HYBRID OPTIMIZATION: OPTIMAL STATIC TRAFFIC CONTROL CONSTRAINED BY DRIVERS' ROUTE CHOICE BEHAVIOR

by

S.B. Gershwin and H.-N. Tan

Massachusetts Institute of Technology
Laboratory for Information and Decision Systems
Cambridge, Massachusetts 02139

ABSTRACT

This paper considers the static traffic signal control problem in a network with the relaxation of the usual fixed flow assumption implicit in most such studies. Link flows in a traffic network are realistically assumed to be variable because given a set of network control parameters such as signal settings, drivers are free to choose among alternate paths. This is called "hybrid optimization" because it combines the usual notions of user equilibrium and system optimization. Necessary conditions for the optimum solution are derived and discussed.

These conditions extend user equilibrium: not only is travel time equalized over utilized paths, but also other quantities related to the system-wide objective. Numerical algorithms are proposed.

1. Introduction

The traffic control problem is related to the operational aspects of an automotive transportation system. The objective is to regulate traffic flows by using available control devices so that the existing facilities can be most efficiently utilized. A large amount of research has been undertaken. For example, dynamic control problems have been formulated for urban network traffic [1], [6], [7] and freeway corridor traffic [2], [3].

In this paper we focus on a special case where a steady state model of traffic flow is assumed. The vehicle traffic in a network is never at rest, but there are situations where certain quantities such as the rate of traffic demand and the traffic flow distribution can be assumed approximately constant for a relatively long period of time [4]. This kind of situation typically arises in the morning and the evening rush hours [11]. The steady state traffic model has traditionally been used to simplify the analysis of transportation networks [4], [5].

Fixed-time signal control policy has been widely used for traffic control due to its simplicity in implementation [6], [8], [9], [10]. This is an open-loop control policy where for a certain period of the day, the signals operate on a fixed cycle time, with fixed phases and offsets predetermined by some offline computation [10]. A large number of fixed-time signal optimization methods have been developed [8]. Typically in all these methods, optimal traffic controls are chosen to minimize total travel time at fixed signal timings for delay and traffic behavior at a signalized intersection as a function of the control parameters.

Another very important assumption, which is seldom explicitly stated, is the assumption of fixed route choice of the drivers. This implies constant traffic volume on every link in the network under consideration. Under the fixed link traffic on any particular link is constant regardless of the level of service offered by that link. This assumption is invalid in view of the fact that no individual driver can be prevented from taking an alternate route which could have been made more desirable, i.e., faster, by the implementation of a new control policy. It seems intuitively convincing that in a fixed-time signal control system, drivers can learn to adapt their routes and speeds to advantage. In fact these redistributal effects of traffic resulting from implementation of an area traffic control policy has been confirmed in a series of field experiments conducted in the City of Glasgow [12].

It is observed [12], [15] that the new traffic pattern indirectly induced by some "optimal" traffic control policy destroys the original optimality. It would thus seem desirable to periodically reoptimize the controls based on new survey information on the traffic distribution [10]. However, this process of updating controls has seldom been carried out more than once or twice in practice due to the amount of effort and resources involved [10]. On the other hand, it has also been shown that different signal timings induce different traffic patterns [13]. In a rather different approach, Allison [14], recognizing the interdependence between signal timing plan and flow pattern, suggested the idea of using control schemes to influence traffic patterns.

Given the fact that the system has little control over the route selection decisions of individual drivers, can one hope to achieve a flow distribution which is optimal from the system's point of view using the available control? Given all the resources and effects, does the iterative reoptimization procedure necessarily lead to an optimal solution? In more generally, given certain predictive model of driver's route selection behavior, how should one go about choosing a set of controls which, together with the eventual induced traffic pattern, is optimal with
respect to certain system cost criterion? This class of problems is of fundamental importance in transportation network planning.

The hybrid optimization problem has the following essential features. The objectives of the traffic authority and the drivers are different. On the system level, the problem for the traffic authority is to minimize some overall cost in the network, e.g., total travel time, or total fuel consumption. On the other hand, the individual driver wishes to minimize his trip cost in travelling through the network. Another important aspect in the hybrid optimization problem is the role of the individual drivers as independent decision makers in choosing among different available paths. This means that it is beyond the power of the traffic authority to establish link flow at any desirable volume. Consequently, the capability of the traffic authority is limited to the command of traffic control devices only. In most cases, the capability of the traffic authority is further restrained because practical limitations dictate that the traffic authority can only exercise control over a subset of the network.

This paper represents an initial effort in this area of research. It should be emphasized that it does not lead immediately to a new design tool applicable for the solution of practical problems. The main purpose here is to emphasize the importance of drivers' role as independent decision makers in transportation planning. We hope to provide a unified approach, a better understanding (at least qualitatively), an appropriate formulation and possible directions for algorithm development for this class of problem.

In section 2 we present a mathematical formulation of the hybrid problem. A heuristic procedure which has been proposed for the solution of the hybrid optimization problem is discussed in section 3. Two simple numerical examples are solved using the mathematical formulation in section 4. In section 5, we present a discussion on the possible directions for the development of algorithms applicable for larger systems.

2. Formulation of the Hybrid Optimization Problem

An important and perhaps the most complicating issue in the hybrid optimization problem is the role of individual drivers as independent decision makers in choosing among different paths. In this section we first briefly review the flow distribution model to be used in this paper: user equilibrium. A mathematical formulation of the hybrid optimization problem is then presented.

Equilibrium flow distribution

We assume in this paper that traffic distributes itself according to Wardrop’s first principle [25] which states that “The journey time on all routes actually used are equal and less than those which would be experienced by a single vehicle on any unused route.”

Suppose there are K origin-destination pairs and L links in the network under consideration. For the jth 0-D (origin-destination) pair, let \( H^j \) be the total amount of traffic demand to go from the origin to the destination node. We denote the set of all paths connecting the jth 0-D pair by \( P^j \). In this paper we use the notation \( |S| \) for the total number of elements in the set \( S \). Let \( T^j, t^j \) be the jth 0-D pair path time and path flow vector respectively, each of dimension \( |P^j| \). Also let \( f \) and \( t \) be the link flow and link travel time vector respectively, each of dimension \( L \). The \( i \)th elements of \( t^j, h^j \), and \( T^j, t^j \), are respectively the path flow and trip travel time on the \( i \)th path of the jth 0-D pair. The link flow and link travel time on link \( i \) are respectively denoted by \( f_i \), the \( i \)th element of \( f \), and \( t_i \), the \( i \)th element of \( t \). The path flow vectors for all 0-D pairs are ordered to form a path flow vector, \( h \), which is of dimension \( K \cdot |P^j| \). Note that the link flow vector, \( f \), and the \( |P^j| \) set of path flow vectors for all 0-D pairs, \( \{h^j\} \), are related by the equation

\[
F = \sum_{l=1}^{K} A^j_l h^j
\]

where \( A^j \) is the arc-path incidence matrix with dimension \( L \times |P^j| \). For the \( l \)th 0-D pair. The \((i,j)\) element of \( A^j_l \) is defined as follows:

\[
a^j_{ij} = \begin{cases} 1 & \text{if link } i \text{ lies on the } j \text{th path of the } l \text{th } 0-D \text{ pair;} \\ 0 & \text{otherwise.} \end{cases}
\]

Wardrop's first principle can be expressed mathematically as follows:

- Link-path flow relation

\[
f^j = \sum_{l=1}^{K} A^j_l h^j
\]

- Non-negativity flow relation

\[
h^j_l > 0 \quad \forall P^j, \; l = 1, \ldots, K
\]

- Conservation of path flow

\[
h^j_l = \sum_{i} f^j_i \quad l = 1, \ldots, K
\]

- Wardrop's first principle

\[
h^j_l > 0 \Rightarrow T^j_l(f,w) = \min_{h^j_l} T^j_l(f,w)
\]

Definition A set of path flow vectors \( \{h^j\} \) is a feasible path flow if and only if \( h^j \) satisfies (3) and (4) and the correspond-
ing which satisfies (1) is called a feasible link flow.

Definition A feasible path (link) flow is a user optimized or equilibrium path (link) flow if and only if (4) and (5) are satisfied.

For more detailed discussion on equilibrium flow, interested readers are referred to [5], [11], [16], [17], [18]. The problem of computing equilibrium flow has been shown [18], [17] to be equivalent to a convex programming problem under the assumption of separability of link function, and some additional mild assumptions. A recent result [19] using nonlinear complementary theory has shown that an equilibrium flow pattern exists under some very mild conditions:

\[ t_i(a) > 0 \quad \text{if} \quad f_i > f_i' \quad \text{whenever} \quad t_i(f_i, w) \quad \text{continuous.} \]

If the concern is total fuel consumption,

\[ J = \sum_{a=1}^{L} e_i(f, w) \quad (7) \]

where \( e_i(f, w) \) is the amount of fuel consumed per vehicle in travelling through link \( i \). In both cases, the system cost is a function of control parameters and link flows which can also be expressed as a function of path flows and control parameters using the link-path flows relation in equation (1), i.e. \( J = J(f, w) = J(h, w) \).

It should be emphasized that the degree of freedom for the traffic authority is limited to the choice of control parameters from \( W \). More importantly, \( h \) is not an arbitrary feasible flow to be assigned by the traffic authority. Instead \( h \) is required to satisfy a restrictive set of conditions which describe driver behavior. In this paper, Wardrop's principle is used for this purpose. Future research will be devoted to the use of other such descriptions, such as in [21].

Problem Statement

Minimize \( J(h, w) \)

\( w \in W \)

subject to \( h \) an equilibrium flow

A mathematical optimization problem would be formulated in a straightforward manner from this problem statement if given any \( w \in W \), \( h \) were some known explicit function of \( w \) or if \( h \) were required to satisfy a set of equalities and inequalities. The problem is that neither of the two cases is true. Furthermore, the mathematical relations (3) and (6) are neither equalities nor inequalities. They are in fact two logical relations and hence they are not in a form to be posed as constraints to the optimization problem.

It can be shown that (3) to (6) can be transformed to the following equivalent form.

\[ h^T > 0 \quad i \in P, \quad \ell = 1, \ldots, K \quad (3) \]

\[ \Delta h^T l = H^T \quad \ell = 1, \ldots, K \quad (4) \]

\[ f_1^T l(h, w) > ( \sum_{j \in P} T_j^T l(h, w) h_j^T ) / H^T \quad (9) \]

The hybrid optimization problem can now be stated precisely as follows:

\[ \begin{align*}
\text{minimize} & \quad J(h, w) \\
\text{subject to} & \quad h^T > 0 \\
& \quad \Delta h^T l = H^T \\
& \quad f_1^T l(h, w) > ( \sum_{j \in P} T_j^T l(h, w) h_j^T ) / H^T
\end{align*} \]

It should be pointed out that the user optimization model for flow distribution is only one among the many available models [20], [21]. The user optimization model is used in this formulation for several reasons. It is a very common model and has been widely used in transportation planning [22], [23], [24]. Moreover all behavioral models are approximate ones and equilibrium flow has been shown [22] to be a reasonably good approximate of actual traffic distribution. However it should be emphasized that the formulation presented in this section is not restricted to any particular flow distribution model.

3. A Heuristic Procedure

In this section, we study a heuristic procedure which has been proposed and used in a number of studies [4], [13], [14], [12], [15]. This is an iterative procedure consisting of successive alternations between a signal optimizing program and an assignment program as shown in figure 1. The assignment program computes an equilibrium flow assuming the control parameters are fixed. The signal optimizing program computes a set of optimal signal settings with respect to some system cost assuming flows to be fixed. The procedure is initiated by a guess of the optimal control parameters and proceeds by iterating between the two programs until certain stopping criterion is satisfied. Various aspects of this procedure have not been closely examined. For example, does the procedure converge to a solution? If the procedure does converge, what are the properties of the result?

In this section, we establish the fact that the heuristic procedure does not necessarily converge.
to the optimal hybrid solution by using a simple counter example. Consider the network in figure 2 with link travel time assumed to be a linear function of link flow: $t_l(f_{13}) = 15 + 2f_{13}, t_{12} = 15 + 2f_{12}, t_{14} = 30 + f_{14}^{-1}, t_{32} = 30 + f_{32}^{-1}$. Link (3,4) is under the control of the traffic authority and for simplicity we assumed that the control is in the form of delay imposed on the traffic passing through link (3,4). Therefore $t_3^6 = 10 + f_3 + W$, where $w$ is the imposed delay. 10 units of traffic flow are required to go from node 1 to node 2.

Assignment Program: Let $x$ be the path flow along path (1, 3, 2) and (1, 4, 2) are the same and equal to $(10 - x)/2$. Figure 3 shows $x$ at equilibrium as a function of $w$.

Signal Optimizing Program: Assuming that the traffic authority wishes to minimize total travel time the problem of the control optimizing phase of the heuristic procedure can be shown to be the following.

$$
\min \ w \ x + 2.5((x - 3)^2 + 311) \ \text{subject to} \ \ w \geq 0
$$

The solution of this problem can be shown to be

$$
w = 0 \ \text{for} \ x > 0
$$

$$
w > 0 \ \text{for} \ x = 0
$$

We apply the heuristic procedure for this example with initial guess of $w = 10$. The result of the heuristic procedure is listed in table 1 which is constructed using (10) and figure (3). The converged solution is $w = 0, x = 8$.

Figure 4 shows the system cost as a function of $W$. It is clear that the lowest cost is achieved with $w > 20$ and Fig. 3 implies that the optimal flow on link (3,4) is $x = 0$. That is, it is best not to use this link, and there should be a sufficiently large control delay imposed so that no driver chooses to travel on it. This is thus an example of Braess' paradox [5]. The heuristic procedure, however, has done the opposite and converged to the worst possible solution.

4. Numerical Examples

It has been shown in the preceding section that the intuitively appealing action-reaction heuristic procedure converged to the worst possible solution. The main problem with the heuristic procedure is that in the control optimizing phase, controls are chosen without taking into consideration the reaction of the drivers. In this section we describe two simple examples to demonstrate that the formulation of the hybrid optimization is indeed well defined. Details appear in [26].

Example 1

We solve the same problem as in the preceding section using a general nonlinear constrained program, with the same initial guess of $w = 10, x = 4$. A solution of $w = 20$ and $x = 0$ is obtained. Figure 4 shows that this is indeed an optimal solution.

Example 2

Consider the network as shown in figure 5. Node 3 is a signalized intersection. We denote the green split facing link (1,3) by $g$. The link travel time is modelled as a fourth power polynomial of the link flow [27]

$$
t_l(f_l) = t_0(1.0 + 15(f_l/capacity)^4)
$$

The capacity is taken to be 1500 vehicle/hour per lane. All links except link (4,5) are assumed to be single-lane. We use the Webster formula [15] with fixed cycle time of 1.0 minute for the waiting time at the signalized intersection. The traffic demands are 800 veh./hr. for each of the following 0-D pairs: 1 to 5, 1 to 6, 2 to 5, 2 to 6.

This example is solved using the formulation presented in section 2 by a general nonlinear constrained program. The optimal solution obtained is $g = 0.21$ with minimum system cost (total travel time) = 880.5 vehicle-hour/hour. The optimality of $g^*$ has been verified using the results of a series of user equilibrium flow patterns computed at various values of $g$, the green split. Numerical experience on the application of the heuristic procedure to this example also shows that it converges to wrong solution.

5. Extended Equilibrium Principle for Hybrid Optimization

We present in this section an extended equilibrium principle for the hybrid optimization problem without proof. For detail information on the derivation, the readers are referred to [26].

Suppose $\{w^*, f^*\}$ is the optimal hybrid solution, then there exists some $\{\lambda_1^*, \lambda_2^*, \ldots, \lambda_K^*\}$ and $\{\pi_0^*, \pi_1^*, \ldots, \pi_K^*\}$ which are the Lagrange Multipliers of the hybrid optimization problem, such that the following statements are true.

1. The trip times along all used paths between the $j$th o-D pair are the same and equal to $T_{min}^j$. Any path having trip time greater than $T_{min}^j$ carries no flow.

2. Let $\frac{\delta L}{\delta f} = \sum_{j=1}^K (\lambda_1^j f_1^j - \pi_0^j f_0^j) + \sum_{j=1}^K \pi_0^j f_0^j + \lambda_0^j w^* + \lambda_1^j f_1^j + \ldots + \lambda_K^j f_K^j$ for each link and each 0-D pair. Let $\bar{M}_j = \frac{\delta L}{\delta f} \bar{f}$ be a pseudo cost along the $j$th path of the $j$th o-D pair. Note that $f^j = \lambda_0^j h^j j$ is the link flow vector for the $j$th 0-D pair. We have the following additional equilibrium principle:

The pseudo cost along all paths between the $j$th o-D pair are the same and equal to some value, say, $\bar{M}^j$. Any path having a pseudo cost greater than $\bar{M}^j$ carries no flow. The min solution of the hybrid optimization problem requires enumeration of all paths a priori. An algorithm has been studied in [26] which avoids this requirement by making use of the extended equilibrium principle. In [26] it has been shown that paths can be generated sequentially only when required. The solution of the hybrid optimization problem is reduced to a sequence of small nonlinearly constrained master problems and shortest path problems. This is discussed fully in [26].

6. Conclusion

It is the main purpose of this paper to emphasize the importance of drivers' behavior in trans-
portation planning. Numerical examples have been presented to show that proper evaluation of any control policy cannot be made without taking the reactions of the drivers into consideration. Numerical experience shows that the heuristic procedure leads to wrong solutions in both the examples. A practical implication is that if this procedure is implemented in real life as suggested in [15], [10], the amount of excess cost accrued over a long period of operation can be substantial.

The numerical examples presented also show that the formulation of the hybrid optimization problem is indeed well defined. We introduced an extended notion of user equilibrium principle for the hybrid optimization problem and it is shown that how this can be exploited for algorithm developments.

References


Table 1: Result of the Heuristic Procedure

<table>
<thead>
<tr>
<th>STEP</th>
<th>SYSTEM (W)</th>
<th>DRIVERS (X)</th>
</tr>
</thead>
<tbody>
<tr>
<td>INITIALIZATION</td>
<td>10</td>
<td>4</td>
</tr>
<tr>
<td>ASSIGNMENT</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>CONTROL OPTIM.</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

CONVERGED SOLUTION: \( W = 0, X = 8 \)

Figure 2: Network for Example 1

Figure 3: Flow along (1,3,4,2)

Figure 4: System Cost at Equilibrium

Figure 5: Network for Example 2