \in D, \forall j \in O
\]

(31)

For a particular intersection \( (i, j) \) as shown in Fig. 1(b), the traffic signal control state at time step \( t \) is modeled with a binary variable, \( w_t^{(i,j)} \). The intersection capacities for approaches \( i \) and \( j \) are represented by \( q_{i,t} = Q_i w_t^{(i,j)} \) and \( q_{j,t} = Q_j (1 - w_t^{(i,j)}) \), respectively. Here \( w_t^{(i,j)} \) is the only integer variable in the MIP formulation. In the LR formulation below, the integrality constraints for \( w_t^{(i,j)} \) are replaced by \( 0 \leq w_t^{(i,j)} \leq 1 \).

An instance of Fig. 1(a) was solved as an MIP and using the LR respectively over a range of total time steps. The computation times and the objective function values for MIP and LR are shown in Table 1. With the change in total number of time steps from 40 to 200, the computation times for MIP changes drastically from 7 seconds to about 3 hours. The computation times for LR remain within 7 seconds. One can also see that the difference between the total delay generated by MIP and LP is very small in all cases. Though the LR solution of the original MIP usually is not feasible, it provides a lower bound to the original MIP problem and generates traffic states that can be used to reconstruct near-optimal solutions to the MIP problem.

<table>
<thead>
<tr>
<th>Steps</th>
<th>MIP (time)</th>
<th>LR (time)</th>
<th>MIP (delay)</th>
<th>LR (delay)</th>
</tr>
</thead>
<tbody>
<tr>
<td>40</td>
<td>7.1</td>
<td>0.53</td>
<td>7351.29</td>
<td>7161.6</td>
</tr>
<tr>
<td>80</td>
<td>440.2</td>
<td>1.5</td>
<td>15492.7</td>
<td>15184.6</td>
</tr>
<tr>
<td>120</td>
<td>1687</td>
<td>2.9</td>
<td>20782</td>
<td>20407.3</td>
</tr>
<tr>
<td>200</td>
<td>10650</td>
<td>6.4</td>
<td>39055.8</td>
<td>38454.5</td>
</tr>
</tbody>
</table>

Table 1 Comparisons of computation times with MIP & LR
3. Proposed approaches

3.1 Approach I — the Dive-and-fix method

The first algorithm developed is based on a generalized heuristic algorithm, the Dive-and-fix method (Wolsey [24]). The basic idea of this approach is to first solve the LR problem and round a non-integer variable in the range of [0, 1] to the nearest integer, either 0 or 1. This variable is then treated as known data (fixed constant). A new LR with the known data is then solved and another non-integer variable is rounded to the nearest integer. This process is repeated until there are no remaining non-integer decision variables. The LR formulation becomes smaller and smaller and is solved faster and faster. The Dive-and-fix algorithm is:

Step 0 (Initialization): Solve the LR formulation; get a linear relaxation solution \( (w^*_t, 1 - w^*_t) \) \( t = 1 \ldots T \) for each intersection pair.

Step 1: let \( F = \{ t: w^*_t \not\in \{0,1\} \} \) be the set of 0-1 variables that are fractional.

Step 2: Find the position of the most integer-like solution.

- If \( F = \emptyset \), stop.
- Else if \( F \neq \emptyset \), let \( t^* = \arg \min_{t \in F} \min(w^*_t, 1 - w^*_t) \).

Step 3: Fix the solution to be data, 0 or 1.

- If \( w^*_t < 0.5 \), fix \( w^*_t = 0 \) and \( (1 - w^*_t) = 1 \).
- Else fix \( w^*_t = 1 \) and \( (1 - w^*_t) = 0 \).

Step 4: Solve the new LR from step 3. Go to step 1.

An example is shown in Fig. 2. At iteration 1, given the LR solution, the most integer-like solution is \( (0.98, 0.02) \). This solution is fixed to \( (1, 0) \) as input for iteration 2. At iteration 2, the LR formulation is solved and the most integer-like solution is \( (0.1, 0.9) \). This solution is fixed to \( (0, 1) \). The same occurs at iterations 3, 4, 5 and so on. If only one fractional solution exists at iteration \( k \), it will be fixed to binary and the algorithm stops.

The traditional Dive-and-fix method can provide approximate integer solutions. However, the constraints may be violated for preserving maximum/minimum green. It is necessary to adjust the solutions to make them feasible.

3.2 Approach II — The Ratio-cluster method

The LR method can be thought to treat a signalized intersection as if it were a merge junction in that fractional traffic flows are mixed at the junction. Although the solution from the LR is infeasible for
signal control in general, it provides moment-by-moment information about green time allocation in an optimal capacity allocation fashion, similar to the result in Liu et al. ([14]). The capacity allocation can be interpreted as a green time allocation. Suppose the solution obtained for the two different approaches is $w_i^j = 0.8$ and $w_i^0 = 0.2$ for intersection cell (3,10) at time step $t_i$, as shown in Fig. 1. For a given cycle length, the solution suggests that the green time for the main street should be 80% ($= w_i^3$) of the cycle length, whereas the green time for the side street should be 20% ($= w_i^0$) of the cycle length. Clearly, the green time ratio obtained with the LR method can be utilized to generate a feasible MIP solution.

The idea of the ratio-cluster approach is to group time steps into several clusters with relatively stationary green time allocations in each cluster. The average green/red time ratio within a cluster is used to determine the cycle length and generate the integer green time allocation. The average green/red time ratio is approximated by using available MIP green/red time ratios, e.g. $G_{\text{min}}=1$ step, $G_{\text{max}}=3$ steps in Table 2 when each time step represent 10 seconds. Detailed steps are described as follows.

<table>
<thead>
<tr>
<th>Cycle length</th>
<th>Available Ratio 1</th>
<th>Available Ratio 2</th>
<th>Available Ratio 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1:1 = 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1:2 = 0.5</td>
<td>2:1 = 2</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>2:3 = 0.33</td>
<td>3:2 = 1.5</td>
<td>3:1 = 3</td>
</tr>
<tr>
<td>5</td>
<td>1:3 = 0.33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1:1 = 1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Step 0 (initialization): Given the LR solutions $(w_j^i, 1 - w_j^i)$, only consider the green time allocation $w_j^i$ for main street. An exponential weighed moving average (EWMA) ([25]) time series $z_i$ is generated to track the shifts in $w_j^i$:

\[ z_i = \lambda w_j^i + (1-\lambda)z_i-1, \quad t > t_0 \]

\[ z_{t_0} = w_{t_0}^i, \]

Where $0 < \lambda \leq 1$ is a constant which defines the weight assigned to the current sample compared to the historical means. A control limit $L_i$ for the first cluster is defined as $L_i \equiv [z_{t_0} - \sigma, z_{t_0} + \sigma]$, where $\sigma$ is a constant that defines the variation ($0 < \sigma \leq 1$). Let $t = t_0$ and $k = 1$, where $t$ is the index of time steps and $k$ is the index of clusters.

Step 1 (clustering): Update $t = t + 1$.

Step 1.1

If $t = T$, where $T$ is the total time steps.

Let $y_i = \sum_{r=i-k+1}^{i} w_j^r / (T_i - \sum_{r=i-k+1}^{i} w_j^r)$ be the average green/red time ratio for approach $i$ in cluster $k$, where

\[ T_k = t_k - t_{k-1}, \]

$N = k$, where $N$ is the total number of clusters found.

Go to step 2.

End

Step 1.2

If $z_i \in L_i$, go to step 1 again.

Step 1.3

Else if $z_i \notin L_i$,

Let $t_k = t$ be the first time step in the new cluster $k + 1$.
He, Lin, Liu and Head

1. If $T_k < C_{min}$ ($C_{min}$ is the available minimal cycle length), go to step 1.
2. Else, a new cluster starts,
3. Let $\gamma_k = \sum_{t=k-1}^{T_k} w^*/(T_k - \sum_{t=k-1}^{T_k} w^*)$ be the average green/red time ratio in cluster $k$.
4. Update the control limit $L_{k+1} = [z_{\gamma_k} \sigma, z_{\gamma_k} + \sigma]$ for new cluster $k+1$.
5. Update the index of cluster $k = k + 1$.
6. Go to step 1.
7. End.
8. End.
9. Step 2 (binary solution approximation):
10. Let $k = 1$ and $E_k = 0$, where $E_k$ is cumulative green time error between LR solution and Ratio-cluster outputs for the main street.
11. While $k \leq N$
12. Step 2.1
13. Select a cycle $(G_k, R_k) \in C$, where $(G_k, R_k) = \arg \min_{(G,R) \in C} \frac{G_k}{R_k} - \gamma_k$, where $G_k$ and $R_k$ are the green time and red time in one cycle, which has the closest green/red ratio to $\gamma_k$, $C$ is the set of all of the possible cycles.
14. Let $n_k = \left\lfloor \frac{T_k}{G_k + R_k} \right\rfloor$ be the integer rounded down from dividing $T_k$ by $(G_k + R_k)$, and $m_k = T_k - (G_k + R_k) * n_k$ be the remainder.
15. Step 2.2
16. Generate approximate the Ratio-cluster integer solution $w^*_{t_k}$ from $t_{k-1}$ to $t_k - 1$ with binary 0 or 1 by $n_k$ cycles of $(G_k, R_k)$ in cluster $k$ until the rest of time steps are equal to $m_k$.
17. Step 2.3
18. Select a cycle $(G'_k, R'_k) \in C'$, or a half cycle (either $(G'_k, R'_k) = (m_k, 0)$ or $(G'_k, R'_k) = (0, m_k)$) if $m_k \in [G_{min}, G_{max}]$ to fill the rest $m_k$ time steps, where $(G'_k, R'_k) = \arg \min_{(G,R) \in C'} |E_k|$, $E_k = \sum_{t=k-1}^{T_k} (w^*_t - w^*_t) + E_{k-1}$ and $C' = \{(G,R) | G + R = m_k, (G,R) \in C\}$.
19. Step 2.4
20. Update $k = k + 1$ and $E_k = \sum_{t=k-1}^{T_k} (w^*_t - w^*_t) + E_{k-1}$
21. End.
22. The Ratio-cluster algorithm tracks the changes of green time of the LR solution and divides total time steps into several clusters via EWMA data, a concept used in quality control charts. Suitable integer cycles are then inserted into each cluster to match the average green ratio and minimize the cumulative green time error.

3.3 Approach III — The Cumulative-departure method

In this approach, the throughput of an intersection, or the cumulative departures downstream of an intersection, generated from the LR solution is employed to obtain an MIP solution.
As the cumulative departure curve taken at each approach downstream of an intersection is in fact determined by the intersection discharge rate, controlled by a specific signal timing plan and the vehicle arrival flow. The optimal departure curve must be the result of the optimal signal setting. Given that fact, the smooth cumulative departure curve from LR solutions is the optimal departure curve, since LR gives a lower bound on the objective function. In the cumulative-departure method, integer solutions are produced by matching the vehicle discharge rate in each direction with the cumulative departure curve from LR. Detailed steps are as follows (for each intersection):

Step 1: Given the LR solution \((w', 1 - w')\), generate the cumulative vehicle departure series \(D_{LR}(t)\) and \(D'_{LR}(t)\) for the main street and the side street respectively immediately downstream the intersection. Let \(t = 1\).

Step 2:

While \(t \leq T\), where \(T\) is the total time steps.

Determine the binary solution \(w_i^{ca}\) whether 0 or 1 without violating the \(G_{max} / G_{min}\) constraints, and

\[
{w_i^{ca} = \arg \min_{\{0,1\}} |D_{ca}(t) - D_{LR}(t)| + |D'_{ca}(t) - D'_{LR}(t)|,}
\]

where \(D_{ca}(t)\) is the cumulative vehicle departure series guaranteed by \(w_i^{ca}\) for main street, \(D'_{ca}(t)\) for side street.

Update \(t = t + 1\).

End

An example for a single intersection is shown in Figure 3. The solid curves represent \(D_{LR}(t)\) (left) and \(D'_{LR}(t)\) (right), while the dashed curves are \(D_{ca}(t)\) and \(D'_{ca}(t)\). The dashed curves appear to have a slight lag compared with the solid curves due to the solid curves always producing the lower bound (higher vehicle throughputs). The goal of this method is to generate a binary-solution based curve that best matches the LR curves, or to minimize the difference between the lower bound and the feasible binary solution.

Fig. 3 An example of the cumulative departure method (a) CTM based one-intersection layout; (b) Cumulative departure curve match with LR solutions in Approach III
4. Numerical examples

Four numerical examples are presented which are solved using CPLEX 10.1 on a Sun workstation with 2 * 900 MHz CPU and 4G memory. Examples 1 and 2 consist of only one intersection. In example 1, the performances of each approach are shown in Table 3(a). It takes 13 minutes (821 seconds) to get the optimal MIP solution, whereas it only takes 1.4 seconds to obtain the LR solution. Of the three approaches, the Dive-and-fix method takes the most time (32 seconds). Its relative error is also the largest, nearly 5%. The best approach in this case is the Cumulative-departure method with 0.68% relative error and less than 0.1 second computation times. The result of example 2, with more congested traffic condition, is shown in Table 3(b). The computation time of the MIP formulation went up to 23 minutes, while the computation time for the LR method is still 1.5 seconds. The Ratio-cluster method seems to outperform other approaches in example 2 with only 0.58% relative error and 0.5 seconds computation time.

Table 3 Comparison of different approaches

<table>
<thead>
<tr>
<th></th>
<th>Objective Value</th>
<th>Solution Time (sec)</th>
<th>Relative error (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MIP (CPLEX)</td>
<td>18262.4</td>
<td>821</td>
<td>0</td>
</tr>
<tr>
<td>LR</td>
<td>18047.5</td>
<td>1.4</td>
<td>--</td>
</tr>
<tr>
<td>Dive-and-fix</td>
<td>19153</td>
<td>32</td>
<td>4.88</td>
</tr>
<tr>
<td>Ratio cluster</td>
<td>18501</td>
<td>0.0625</td>
<td>1.31</td>
</tr>
<tr>
<td>Cum. departure</td>
<td>18386</td>
<td>0.0768</td>
<td>0.68</td>
</tr>
</tbody>
</table>

(b) Example 2 (congested traffic)

<table>
<thead>
<tr>
<th></th>
<th>Objective</th>
<th>Solution Time (sec)</th>
<th>Relative error (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MIP (CPLEX)</td>
<td>21596.6</td>
<td>1405</td>
<td>0</td>
</tr>
<tr>
<td>LR</td>
<td>21397.1</td>
<td>1.5</td>
<td>--</td>
</tr>
<tr>
<td>Dive-and-fix</td>
<td>22067.4</td>
<td>43</td>
<td>2.18</td>
</tr>
<tr>
<td>Ratio cluster</td>
<td>21722</td>
<td>0.4531</td>
<td>0.58</td>
</tr>
<tr>
<td>Cum. departure</td>
<td>21910</td>
<td>0.2969</td>
<td>1.45</td>
</tr>
</tbody>
</table>

In examples 3 and 4, an arterial with two intersections is considered using the same configuration as in examples 1 and 2. The optimal solution of MIP was not exactly achieved after several hours of computation due to the "curse of dimensionality". The LR solutions are obtained in several seconds for both examples as shown in Tables IV(a) and IV(b). The performance of the Cumulative-departure method is superior to the first two methods in terms of computation time and accuracy. The relative error of is less than 0.3% for the two-intersection examples, which could be treated as near optimal MIP solutions.

Table 4 Comparison of different approaches

<table>
<thead>
<tr>
<th></th>
<th>Objective</th>
<th>Solution Time (sec)</th>
<th>Relative error (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) Example 3 (uncongested traffic)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

11
<table>
<thead>
<tr>
<th></th>
<th>MIP (CPLEX)</th>
<th>LR</th>
<th>Dive-and-fix</th>
<th>Ratio cluster</th>
<th>Cum. departure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Objective</td>
<td>39055.8</td>
<td>38454.5</td>
<td>39863</td>
<td>39851</td>
<td>39158</td>
</tr>
<tr>
<td>Solution Time (sec)</td>
<td>10650</td>
<td>6.4</td>
<td>195</td>
<td>0.4688</td>
<td>0.6563</td>
</tr>
<tr>
<td>Relative error (%)</td>
<td>0</td>
<td>--</td>
<td>2.07</td>
<td>2.04</td>
<td>0.26</td>
</tr>
</tbody>
</table>

(b) Example 4 (congested traffic)

<table>
<thead>
<tr>
<th></th>
<th>Objective</th>
<th>Solution Time (sec)</th>
<th>Relative error (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MIP (CPLEX)</td>
<td>76620</td>
<td>80214</td>
<td>0</td>
</tr>
<tr>
<td>LR</td>
<td>75500.7</td>
<td>8.9</td>
<td>--</td>
</tr>
<tr>
<td>Dive-and-fix</td>
<td>77860</td>
<td>265</td>
<td>1.62</td>
</tr>
<tr>
<td>Ratio cluster</td>
<td>78509</td>
<td>0.28</td>
<td>2.46</td>
</tr>
<tr>
<td>Cum. departure</td>
<td>76731</td>
<td>0.64</td>
<td>0.14</td>
</tr>
</tbody>
</table>

5. Concluding remarks

The computation time for the 0-1 mixed integer linear programming formulation for the traffic signal problem typically increases dramatically with the increase in the number of time steps and the size of the network. The linear relaxation (LR) can be solved very easily. In this paper, three algorithms are proposed by solving the 0-1 mixed-integer linear programming problem with linear relaxation and converting the infeasible non-integer solutions into feasible integer solutions.

It is shown that the corresponding traffic patterns between the optimal solution of LR and the optimal solution of MIP formulation are quite similar. Therefore, the LR solution can serve as the basis for obtaining near-optimal green time allocation for each phase. In particular, three algorithms are developed in the paper: the Dive-and-fix method, the Ratio-cluster method, and the Cumulative-departure method.

The Dive-and-fix is a generalized heuristic algorithm for binary integer programming formulations. This approach usually takes a very long time to get approximate integer solutions since a number of linear programs need to be solved one by one. The last two approaches are based on the interpretation of the LR
solution. Given the LR solution, the best dynamic green time allocations and cumulative vehicle departures right after each intersection are known for each time step, since the LR solutions always represent the lower bounds of the MIP. The second approach, Ratio-cluster, and the third approach, Cumulative-departure, are proposed to approximate the green time allocations and cumulative vehicle departure rates suggested by the LR solutions, respectively. Both perform much faster and have smaller relative errors than the Dive-and-fix method. The Cumulative-departure method outperforms the Ratio-cluster method (relative errors are less than 1%) when the problem grows large (examples 3 and 4) and its approximation mechanism can be easily implemented compared with the Ratio-cluster method.

In the future, the cumulative-departure method will be applied for online near-optimal traffic control for oversaturated intersections, which are considered a very significant but difficult problem in traffic operations. Only a two-intersection network is considered in this paper.

References


