A Mixed Integer Linear Programming Model for Dynamic Route Guidance*

David E. Kaufman

AT&T Labs, Room 2B19, 379 Campus Drive, Somerset, NJ 08873

Jason Nonis, Robert L. Smith

Department of Industrial and Operations Engineering

University of Michigan, Ann Arbor, Michigan 48109

December 16, 1997

Abstract

One of the major challenges facing ITS (Intelligent Transportation Systems) today is to offer route guidance to vehicular traffic so as to reduce trip time experienced. In a cooperative route guidance system, the problem becomes one of assigning routes to vehicles departing at given times from a set of origins to a set of destinations so as to minimize the average trip time experienced (a so-called system optimal criterion) Since the time to traverse a link will depend upon traffic volume encountered on that link, link times are dynamic. The complex interaction resulting between objective function and constraints makes the dynamic problem significantly more difficult to formulate and solve than the static version. We present a mixed integer linear programming

*This work was supported in part by the Intelligent Transportation Systems Research Center of Excellence at the University of Michigan

formulation of the problem which is formally derived from a set of traffic flow assumptions. Principal among these is the simplifying assumption that vehicles upon entering a link assume the speed that traffic would attain were the traffic volume encountered on that link in steady-state. The integer variables correspond to selection of vehicle capacity constraints on the link while the continuous variables correspond to selection of vehicle routes. Implicit within this MILP formulation of the dynamic traffic assignment problem is therefore a decomposition of the problem that results in a conventional capacitated linear programming network flow problem. A small illustrative subnetwork extracted from the city of Sioux Falls is solved to optimality by IBM's OSL Branch-and-Bound algorithm.

1 Introduction

The problem of cooperative route guidance in the context of Intelligent Transportation Systems is the problem of assigning vehicular traffic to paths or routes from given origins to given destinations in such a way that average trip time is minimized, a so-called system optimal solution. (The user optimal criterion for non-cooperative route guidance leads to a descriptive, as opposed to normative, model that assigns routes so that no single vehicle can change its route and achieve a strictly lower trip time). The static version of this traffic assignment problem assumes that traffic is in a steady state so that link volumes are time invariant and the time to traverse a link depends only upon the number of vehicle routes that include that link. In the dynamic case, the origin-destination demand is allowed to be time-varying so that the number of vehicles passing through a link and the corresponding link travel times become time dependent. Unlike the elegant and complete treatment of the classic static case (Potts and Oliver [1972], Sheffi [1985]), the dynamic traffic assignment problem is still largely unexplored, at least from a formal point of view, where even its proper formulation is not clearly understood.

One of the more controversial issues in modeling the dynamic traffic assignment problem is the question of how the dynamic character of traffic speeds should be formulated. McGurrin and Wang [1991] have constructed a microscopic traffic simulation built on car-following models. The penalty for this microscopic level of detail is significant computational times.

Chen and Mahmassani [1993] and Van Aerde et al [1989] have constructed more macroscopic simulations in which vehicle speeds are determined by static speed functions applied to some average congestion level, but these approaches may produce link inconsistency (i.e. violation of first-in first-out, which is unreasonable in a deterministic model) and take limited account of the dynamic character of the problem. Janson [1989,1991] takes a similar approach in a nonlinear integer programming model with a mixed discrete/continuous time character; the model can also suffer from inconsistency and the discretization is based on relatively long intervals, again limiting the degree of dynamic character. Merchant and Nemhauser [1978], Carey [1987, 1990], Friesz et al 1989], Wie et al [1990], and Wie [1991] have proposed models where nonunique (i.e. dispersionary) link travel times are determined implicitly by link outflow functions, which give the number of vehicles to depart a link as a function of instantaneous link volume. Link outflow functions, although relatively tractable analytically, have the undesirable tendency to manifest flow rates which increase with congestion. Ran et al [1993, 1994] represent outflow rates as decision variables constrained by compound link travel time functions (distance traversal time plus queueing delay). However, the values of these link travel times are approximated by an iterative process in which each iteration itself requires iterative solution of an nonlinear optimization problem of traffic assignment in the network of interest.

In this paper, we formulate a macroscopic model that implicitly views the assignment problem as decomposed into two stages. First is the selection of time dependent link volumes, and second is the assignment of routes that optimally utilize those volumes. Since link volumes and speeds are related in a one to one fashion through link impedance functions, the first choice is equivalent to selecting a time dependent number of time periods corresponding to desired link travel times. This choice is modeled by integer valued variables. The resulting assignment problem then becomes a multicommodity network flow problem which we formulate as an ordinary linear program. The resulting mathematical program is called a mixed integer linear program or MILP. There exists an algorithm for solving such a formulation called the Branch-and-Bound algorithm (Murty [1976]). However, based upon our computational experience in solving a four node network problem as reported in section 4, an exact solution by Branch-and-Bound can only be expected for small scale networks. The principal contributions of the model offered in this paper derive at this point from its mathematically precise statement and axiomatic justification, together with its associated linear programming relaxation allowing for an exact solution through Branch-and-Bound, at least for small scale networks. However, the opportunity to solve for a true system optimal solution for even small scale network problems can result in benchmark problems against which more efficient heuristic algorithms may be tested for ability to recover most of the times savings possible.

2 The system-optimal time-expanded network model

In this section we model the problem of dynamic traffic assignment, where we seek to determine the time-dependent traffic volumes and link travel times that occur in a spatial traffic network with known topology and time-dependent origin-destination travel demands. We discretize the finite interval of time to be studied and model traffic as a continuous-valued multicommodity flow in a time-expanded version of the spatial network. Traffic congestion is modeled by simple capacity constraints, with upper bounds on instantaneous link congestion determined by impedance functions. Queueing effects on downstream links leading to spill-back are indirectly modeled through the finite capacities of these links. It is assumed that platoons of vehicles entering a link during a single time period exit the link together during a single later time period; this is modeled by multiple-choice constraints on 0-1 variables which correspond to the time-dependent travel arcs. Note here that this assumption of no platoon splitting is not so severe as it would be in the static case, since platoons are already separated by time of entry to the network.

We model the flows that would occur under cooperative behavior, i.e., dynamic system optimality, by attaching total network travel time (or equivalently average trip time) as an objective function to be minimized.

We present a general dynamic impedance model as an extension of static impedance functions. Vehicle movements are assumed to be static (i.e. in steady state) within links but dynamic (i.e. transient state) across links. The fundamental assumption is that travel time along a link is that which would be experienced were the current link loading in steady sate conditions. Hence the dynamic character occurs across and not within links. In particular,

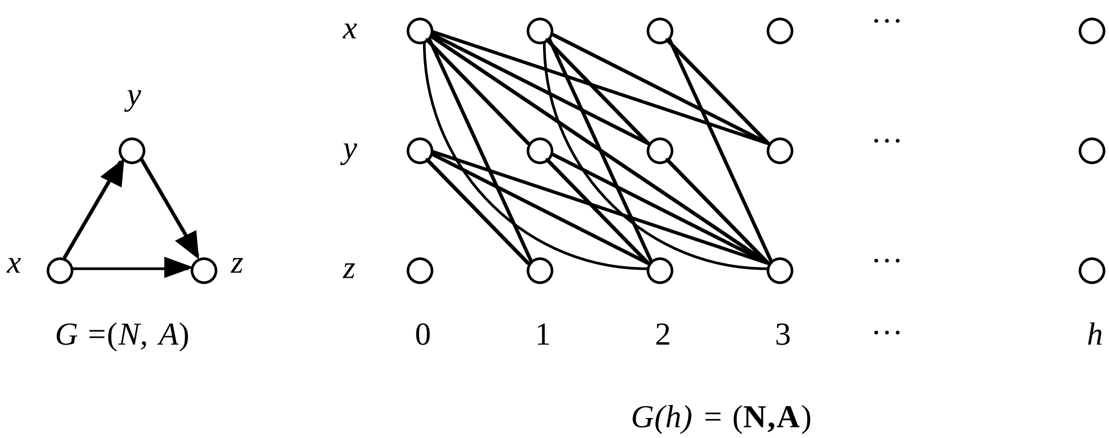

Figure 1: Time expansion of the spatial network

link congestion behaves under stationary dynamic traffic flows as it would under static flows. Daganzo [1995] has noted that dynamic traffic flow models that are driven by link impedance functions can suffer from various anomalous behaviors unless the impedance functions are severely restricted in scope. We would expect our model is subject to the same behaviors. The stability of the model as the period decreases to zero is another interesting issue which we expect to explore in a future investigation.

2.1 Network definition

Let $G = (N, A)$ be the traffic network with node set N and link set A. We model the problem with a horizon of h periods, each of duration Δt, as a time-expanded version of G, $G(h) = (\mathcal{N}, \mathcal{A})$. Corresponding to each node x of N, $\mathcal{N}$ has $h+1$ nodes $x(\tau)$, $\tau = 0, 1, \ldots, h$. Corresponding to arc (x, y) of A, $\mathcal{A}$ has arcs $(x(\tau), y(\tau + s))$, $\tau = 0, \ldots, h-1, s = 1, \ldots, h$. The time expansion is illustrated in Figure 1. Since N is assumed to contain no self-loops (x, x), we prevent stalling at nodes by not including any arcs $(x(\tau), x(\tau + s))$ in $G(h)$. (Stalling arcs would appear as horizontal arrows in Figure 1.)

We represent vehicle flows by nonnegative continuous-valued flow variables $f_d(x(\tau), y(\tau + s))$, representing the flow *volume* of traffic with final destination $d \in N$ which enters link (x, y) at time τ and exits at time $\tau + s$. We represent the total such flow volume irrespective of destination by

$$f(x(\tau), y(\tau + s)) = \sum_{d \in N} f_d(x(\tau), y(\tau + s)).$$

We adopt the general notation $g(Z) \equiv \sum_{z \in Z} g(z)$ for vectors g and sets Z, and we define $S = \{1, 2, 3, \ldots\}$ and $\bar{S} = S \cup \{0\}$. Sums over time expressed in this way will have upper limits apparent from context, determined by the time horizon. Multiple sets occurring in a single expression denote multiple summation; for example, the system-optimal objective function will include terms

$$f(x(\tau - \bar{S}), y(\tau + S)) \equiv \sum_{u=0}^{\tau} \sum_{s=1}^{h-\tau} f(x(\tau - u), y(\tau + s))$$

which give the volume, i.e., number of vehicles, on link (x, y) during the τ^{th} time period.

2.2 Traffic modeling assumptions and implied constraints

We now list our modeling assumptions governing vehicle dynamics and translate them into mathematical programming constraints.

I. *No dispersion of platoons within links.* We assume that all vehicles entering a link in a single period τ experience the same link travel time. We enforce the assumption by introducing 0-1 integer variables $\delta(x(\tau), y(\tau + s))$ for each $(x(\tau), y(\tau + s)) \in \mathcal{A}$. A vehicle platoon entering link (x, y) at time τ experiences link travel time s only if $\delta(x(\tau), y(\tau + s)) = 1$, by virtue of the constraints

$$f(x(\tau), y(\tau + s)) \leq M\delta(x(\tau), y(\tau + s)) \qquad (x(\tau), y(\tau + s)) \in \mathcal{A} \tag{1}$$

where M is an arbitrarily large constant. We prevent dispersion of the platoon by the *multiple choice* constraints

$$\delta(x(\tau), y(\tau + S)) \leq 1 \qquad (x, y) \in A, 0 \leq \tau < h \tag{2}$$

permitting at most one link travel time s (i.e. one arc $(x(\tau), y(\tau + S))$) to be chosen from S for link (x, y) at time τ). Note that if no vehicles enter link (x, y) at τ, then it is feasible not to choose any arc, i.e., $\delta(x(\tau), y(\tau + S)) = 0$. It might be preferable to require a choice

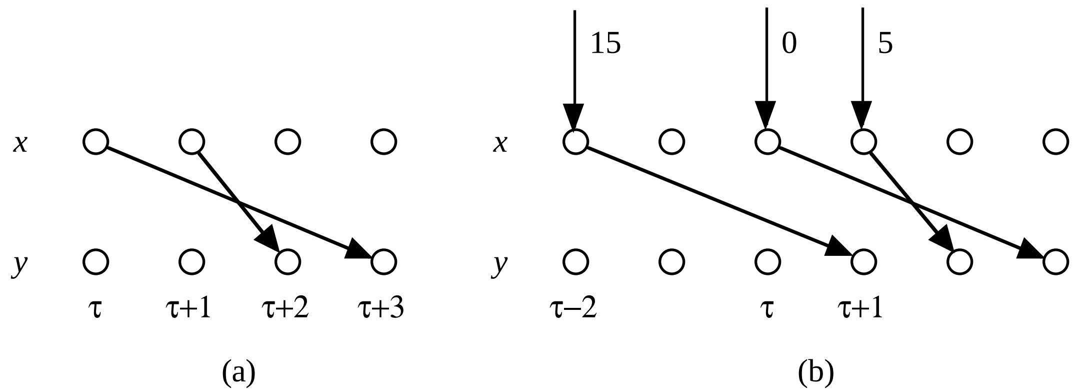

Figure 2: Link inconsistency

so that we would know what link travel time would occur if some increment of flow were rerouted onto (x, y), but the corresponding revision of (2) as strict equality constraints has a drawback, illustrated below, in connection with our next assumption.

II. *Link consistency.* We assume that vehicles do not pass one another, i.e., that among two platoons traversing a link, the one that enters later does not leave earlier. An inconsistent arc choice is illustrated in Figure 2(a). (The terminology is from Kaufman and Smith [1993], where link consistency is discussed in more detail.)

By observing that a nonzero traffic flow entering (x, y) at τ experiences link travel time equal to $\sum_{s=1}^{h-\tau} s\delta(x(\tau), y(\tau + s))$, we write the link consistency constraints

$$\tau + \sum_{s=1}^{h-\tau} s\delta(x(\tau), y(\tau + s)) \leq \omega + \sum_{s=1}^{h-\omega} s\delta(x(\omega), y(\omega + s)) \tag{3}$$

for all $(x, y) \in A$, $0 \leq \tau < \omega < h$ such that $\delta(x(\omega), y(\omega + S)) = 1$.

The constraint cannot be applied when $\delta(x(\omega), y(\omega + S)) = 0$ (whereas, if $\delta(x(\tau), y(\tau + S)) = 0$, it is vacuous). This restriction is technically nonlinear, but has the simple linear reformulation

$$\tau + \sum_{s=1}^{h-\tau} s\delta(x(\tau), y(\tau + s)) \quad \leq \quad \omega + \sum_{s=1}^{h-\omega} s\delta(x(\omega), y(\omega + s)) + M(1 - \delta(x(\omega), y(\omega + S))) \tag{4}$$

for all $(x, y) \in A, 0 \leq \tau < \omega < h$ for a suitably large value M.

We can now explain why we have chosen the inequality form of the multiple-choice constraint by discussing the example in Figure 2(b). Consider link (x, y) loaded by only those platoons shown descending vertically in the figure. Say that the platoons of 15 and 5 vehicles entering empty link (x, y) at times $\tau-2$ and $\tau+1$ have link travel times of 3 and 1 periods, respectively. In the congestion model we will present, link travel times are determined simply by the number of vehicles anywhere on the link at the moment of entry, including the entering platoon. Thus the zero-flow platoon entering at time τ would see a total of 15 vehicles on the link and choose a link travel time of three periods (shown by the dashed arc) if the equality multiple-choice form were in force. To resolve the resulting inconsistency, we would have to delay the link exit of the five-vehicle platoon by one period. We have used the inequality form in (2) to prevent this "phantom-vehicle" delay effect.

III. *Finite horizon addressed with fixed trip-completion penalties.* We will permit vehicles to enter the network throughout the study horizon, so vehicles that enter near time h may not be able to finish their trips by h. We require that these vehicles occupy some time-expanded node $y(h)$ at time h, rather than being left in the middle of a time-expanded arc. This will be accomplished by leaving all arcs of the form $(x(\tau), y(h))$ uncapacitated. We then penalize each vehicle that failed to complete its trip to node d and instead finished at $y(h)$ by an estimate $\beta(y, d)$ of the travel time required to finish the trip. For example, $\beta(y, d)$ can be chosen to be the free-flow time from node y to node d. There are at least two ways to circumvent the necessity to specify these end-of-study penalties. The first is to eventually prevent vehicles from entering the network and set the horizon sufficiently large to allow all vehicles to clear. The second is to set the horizon sufficiently distant that the effect of these penalties becomes negligible with respect to the early routing decisions. Although such a horizon is difficult to a priori compute, its existence is assured (Schochetman and Smith [1997]).

Thus the complete system-optimal objective function is

$$\min \sum_{\tau=0}^{h-1} \sum_{(x,y)\in A} f(x(\tau-\bar{S}), y(\tau+S)) + \sum_{d\in N} \sum_{(x,y)\in A} \beta(y,d) f_d(x(h-S), y(h)) \tag{5}$$

where the first term expresses the actual total travel time in the system (as explained in section 2.1), and the second term adds the trip completion penalties.

IV. *Flow conservation except at trip completion.* We require that for vehicles with any particular destination d, the number departing any time-expanded node $x(\tau)$ ($x \neq d$) is equal to the number entering $x(\tau)$ plus the number that begin their trips at $x(\tau)$. Given that $v_d(x(\tau))$ vehicles enter the network at $x(\tau)$ with destination d, we write the corresponding constraints as

$$f_d(x(\tau), N(\tau + S)) - f_d(N(\tau - S), x(\tau)) = v_d(x(\tau)) \qquad x, d \in N, x \neq d, 0 \leq \tau < h. \quad (6)$$

We omit conservation constraints for $\tau = h$ since vehicles still on the network at time h are handled via trip-completion penalties. We also omit constraints for $x = d$, allowing vehicles that reach their destinations to drop out of the network and cease contributing to the total travel time component of the objective function.

V. *Capacitated congestion modeling.* We now model the delay caused by increasing traffic loads. Our congestion modeling is determined by three assumptions:

V.1 Feasible link travel times for a platoon entering link (x, y) at time τ depend only on the volume of traffic on (x, y) at time τ, including the entering platoon.

V.2 The dynamic model of link travel time, applied to a traffic network in steady state, agrees with standard static models.

V.3 Congestion can be modeled by capacity constraints

$$f(x(\tau), y(\tau + s)) + f(x(\tau - S), y(\tau + S)) \leq c(x(\tau), y(\tau + s)) \quad (7)$$
$$\text{for all } (x(\tau), y(\tau + s)) \in \mathcal{A},\ \tau + s < h \text{ such that } \delta(x(\tau), y(\tau + s)) = 1$$

where $c(x(\tau), y(\tau + s))$ is the capacity of spatial link (x, y) given that flows entering at time τ can achieve link travel time s.

Assumption V.1 is reflected in the structure of constraint (7), which takes into account only flows $f(x(\tau - \bar{S}), y(\tau + S))$ which entered (x, y) at or before time τ and will exit after τ. Assumption V.2 helps to determine capacity values c; we defer this topic to Section 3.

The multiple choice structure allows us to sum constraints (7) over s, reducing the number of constraints by $O(h)$ and yielding

$$\begin{aligned} f(x(\tau), y(\tau+S)) &+ f(x(\tau-S), y(\tau+S)) \\ &\leq \sum_{s=1}^{h-\tau} \delta(x(\tau), y(\tau+s)) c(x(\tau), y(\tau+s)) \end{aligned} \tag{8}$$
$$\text{for all } (x,y) \in A,\ 0 \leq \tau < h \text{ such that } \delta(x(\tau), y(\tau+S)) = 1$$

which is a nonlinear restriction as was (3), but which can be reformulated linearly in the manner of (4). Also, the compressed version requires $c(x(\tau), y(h)) = M$ for all $(x,y) \in A$, $0 \leq \tau < h$ so that arcs ending at time h are effectively uncapacitated, as specified under Assumption III.

2.3 The mathematical program

The mixed integer-linear mathematical program for system-optimal dynamic traffic assignment consists of objective function (5) subject to non-dispersion ((1) and (2)), consistency (4), flow conservation (6), and capacitated congestion modeling (the linear version of (8)), given the decision variables as specified in Section 2.1.

The program requires time-dependent travel demand data $v_x(d(\tau))$ and link capacities $c(x(\tau), y(\tau+s))$. We do not further discuss the issue of travel demand data which must be provided from the field. We discuss link capacities in the following section.

3 The capacitated congestion model

In this section, we demonstrate that a standard model of the steady-state behavior of the link, in combination with Assumption V.2, uniquely determines the capacity data required for the mathematical program specified in Section 2.3.

Under Assumption V.2, we require that for a link in steady state with constant inflow rate and total volume over time, the dynamic model predicts the same link travel time that would arise in static modeling. Static traffic modeling provides speed functions $\sigma_{xy}(\cdot)$, which give the vehicle speed on link (x,y) as a function of the time rate of flow across the link, assumed constant over time. We assume that σ is a positive, continuous, decreasing function.

To make the dynamic model act as a true generalization of the static model, we construct a steady-state loading in our time-expanded network $G(h)$ where λ vehicles enter link (x, y) in each period. The speed on (x, y) is therefore $\sigma_{xy}(\lambda)$, constant over time. We denote the physical length of (x, y) by d_{xy}, and thus

$$\sigma_{xy}(\lambda) = \frac{d_{xy}}{s} \tag{9}$$

where s is the steady-state link travel time. The speed function σ has a well-defined inverse, thus $\lambda = \sigma^{-1}(d_{xy}/s)$.

Theorem 1 *Assumption V.2 implies*

$$c(x(\tau), y(\tau + s)) = s\sigma_{xy}^{-1}\left(\frac{d_{xy}}{s}\right) \qquad (x(\tau), y(\tau + s)) \in \mathcal{A}.$$

Proof: With a flow rate of $\lambda = \sigma^{-1}(d_{xy}/s)$ vehicles per period, assumed feasible, constraint (8) requires

$$c(x(\tau), y(\tau + s)) \geq s\sigma_{xy}^{-1}\left(\frac{d_{xy}}{s}\right), \tag{10}$$

the right-hand side of (10) being the number of vehicles on link (x, y) at any time instant. Now assume strict inequality holds in (10). Then $\sigma_{xy}(c(x(\tau), y(\tau + s))/s) < d_{xy}/s$, contradicting (9). ∎

4 Computational example

The formulation we have developed presents a very challenging mixed integer linear program to solve. Given a particular choice of the values of the 0-1 variables δ, many flow variables are eliminated by constraints (1). However, $O(|N||A|h)$ columns remain in the underlying linear program. More importantly, there are $O(h^{|A|h})$ ways to specify the complete set of δ values, ruling out optimization strategies based on complete enumeration.

One possible use of the MILP in large networks is to optimize small subnetworks in isolation. Accordingly, we removed a centrally located subnetwork of four nodes, as shown

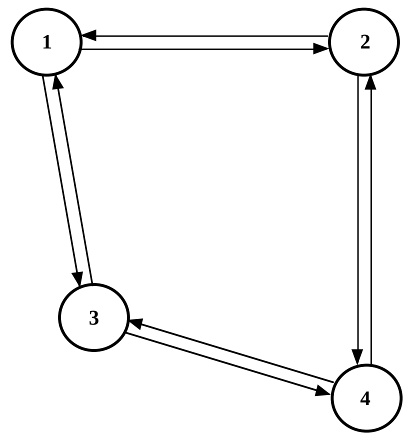

Figure 3: Sioux Falls subnetwork

in Figure 3, from the Sioux Falls network of 24 nodes studied by Leblanc [1975]. Our small test network has four nodes, eight physical links, and four destinations. We chose to study a five period problem with the standard time interval representing 1.5 minutes.

4.1 Travel demand data

We produced an imputed set of travel demand data by identifying all shortest paths in the original network under freeflow (i.e., zero-flow) link travel times. For each travel demand for flow from node x' to node d' in the full network, we counted those vehicles as demanding travel from node x to node d in the subnetwork if and only if the freeflow shortest path from x' to d' intersected the subnetwork, entering at x and exiting at d. We accepted the resulting totals as mean entry rates $\bar{v}_{xd}$ per period for each origin-destination pair, and generated a set of time-varying travel demands by sampling an entering flow $v_d(x(\tau))$ for each period τ from the normal distribution with mean $\bar{v}_{xd}$ and variance chosen nominally as $0.1\bar{v}_{xd}^2$.

The resulting time-varying origin-destination demands in vehicles per period are shown in Table 1, along with trip completion penalties β estimated as freeflow trip completion times. (Recall that the penalty $\beta(y, d)$ is applied to any vehicle with destination d which is at node y at the end-of-horizon time h. Recall also that all nodes $(x(\tau), y(h))$ are uncapacitated, so that vehicles always occupy some node y at time h rather than being left in mid-link. This has the effect of truncating all link travel times which would otherwise extend beyond h.)

O-D for time 0

	1	2	3	4
1	0	39	17	10
2	40	0	7	38
3	14	5	0	14
4	10	36	14	0

O-D for time 1.5

	1	2	3	4
1	0	45	17	11
2	44	0	5	31
3	21	5	0	15
4	12	44	15	0

O-D for time 3.0

	1	2	3	4
1	0	41	16	11
2	38	0	7	29
3	17	6	0	13
4	13	36	13	0

O-D for time 4.5

	1	2	3	4
1	0	37	14	12
2	38	0	7	42
3	15	6	0	13
4	9	35	12	0

O-D for time 6.0

	1	2	3	4
1	0	44	17	12
2	42	0	5	41
3	17	6	0	16
4	12	30	13	0

$\beta(x, y)$ in Δt units

	1	2	3	4
1	0.000	2.000	1.768	3.576
2	2.000	0.000	3.768	2.348
3	1.768	3.768	0.000	1.808
4	3.576	2.348	1.808	0.000

Table 1: Actual travel demands and trip completion penalties used in the test problem

4.2 Capacity coefficients

The congestion behavior of the network is given in a different form than was assumed for our discussion in Section 3, but as we now demonstrate, the principle by which we choose capacities $c(x(\tau), y(\tau + s))$ is sufficiently flexible to handle this variation.

Rather than being given speed functions $\sigma_{xy}(\lambda)$ and physical link lengths d_{xy}, the Sioux Falls network has associated link travel time functions which follow the well-known BPR form (cf. Branston [1976]):

$$T_{xy}(\lambda) = T^0_{xy}\left[1 + 0.15\left(\frac{\lambda}{C_{xy}}\right)^4\right]$$

where λ is a vehicle flow rate and $T_{xy}(\lambda)$ is its associated link travel time, given the freeflow link travel time T^0_{xy} and the practical link capacity C_{xy}. This data is given in Table 2 for the links in the subnetwork, with travel times in units of 1.5-minute periods and practical capacities as vehicles per period. (The links in this network are assumed to have identical characteristics in either direction. For example, link (2,1) has freeflow duration and capacity equal to that of link (1,2).)

Link (x,y)	Freeflow travel time T^0_{xy}	Practical capacity C_{xy}
(1,2)	2.000	25.00
(1,3)	1.768	12.50
(2,4)	2.348	33.75
(3,4)	1.808	12.50

Table 2: Link travel time function data

To apply our prior reasoning to this form, we write the associated speed function as

$$\sigma_{xy}(\lambda) = \frac{d_{xy}}{T(\lambda)}$$

where d_{xy} is now an *effective* link length, rather than a physical length. (We need not determine a value for the effective link length, as it is about to drop out of the capacity expression.) We can easily find the inverse speed function to be

$$\sigma^{-1}_{xy}(g) = C_{xy} \left[\frac{1}{0.15} \left(\frac{d_{xy}}{gT^0_{xy}} - 1 \right) \right]^{0.25}$$

and thus by Theorem 1, we find

$$c(x(\tau), y(\tau + s)) = sC_{xy} \left[\frac{1}{0.15} \left(\frac{s}{T^0_{xy}} - 1 \right) \right]^{0.25},$$

into which we substitute $s = 1, 2, \ldots, h$ for each link to complete the data needed for the MILP.

4.3 Solution results

The MILP formulation for this example has 116 integer variables, 304 continuous variables, and 296 constraint rows. To gauge the significance of the optimal solution, we first constructed an initial heuristic solution by assigning all vehicles to their freeflow paths. The resulting objective function value was 4866.0 minutes.

Using the branch-and-bound MILP solution package of the IBM Optimization Subroutine Library, we obtained an optimal solution with objective function value 4803.0 minutes, an improvement of 1.3%. The link travel times in effect in the optimal solution are illustrated in Figure 4, by the device of showing a link $(x(\tau), y(\tau+s))$ of the time-expanded network if and only if $\delta(x(\tau), y(\tau+s)) = 1$. For ease of reading, the upper graph shows link travel times for vehicles at nodes 1 and 4, and the lower graph gives the same information for nodes 2 and 3.

Note that the model does have the ability to model the balancing of flows by route splitting, i.e., giving different routes to vehicles with identical characteristics (location and destination). For example, although we have not shown the flow volumes divided by destination class, it can be verified that some vehicles enter at node 4 at time 1:30 and take the long route through nodes 2 and 1 to reach node 3, rather than going on link $(4, 3)$ directly. This allows a reduction in link travel time on $(4, 3)$ from 4.5 minutes to 3 minutes, giving a clear indication that the system-optimal solution is distinct from a user-optimal, or Nash equilibrium, solution.

We next repeated the generation of random travel demands to produce ten data sets. The optimal solutions of the resulting programs improved on the associated freeflow-heuristic solutions by an average of only 2.2%. However, we feel that this seeming inutility of the MILP solution process is substantially due to end-of-horizon effects. Although the solution required only six minutes on an IBM RS/6000 workstation, the size of the problem strained the available memory, inducing our choice of a horizon of only five periods. As a result, many of the vehicles which enter the network are unable to complete their trips by the end of the horizon, thus incurring the trip completion penalties β, which were derived from freeflow link travel times. Therefore, the cost function of the MILP bears a strong resemblance to pure freeflow travel time, which was minimized in generating the heuristic solution. The optimal MILP solution value would thus be expected to be close to the minimal freeflow travel time.

If this modest improvement over a freeflow solution were to persist in more extensive empirical studies, this would suggest the important conclusion that heuristic methods could be expected to capture most of the travel time savings possible. (Another more disturbing implication would be that potential savings through route guidance are not substantial.) An implementation of Branch-and-Bound that specifically exploited the network structure of the

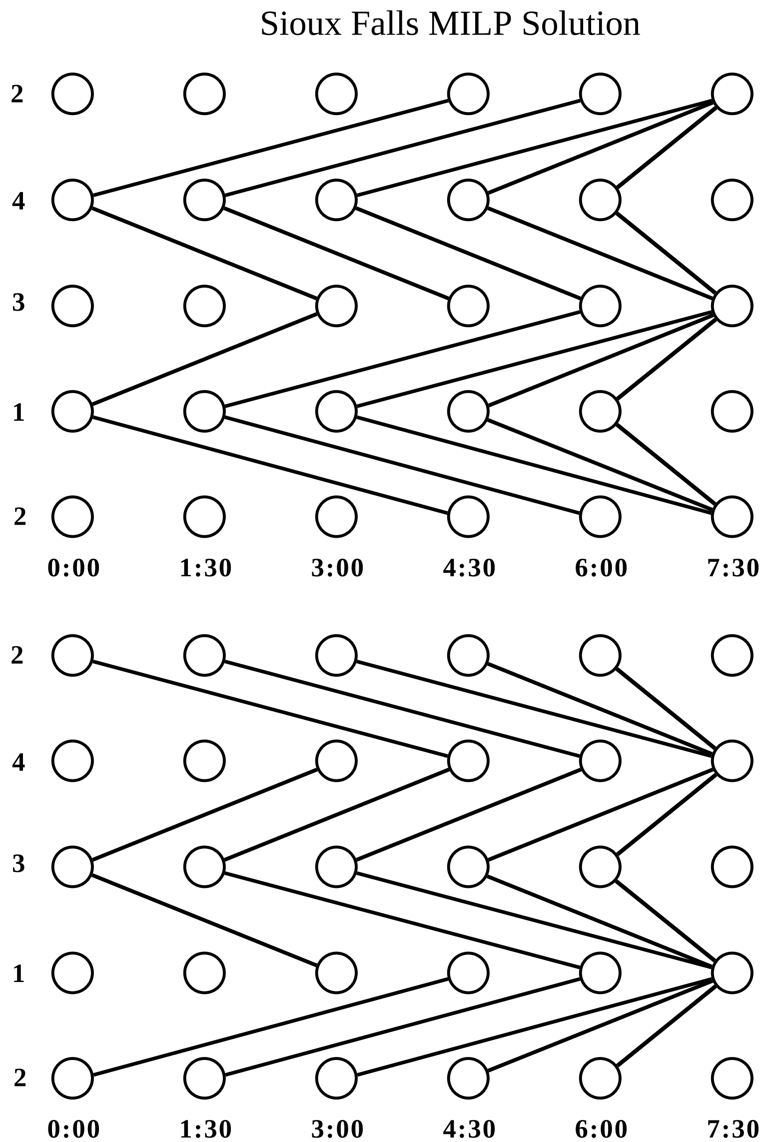

Figure 4: Link travel times in system-optimal solution

subproblems to be solved during its execution would be in a position to solve a considerably longer horizon problem that we were able to address in this initial study.

References

[1] Boyce, D.E., B. Ran, and L.J. LeBlanc. 1995. Solving an Instantaneous Dynamic User-Optimal Route Choice Model. *Transportation Science* **29** 2 128–42.

[2] Branston, D. 1976. Link Capacity Functions: A Review. *Transportation Research* **10** 223–236.

[3] Carey, M. 1987. Optimal Time-Varying Flows on Congested Networks. *Operations Research* **35** 1 58–69.

[4] Carey, M. 1990. Extending and Solving a Multiperiod Congested Network Flow Model. *Computers & Operations Research* **17** 5 495–507.

[5] Chen, P.S. and H.S. Mahmassani. 1993. An Investigation of the Reliability of Real-Time Information Systems for Route Choice Decisions in a Congested Traffic System. *Transportation* **20** 2 157–178.

[6] Friesz, T.L., J. Luque, R.L. Tobin, and B.W. Wie. 1989. Dynamic Network Traffic Assignment Considered as a Continuous Time Optimal Control Problem. *Operations Research* **37** 6 893–901.

[7] Janson, B.N. 1989. Dynamic Traffic Assignment for Urban Road Networks. *Transportation Research B* **25B** 2/3 143–161.

[8] Janson, B.N. 1991. A Convergent Algorithm for Dynamic Traffic Assignment. *Transportation Research Record* **1328** 69–80.

[9] Daganzo, C.F. 1995. Properties of Link Travel Time Functions Under Dynamic Loads. *Transportation Research B* **29B** 95–98.

[10] Kaufman, D.E. and R.L. Smith. 1993. Fastest Paths in Time-Dependent Networks for Intelligent Vehicle-Highway Systems Application. *IVHS Journal* **1** 1 1–11.

[11] LeBlanc, L.J. 1975. An Algorithm for the Discrete Network Design Problem. *Transportation Science* **9** 183–199.

[12] McGurrin, M.F. and P.T.R. Wang. 1991. An Object-Oriented Traffic Simulation with IVHS Applications. *SAE Vehicle Navigation and Information Systems Conference Proceedings* P-253, 551–561.

[13] Merchant, D.K. and G.L. Nemhauser. 1978. A Model and an Algorithm for the Dynamic Traffic Assignment Problems. *Transportation Science* **12** 3 183–199.

[14] Murty, K. 1976. *Linear and Combinatorial Programming*. John Wiley, New York.

[15] Potts, R. B. and R.M. Oliver. 1972. *Flows in Transportation Networks*. Academic Press, New York.

[16] Ran, B., D.E. Boyce, and L.J. LeBlanc. 1993. A New Class of Instantaneous Dynamic User-Optimal Traffic Assignment Models. *Operations Research* **41** 1 192–202.

[17] Ran, B. and D.E. Boyce. 1994. *Dynamic Urban Transportation Network Models: Theory and Implications for Intelligent Vehicle-Highway Systems*. Springer-Verlag, Berlin.

[18] Schochetman, I. E. and R.L. Smith. 1997. Existence and Discovery of Average Optimal Solutions in Infinite Horizon Optimization. *Mathematics of Operations Research*, in press.

[19] Sheffi, Y. 1985. *Urban Transportation Networks*. Prentice-Hall, Englewood Cliffs, N.J.

[20] Van Aerde, M., J. Voss, A. Ugge, and E.R. Case. 1989. An Integrated Approach to Managing Traffic Congestion in Combined Freeway and Traffic Signal Networks. *ITE Journal* **59** 2 36–42.

[21] Wie, B.W., T.L. Friesz, and R.L. Tobin. 1990. Dynamic User Optimal Traffic Assignment on Congested Multidestination Networks. *Transportation Research B* **24** 6 431–442.

[22] Wie, B.W. 1991. Dynamic Analysis of User-Optimized Network Flows with Elastic Travel Demand. *Transportation Research Record* **1328** 81–87.