Adaptive dynamic control for road traffic signals

BG Heydecker, Chen Cai and CK Wong

Centre for Transport Studies
University College London
Gower Street
London WC1E 6BT
England

EXTENDED ABSTRACT

1. Introduction
The control of road traffic at traffic signals provides an effective means to manage conflicting requirements for road space. In planning and applying control of this kind, decisions are required on the sequence and durations of the periods for which the different signals display green. In order to achieve effective control that provides adequate capacity, and acceptable queue lengths and delays, these control variables should be matched to the flows of traffic. Various strategies have been established for this, ranging from fixed time control in which the sequence and duration of the green indications is the same from cycle to cycle, through heuristic decision approaches based directly upon outputs from traffic detectors, to optimisation approaches based upon dynamic estimates of traffic states and arrival rates. In the present paper, we report on development of a novel adaptive optimisation-based control strategy that responds to measured variations in traffic flows. This approach is based upon the methodology of adaptive dynamic programming. We present the background to the formulation, the approach that is adopted, and some preliminary results from dynamic microscopic simulation.

2. Control of road traffic
Traffic signals work by providing an opportunity for mutually compatible traffic movements to have exclusive use of space in a road junction in turn for periods of time, known as stages, that are separated by periods of time during which the individual signals change between green and red, and clearance times elapse between movements losing right of way and those gaining it. While a stage is running, and after a certain minimum green time has elapsed for it, the control system can change to a different stage. The structure of the decisions to be made for this control can be specified at each instant by two pieces of information:
whether or not a change of stage should be initiated;
if so, which stage it is that should be called next.
If a change of stage is initiated, then this will commit the control actions for a period of several seconds – typically about 12 in the UK, comprising 3 seconds stopping amber, 2 seconds starting amber and 7 seconds minimum green (Department of Transport, 1991a). On the other hand, the decision not to initiate a change of stage can be reviewed shortly (eg after 0.5 s). These decisions are informed by detector data that relate typically to arrivals that will take place in the next 10 seconds or so.

At simple junctions, the physical layout and the logic of control can limit the number of stages to 2, so that the decision is reduced to whether or not to initiate a call of the other stage. However, at more complicated junctions, or ones where special provision
is made of opposed turning traffic (right turners in the UK), then several different stages can be used from time to time according to demand.

The traffic that is controlled by the signals at a single junction will, in general, vary on three distinct timescales. The first of these is at flows vary stochastically from moment to moment due to fluctuations in demand and driver behaviour, flows vary systematically within each day due to peak periods, and flows change over protracted periods of time due to developing demand for travel and traffic management in the vicinity. In practice, each of these separate reasons will apply to some degree.

Fixed-time traffic control systems have been developed (Allsop, 1971; 1976) to provide a good allocation of green times. Traffic responsive systems (eg MOVA, Department of Transport, 1991b) take advantage of stochastic fluctuations in traffic arrivals to allocate green time effectively according to the current and short-term future state. However, both of these approaches require knowledge of the mean arrival rates of traffic in order to calculate plans. In the latter case of responsive control, the way in which the control responds to variations in traffic arrivals can be adapted according to estimates of the mean arrival rates.

3. Dynamic control
In the present paper, we develop a dynamic approach to control of road traffic based upon Bellman’s (1957) principal of optimality. We state this formally as:

\[
C(x) = \min_{\psi} \left[ E \left[ c(\psi, x, \omega) + e^{-\gamma \tau(\psi)} C(y(\psi, x, \omega)) \right] \right]
\]

where \( C(x) \) is the total future discounted cost starting from state \( x \),
\( \psi \) is the decision (eg initiate change of stage or defer decision),
\( x \) is the state of the traffic and the controller,
\( \omega \) is information about traffic arrivals,
\( \gamma \) is a discount rate,
\( \tau(\psi) \) is the time over which decision \( \psi \) is implemented,
\( y(\psi, x, \omega) \) is the state at time \( \tau(\psi) \) after decision \( \psi \) is implemented, and
\( c(\psi, x, \omega) \) is the cost of implementing decision \( \psi \).

Several approaches have been developed for this in the past. Robertson and Bretherton (1974) used backward dynamic programming formulations in which they used hypothetical knowledge of individual arrivals over a finite time horizon: this approach calculates the future discounted cost \( C(x) \) recursively backwards from a future terminal time, using the values that have already been calculated in the right-hand side of Bellman’s equation. Although they recognised that this implies an impractical data requirement, they promoted this approach as providing an absolute minimum of cost for the arrival pattern that arises; they showed that the optimal decisions in the short run were insensitive to variations in traffic arrivals at times after about 25 seconds into the future. Gartner (1983) developed the OPAC rolling optimisation procedure that uses a direct search method over estimates of delay based upon detected arrivals for the short-term future and estimated arrivals thereafter using an expression for \( C(x) \) that is linear in queue lengths. The value of an expression of this kind will depend on the mean rate of arrivals in the future.

In the present approach, we consider future discounted delays over an infinite planning horizon. However, rather than estimate these backwards in time, they are
estimated endogenously using a method of approximate dynamic programming (ADP) (Cheung and Powell 1996). In this case, approximate values of the cost function \( C(x) \) are used on the right-hand side of Bellman’s equation to inform decisions. The resulting optimised value of \( C(x) \) on the left-hand side is then used to update the approximation for use in the next decision. This approach requires either a discrete state space for which the values can be tabulated and updated, or a parametric form for the function \( C(x) \). In the present approach, we adopt a linear form

\[
C(x) = \alpha_r \sum_{i \in R} x_i + \alpha_g \sum_{i \in G} x_i
\]

where \( R \) is the set of streams that have a red indication, \( G \) is the set of streams that have a green indication, \( x_i \) is the queue length in stream \( i \), and \( \alpha \) are parameters. The information that is used to estimate the parameters \( \alpha \) enters in the term \( c(\psi, x, \omega) \) and thus depends on the arrival rate that is manifested in \( \omega \). Through this mechanism, the way in which the control responds to traffic will be adapted according to the prevailing flows. This form of ADP therefore corresponds to an adaptive learning approach.

4. Example calculations

In order to illustrate use of this approach, we have undertaken calculations for a simple example junction. In this case, there are 2 approaches and hence 2 signal stages. We adopt a 5 second time increment throughout, with interstage time and minimum green time each equal to 1 time increment. We undertook 10 simulations with each of 8 combinations of constant mean arrival rates and evaluated 4 different control strategies:

Backward dynamic programming (Robertson and Bretherton) (BDP), Approximate dynamic programming with linear value function approximation (ADP), Robertson and Bretherton’s empirical near-optimum strategy (RB), and Gartner’s optimum sequential constrained algorithm (OPAC).

The results of this are shown in the Table below. The BDP method, here implemented after Robertson and Bretherton’s DYPIC, provides the optimal (if unrealisable) performance for each sequence of arrivals that is simulated, and therefore provides a lower bound for all delays. The performance of the ADP algorithm is in many cases better than that of the other approximate methods investigated here.

<table>
<thead>
<tr>
<th>Method</th>
<th>Flow (veh/h)</th>
<th>Arm A</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Arm B</td>
<td></td>
</tr>
<tr>
<td>BDP</td>
<td>252</td>
<td>396</td>
</tr>
<tr>
<td></td>
<td>61</td>
<td>105</td>
</tr>
<tr>
<td>ADP</td>
<td>71</td>
<td>123</td>
</tr>
<tr>
<td>RB</td>
<td>78</td>
<td>137</td>
</tr>
<tr>
<td>OPAC</td>
<td>83</td>
<td>135</td>
</tr>
<tr>
<td>BDP</td>
<td>240</td>
<td>119</td>
</tr>
<tr>
<td></td>
<td>146</td>
<td>255</td>
</tr>
<tr>
<td>RB</td>
<td>145</td>
<td>252</td>
</tr>
<tr>
<td>OPAC</td>
<td>158</td>
<td>263</td>
</tr>
</tbody>
</table>

Table: Simulation results: Total delay over 10 minutes (mean of 10 runs; vehicle-intervals)
The adaptive estimates of the parameters $\alpha$ of the total cost function are shown in Figure 1. These show good convergence within the time frame of the simulation, with ultimate values of about $\alpha = (1.92, 2.85)$ for a single case with mean arrival rates (678, 432) vehicles/h.

These results show that the control method proposed here has the capability to adapt the parameter values according to prevailing traffic flows. The resulting control performance is good by comparison with other realistic traffic responsive control methods.

References


