Implementing Travel Forecasting with Traffic Operational Strategies

ALAN J. HOROWITZ

An "adaptive" travel forecasting model that has the ability to account for the ways in which the traffic system operates is described. Within the model, trip distribution, mode split, and traffic assignment are all sensitive to delays at intersections as well as delays along uncontrolled road segments. To the extent possible, delay relationships were adapted from the 1985 *Highway Capacity Manual* (HCM). Separate relationships were included for all-way stop controlled, priority, and signalized intersections. The model was implemented as a specially modified version of the Quick Response System II (QRS II) software. The HCM signalized intersection procedures were difficult to incorporate because they result in delay/volume relationships that are nonmonotonic and discontinuous. Even with better-behaved delay relationships, it is unlikely that a unique solution could be obtained. Operational strategies can be either automatically calculated or specified by users, depending on their nature. The inclusion of operation strategies in the forecast does not greatly increase computation time, data requirements, or the needed level of user expertise. A form of elastic-demand incremental assignment could find at least one user-optimal equilibrium solution. An adaptive model can reduce dependence on base-year calibration for incorporating operational effects.

Long-term transportation plans are traditionally prepared by evaluating several fixed alternatives against expected future conditions. Sometimes a travel forecast reveals a major oversight in assumptions about the way an alternative is operating, in which case the alternative might be reworked and reevaluated. The determination of when an oversight in operation has occurred is entirely judgmental; many inconsistencies between a travel forecast and an alternative are routinely tolerated.

It is presently feasible to build travel forecasting models that can automatically account for the way facilities are operated. Conceivably, such models could set signal timing, remove on-street parking, determine signal coordination in a corridor, or choose traffic control devices for intersections. The model could make a large number of short-term, microscale design decisions that could not have been made without knowledge of future traffic volumes.

For example, planners normally expect travel forecasts to reflect delays at intersections. Intersection delays depend on signal timing, and, operationally, signal timing depends on traffic volumes. Consequently, a forecasting model should be capable of calculating important aspects of signal timing (number of phases, length of phases, and cycle length) for every intersection in the network as traffic is being assigned (*1–4*). In essence, the travel forecast can be adaptive in the same sense that an actuated signal control system is adaptive. Unless the model can make consistently good assumptions about signal timing, the implied capacities of arterials will be wrong and the resulting forecast will be wrong. This type of forecasting will be referred to here as "adaptive travel forecasting."

If one were to try to create a traffic-optimizing travel forecasting model, it would look something like a hybrid between a traditional urban transportation planning (UTP) model and an operations-level traffic model (such as SOAP, PASSER-II, or TRANSYT). Such a model would be very complex, requiring considerably more data, computer resources, and user expertise than either of its two constituents. A notable example of this type of model is Continuous Traffic Assignment Model (CONTRAM) from the Transport and Road Research Laboratory (*5*). The essential differences between a CONTRAM-like model and an adaptive travel forecasting model are discussed later in this paper.

Adaptive travel forecasting is a special case of the network design problem (*6,7*), which attempts to find societally optimized networks, perhaps involving new facilities as well as operational strategies. Unlike the network design problem, adaptive travel forecasting seeks to be entirely predictive of the impacts of specific alternatives. That is, the model attempts to forecast what will be, not what should be. Clearly, the model must deal with short-term operational strategies to arrive at reasonable link volumes, but these strategies have little value by themselves. The network design problem is inherently more demanding than adaptive travel forecasting.

This paper describes initial experiences with an adaptive travel forecasting model, which has the ability to modify traffic controls at numerous isolated intersections and gives users the ability to make other operational adjustments. The focus is on applications rather than theoretical issues. How much adaptation can reasonably be included? Can the model be made sufficiently consistent with existing methods of travel forecasting, traffic theory, the *Highway Capacity Manual* (HCM) (*8*), and traffic engineering practice? Can an equilibrium solution be found? Are data and computer requirements reasonable? Conversely, does adaptation impose limitations on the size of networks? Does adaptation interfere with elastic-demand assignments? Does adaptation make forecasts more difficult to interpret? Each of these questions will be addressed in the following sections.

ADAPTIVE TRAVEL FORECASTING MODEL

Adaptation was added to an existing travel forecasting package, Quick Response System II (QRS II). QRS II is typical

Center for Urban Transportation Studies and Civil Engineering and Mechanics, University of Wisconsin-Milwaukee, Room E387, EMS Building, P.O. Box 784, Milwaukee, Wis. 53201.

of the UTP models currently used by planning agencies and consultants in the United States, so it provides a fair test of the difficulties involved in implementing adaptation. Extensive revisions to the source code were required to establish signal timings, compute delays, find incrementally averaged trip tables, account for all turning movements, and identify the conflicting and opposing traffic for every approach.

Because of uncertainties of the effects of adaptation on forecasts, only phasing and signal timing at signalized intersections were made automatic. Cycle length, quality of progression, and the placement of stop signs were kept as manual (but explicit) adjustments to the network.

An adaptive model might have to simultaneously handle thousands of traffic-controlled intersections, so an early decision was made to hold the amount of data required for any given intersection to a bare minimum. To calculate intersection delay the model was designed to only require information about an approach's lane geometry, the saturation flow rate of its through lanes, and the form of traffic control. Additional data, such as link speed and street continuity, are already required by QRS II for the purposes of travel forecasting.

Delay

Separate delay procedures were implemented for signalized intersections, all-way stop controlled (AWSC) intersections, and priority intersections (one-way and two-way stops). To the extent possible, consistency with the 1985 HCM was maintained. The choice of the HCM procedures was made on the basis of their wide adoption by planners and traffic engineers and can be considered arbitrary from the standpoint of traffic flow theory. It is possible that delay relations from other countries could perform better in a travel forecasting application.

At this writing the HCM provides insufficient treatment of AWSC intersections. So instead, delay at AWSC intersections was calculated by an enhanced form of Richardson's M/G/1 queuing model (*9*). Additional terms were provided to handle delays from turning and from coordination between drivers on subject and opposing approaches. The M/G/1 model was successfully calibrated to data provided by Kyte (*10*), and it produces results that are consistent with the recently released interim procedures for AWSC capacity (*11*).

The HCM's procedure for priority intersections omits calculation of delay, does not account for equilibrium in lane utilization for multilane approaches, and does not provide capacities when there are large conflicting volumes. These problems were remedied. Delay was calculated as if stops were a random, single-server queue as suggested by the Swedish Highway Capacity Manual (*12*). More complete capacity relations were obtained from Baass (*13*).

Total approach delay for signalized intersections is calculated consistently with the HCM, except as follows:

1. As an expedient, delay is not separately calculated for exclusive right lanes. Rather, sufficient capacity to handle just right-turning traffic is added to the capacity of the TR or LTR group before application of the delay formula.
2. The possible presence of pedestrians is ignored in order to reduce data requirements.
3. Acceleration delay is calculated from link speed and an estimate of the number of stopping vehicles.
4. To avoid a serious discontinuity in the delay function, shared left lanes were allowed to act as exclusive left lanes only when the model determines that a protected left phase is required.
5. The HCM's stopped delay formula was slightly modified for volume-to-capacity ratios (X) greater than 1.0 to eliminate the possibility of infinite or negative delay values in the uniform delay term when a green time constitutes nearly the full cycle.

Total approach delay is found by taking a volume-weighted average of delay across all lane groups and all phases. A separate delay value was not calculated for left turns, because the HCM procedure cannot provide this value in all cases. The desirability of using separate left turn delays when provided by the procedure was not investigated in this study. Users retain the ability to add a left-turn penalty at individual approaches, if desired.

Phasing and Timing

Green times, saturation flow rates, and delays are intrinsically linked. As noted previously, the calculation of delay requires knowledge of lane-group capacity as given by saturation flow rates and green times. Green times depend on saturation flow rates as parts of flow ratios. Of course, the saturation flow rate for a lane group depends on left-turn volumes and the amount of opposing traffic.

Consequently, it becomes necessary to simultaneously solve for green times and saturation flow rates. Once these have been established, delay can be ascertained. This calculation must be performed for each of the many signalized intersections, for each hour in the analysis period, and at each traffic assignment iteration. Recognizing that the amount of calculation could be prohibitively large, it is necessary to place limits on the range of signalization strategies available for any given intersection.

These rules have been adopted:

1. Green time for a TR or LTR phase is allocated in proportion to the critical flow ratio for the phase.
2. To avoid very small green times, a minimum flow ratio can be established by the user for the purpose of signal timing.
3. A protected left phase is provided only if there is insufficient left-turn capacity during an LTR phase, considering both sneakers and gaps in opposing traffic.
4. If protection is required, the phasing is always equivalent to "dual leading lefts with overlap."

It is recognized that these rules do not produce the best signal timing, but they at least find an acceptable timing in accordance with traffic engineering practice. The issue of whether an adaptive model should find optimal, rather than conventional, signalization remains unresolved.

The concept of link capacity within UTP models is seriously weakened in an adaptive travel forecasting model. It is only known that traffic volumes should not exceed the saturation flow rate, less any approach capacity lost during phase changes. Otherwise, links compete for slices of time at intersections.

CREATING AND INTERPRETING AN ADAPTIVE TRAVEL FORECAST

Adaptive Model Operation

Figure 1 shows one possible way of operating an adaptive model. There are two loops. The inner loop automatically adjusts signal timing and incorporates congestion effects. The outer loop provides the user with opportunities to modify traffic control devices in accordance with traffic engineering practice.

This method of operating the model will produce an elastic-demand assignment. That is, both the distribution of trips throughout the urban area and the level of transit ridership will reflect vehicular flow conditions on the highways.

The allocation of traffic engineering principles to the two loops is somewhat arbitrary. As more is learned about adaptive travel forecasting, principles can be transferred from the outer loop to the inner loop. For example, textbooks provide simple rules for determining optimal cycle length at isolated intersections; these rules could be moved to the inner loop provided we also have some way to determine which intersections are truly isolated.

This iterative procedure raises the serious methodological issue of whether it is possible to obtain an equilibrium solution in accordance with Wardrop's first principle—a user-optimal assignment—in a reasonable amount of time on a very large network. (Networks in QRS II could be as big as 585 zones and 4,500 links on a microcomputer; other packages permit nonadaptive networks many times this size). It will be shown later in this paper that it is at least sometimes possible to obtain an equilibrium solution with large-network methods and that it is always possible to determine whether any given solution is an equilibrium one.

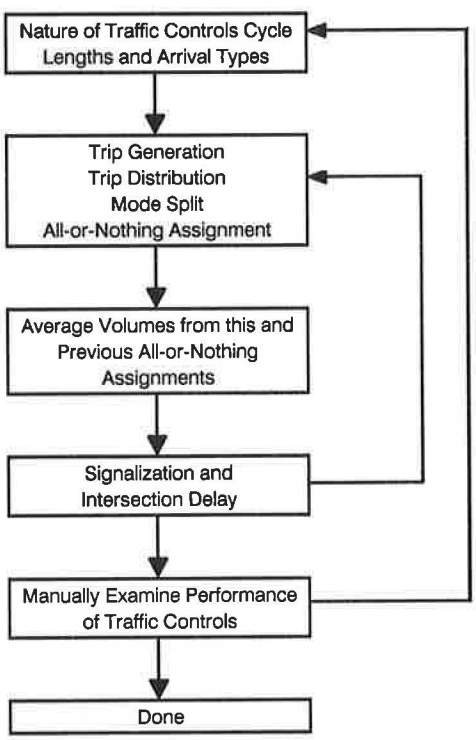

FIGURE 1 Operating an adaptive travel forecasting model.

Obtaining an Equilibrium Solution

A promising heuristic for obtaining an equilibrium solution involves an averaging of traffic volumes from many all-or-nothing assignments. Such an averaging step is fundamental to the most widely cited equilibrium assignment techniques: Frank-Wolfe decomposition for fixed-demand assignments (*14*), Evans's algorithm for elastic-demand assignments (*15*), and convergent incremental assignment (*16*).

Nonlinear optimization methods for large networks, such as Frank-Wolfe decomposition or Evans's algorithm, cannot be applied for two reasons. First, the adaptive model lacks a closed-form and well-behaved delay/volume function. Second, delay at any given approach is a function of all possible movements at the intersection. The second problem by itself could possibly be overcome by assignment methods designed to handle "asymmetric" networks (*17*).

Of the three aforementioned techniques, only incremental assignment can be implemented with adaptive travel forecasting, as it is now understood. In effect, incremental assignment creates a weighted average of many all-or-nothing assignments. The weights are predetermined; they do not depend on knowledge of the delay/volume function. Incremental assignment converges to a user-optimal equilibrium solution for fixed-demand assignments (*16*), runs only slightly slower than Frank-Wolfe decomposition (*18*), and works well on a broad range of elastic-demand problems (*19,20*). Incremental assignment has already been tested in the United Kingdom on networks with traffic controls by the authors of JAM as reported by Lewis and McNeil (*21*). Since there does not yet exist theory to suggest that incremental assignment works properly on adaptive networks, its usefulness must be established empirically.

A solution to an adaptive network may not be unique, even if it is an equilibrium solution. Because the model reallocates limited resources (such as intersection capacity or favorable coordination) across facilities, it is entirely possible to have many good solutions (*1,22*).

Consider the network of Figure 2. It consists of a trip origin, a trip destination, a single signalized intersection, and four

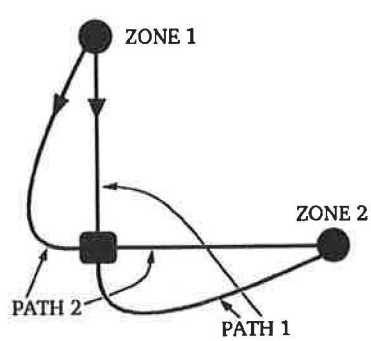

FIGURE 2 Two-zone, single-intersection network.

two-way links with capacities only at the intersection approaches. Turns at this intersection are fully restricted, so there are only two paths from the origin to the destination. Traffic is uncongested. Both paths have identical characteristics, but they compete for green time at the intersection. There are exactly three assignment solutions that could reasonably satisfy Wardrop's first principle:

A. All vehicles use Path 1,
B. All vehicles use Path 2, and
C. Both paths are used equally.

Further assume that the intersection is operating below capacity and that no minimum is established for the length of green phases. Solutions A and B would cause nearly the full cycle to be allocated to one path or another, thereby minimizing travel time for all vehicles. Solution C allocates green time equally to both streets, and it is the solution that would have been obtained in a nonadaptive network with a monotonically increasing delay/volume function. A quick inspection of the uniform delay term of the HCM delay formula reveals that Solution C is by far the least desirable from the standpoint of user cost.

It is possible to conclude from this example that equilibrium solutions to an adaptive network are not unique, and they would be likely to differ from those of a nonadaptive network. Furthermore, this example raises some tough methodological questions, which cannot be definitively answered here. How can we deal with multiple solutions to the same problem? Do we care to find more than one solution? Is there a tendency in adaptive travel forecasts toward all-or-nothing assignments?

Comparison with CONTRAM-Like Models

Since many traffic engineers are familiar with CONTRAM, it provides a good basis for comparing the current work. CONTRAM and QRS II are approaching the same methodological position from opposite directions but remain some distance apart. CONTRAM is fundamentally a traffic-optimization model that permits path choice; QRS II is fundamentally a travel demand/assignment model that contains explicit traffic flow relations. The two models have similar philosophical underpinnings but differ considerably in the types of problems they can address. CONTRAM is geared to small networks with fixed demands; QRS II is geared to large networks with elastic demands. CONTRAM's primary outputs are optimized traffic controls and indicators of the performance of the traffic system; QRS II's primary output is assigned traffic volumes. CONTRAM will yield results that are of little interest to those doing medium- or long-term travel forecasts, such as queue lengths and optimized green times. Interestingly, CONTRAM is still well ahead of conventional travel forecasting models by implementing a form of dynamic traffic assignment, something that planners are just now recognizing as important. However, CONTRAM's assignment algorithm cannot ensure that assigned volumes represent equilibrium conditions.

DELAY/VOLUME RELATIONSHIPS AT ADAPTIVE SIGNALIZED INTERSECTIONS

As indicated previously, delay must be found by simultaneously solving for saturation flow rates and green times. For this research the solution was found by the method of successive approximations.

Step 0. Determine the need for a protected left phase at each approach, so that the number of phases is known for the remaining steps. A by-product of this step is an initial estimate of saturation flow rates.
Step 1. Estimate green times.
Step 2. Estimate saturation flow rates according to the HCM procedure.
Step 3. Check for convergence. If not converged, go to Step 1.
Step 4. Calculate stopped and acceleration delay for each phase and lane group. Find the volume-weighted average for each approach.

There are other possible algorithms. Because the number and type of phases are determined before finding the length of phases, this algorithm will result in only one of many possible signal timings. No attempt is made to select an optimal timing. The ability of this algorithm to converge was individually checked on approximately 800 intersections of varying characteristics. Approximately six iterations are required to obtain four significant digits in the saturation flow rates.

Delay/volume relationships for approaches to signalized intersections in an adaptive model are considerably different from those typically seen for fixed-capacity facilities. Figure 3 shows a typical intersection with delays on all approaches varying as volume changes on a single, subject approach. Each approach at this four-way intersection has two shared lanes, there are no turning vehicles, and volumes at conflicting and opposing approaches are held at 800 vph. The volume on the

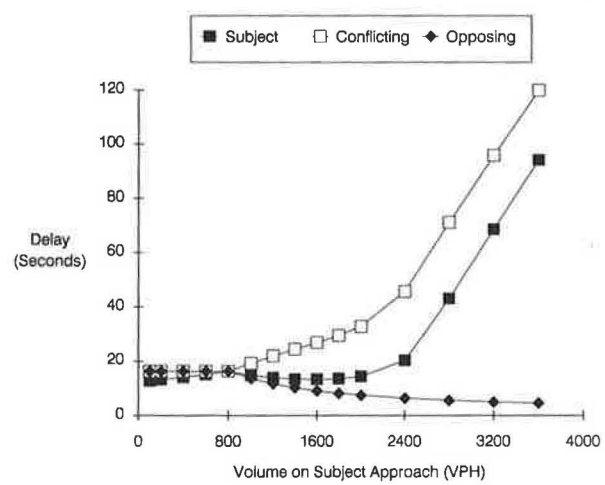

FIGURE 3 Delay on all approaches of a signalized intersection as a function of volume on a single approach (0 percent right turns, 0 percent left turns, 800 vph at opposing and conflicting approaches, no exclusive lanes, 3,600 vph ideal saturation flow rate, 20 mph speed).

subject approach is varied from 50 to 3,600 vph (the ideal saturation flow rate).

The delay at all approaches is strongly affected by variations of volume on the subject approach. The fact that the subject and opposing delay curves have roughly similar shapes is coincidental. Delay on the opposing approach declines with increasing subject approach volume beyond 800 vph. The decline is due to the increasingly ample green time to handle a given volume. Figure 3 also shows that delay on the subject approach is not even monotonic. Subject approach delay rises to a local maximum at 800 vph (the fixed volume on conflicting and opposing approaches), then declines to a local minimum at 1,600 vph, before increasing again.

Other tested intersections exhibit different curve shapes, different positions of local maxima and minima, and different values of delay. However, all tested intersections show a direct relationship between delay on the conflicting approaches and volume on the subject approach, declining values of delay on the opposing approach, and a clear lack of monotonicity.

Although they are not readily seen in Figure 3, multiple discontinuities can exist in the delay/volume relationships. Major culprits are the steps in the HCM's procedure dealing with left turns from shared lanes and the need to decide on an integer number of phases. Discontinuities can be eliminated only by deviating substantially from the HCM and accepted traffic flow theory.

Nonmonotone and discontinuous delay/volume relationships further complicate the assignment algorithm. At best, such messy relationships can introduce additional equilibrium solutions. At worst, they can prevent convergence to any equilibrium solution.

TEST OF ADAPTIVE TRAVEL FORECASTING

The UTOWN network, originally created for testing UTPS, did not contain traffic controls (see Figure 4). This network is known to be hostile to assignment algorithms. It was modified by incorporating signalized intersections and two-way stops (primarily at freeway off-ramps). Otherwise, an attempt was made to keep the two networks as similar as possible.

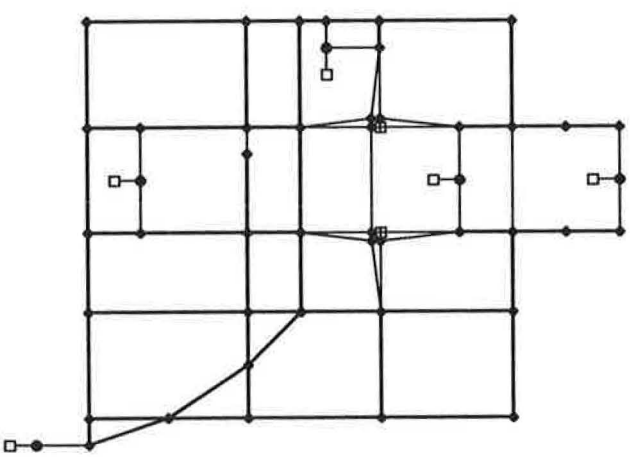

FIGURE 4 Nonadaptive UTOWN network without traffic control.

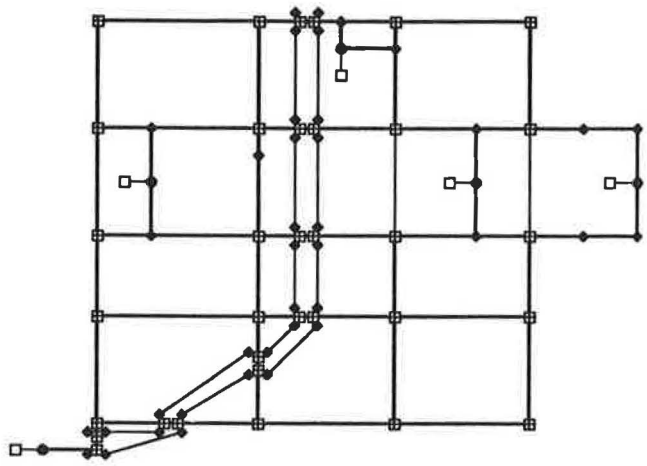

FIGURE 5 Adaptive UTOWN network with traffic control.

The modified, adaptive UTOWN network is shown in Figure 5. Both networks simulated a single peak hour with three trip purposes, but without mode split.

Convergence to an equilibrium solution needs to be checked, but the standard methods derived from Frank-Wolfe decomposition will not work in this case. Since we are interested in a user-optimal assignment, each trip should be assigned to a shortest path between its origin and destination. Therefore, it is possible to determine when equilibrium has been achieved by checking whether the used paths are indeed the shortest paths. A simple test can be devised that compares total travel time between two assignments.

Step 1. Run the model through the desired number of iterations. Obtain estimates of volumes. Recalculate the link travel times. Compute total travel time with the estimates of link volumes and the new travel times.

Step 2. Using the averaged trip table and new travel times from Step 1, run an all-or-nothing assignment. Do not recalculate link travel times. Compute total travel time.

Step 3. Compare the total travel times from Steps 1 and 2. The total travel time from Step 2 will always be the smaller. If they are nearly the same, convergence to an equilibrium solution has been achieved. If they differ significantly, there could be two causes: (*a*) more iterations are required or (*b*) the algorithm failed.

A similar test is provided in UTPS for fixed-demand assignments.

The test was performed on both UTOWN networks for varying numbers of iterations of elastic-demand, incremental assignment. As seen in Table 1, incremental assignment can produce an equilibrium solution on a network with explicit traffic controls. After 200 iterations the difference between Steps 1 and 2 was inconsequential. Equilibrium was effectively achieved after about 20 iterations. A comparison of Table 1 with Table 2 indicates that the adaptive travel forecast converges faster than the nonadaptive one.

A comparison of assigned volumes between the two networks indicated little agreement. The introduction of explicit traffic controls had a large effect on the forecast. In the original nonadaptive network the capacities of many intersections

TABLE 1 Convergence of Incremental Assignment on the Adaptive UTOWN Network

	Total Travel Time		
Iterations	Step 1	Step 2	% Difference
1	1141630	1017512	12.220
2	1288223	1044020	23.339
5	1072852	1018350	5.352
10	1028061	1014823	1.013
20	1012969	1009968	0.297
50	1019219	1015559	0.360
100	1016405	1015892	0.050
200	1015288	1014971	0.031

TABLE 2 Convergence of Incremental Assignment on the Nonadaptive UTOWN Network

	Total Travel Time		
Iterations	Step 1	Step 2	% Difference
1	2188242	1080782	102.468
2	1405340	1086944	29.293
5	1181057	1120440	5.410
10	1162928	1137757	2.212
20	1148820	1117023	2.847
50	1137249	1131264	0.529
100	1133918	1131264	0.235
200	1132658	1130776	0.166

were greatly exceeded even though individual link capacities were maintained.

For the adaptive network, approximately 11 percent of the computation time was devoted to intersection simulations. A greater amount of computation time would be required for a multihour assignment. Memory requirements increased slightly because of the need to keep track of the many turning movements for each all-or-nothing assignment.

One means of judging whether an adaptive assignment tends toward an all-or-nothing assignment is to count the number of used links. Both UTOWN networks only have 20 assignable origin-destination pairs (discounting intrazonal trips), so an all-or-nothing assignment requires less than half of the network. Table 3 gives the percentage of feasible link directions that were assigned a meaningful amount (defined as 1 percent of the saturation flow rate after 200 iterations) of traffic. Freeway ramps in the adaptive network were ignored for consistency. The results suggest that adaptive networks require a smaller percentage of link directions, but not as few as all-or-nothing assignment.

A version of QRS II that contains adaptation has recently been distributed to many users, so a considerable amount of experience is starting to be accumulated. In general, the observations obtained from tests of small networks have been confirmed on full-sized networks.

TABLE 3 Number of Possible and Used Directions

Network	Possible	Assigned	%Assigned
UTOWN Adaptive AON	126	58	46
UTOWN Adaptive Incremental	126	78	62
UTOWN NonAdaptive AON	112	51	46
UTOWN Adaptive Incremental	112	78	70

CHOOSING TRAFFIC CONTROLS

An adaptive travel forecasting model encourages planners to alter the types of traffic controls at intersections, if warranted by the assigned volumes. For instance, a choice could be made between signals and signs. If signs are chosen, a further choice must be made as to which approaches will get them. It is easy to imagine an adaptive network that automates this choice process in accordance with the *Manual on Uniform Traffic Control Devices* (MUTCD).

A simpler rule, often seen in traffic operations models, is to minimize delay. For example, the intersection of Figure 2 would be best served by a two-way stop, not a signal. For the preferred solutions, A and B, signs could be located on the one path with no assigned volume. None of the vehicles would ever be required to stop, so delay is further minimized.

Figure 6 shows total delay at an intersection with three alternative forms of traffic control. Subject and opposing volumes are varied together. Conflicting volumes are held constant at 200 vph, each approach has only one lane, and half of the volume makes a turn. It is seen that the three types of traffic controls perform almost equally well at a volume of 400 vph on the subject and opposing approaches. Below 400 vph the two-way stop is superior; above 400 vph the signal is superior. As expected from the MUTCD, other similar tests indicate that the point at which all controls are equally effective varies with the amount of conflicting volume.

It is outwardly practical to let the model decide on the nature of traffic control, provided that computer resources are sufficient to evaluate each arrangement of signs and signals. Such a model could be made consistent with the MUTCD. Unfortunately, allowing the model to choose signs or signals introduces additional discontinuities in the delay function. The discontinuities would further hamper the search for a good, stable solution.

Assuming that we are able to reasonably and automatically replicate the MUTCD at a single intersection, it is possible for both the alternatives of signs and signals to represent a

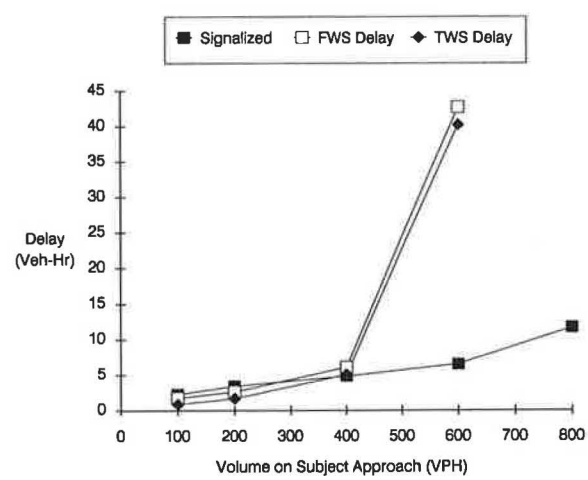

FIGURE 6 Total delay on all approaches for a four-way stop, a two-way stop, and a signal (opposing volume same as subject volume, conflicting volumes at 200 vph, 25 percent right turns, 25 percent left turns, one lane at all approaches, 20 mph speed).

user-optimal equilibrium solution. A signal is not just an upgraded sign; it can influence travelers' paths. Consider an intersection with two-way signing, where the volumes are large on the major street and potentially large on the minor street. Traffic on the minor street is subject to long delays from conflicting traffic, so the assignment algorithm holds the minor street volumes to a level below any of the MUTCD warrants. If the signs are replaced by a signal, the minor street volumes would increase dramatically. The signal is now warranted, simply because it is there. Users of QRS II report that this dilemma arises often enough in real networks to be worrisome.

Future-year networks are influenced by tradition and habit, especially in the placement of traffic control devices. This inertia can be built into the model by using the base-year delays as a starting point for future-year forecasts. For example, the network of Figure 2 is highly sensitive to the starting delays. The slightest delay disadvantage for a given street at the beginning of the algorithm would be sufficient for it to lose all of its traffic. Even if the placement of signs were made to be automatic (moved to the inner loop in Figure 1), the model would be reluctant to rearrange them or to upgrade them to a signal. It may be best to apply adaptive travel forecasting to make incremental adjustments to the current configuration rather than to attempt to consider all possible arrangements of traffic controls.

RELATIONSHIP OF ADAPTATION TO NETWORK CALIBRATION

Planners routinely "calibrate" their nonadaptive networks. That is, they adjust turn penalties, link speeds, and link capacities to obtain better agreement with base-year traffic counts. An extensive calibration exercise eliminates forecast/operational inconsistencies in the base-year network.

The need to calibrate is disturbing in itself; a truly good model should be able to provide accurate forecasts without much fiddling. Even more disturbing is the common practice of using the calibrated penalties, free speeds, and capacities for future-year forecasts. It cannot be assumed that these network attributes will be stable over time.

There are many possible alternatives to calibration for improving the match to existing traffic counts. Better delay/volume relationships would certainly help. The ability to forecast operational characteristics (such as cycle length, phase lengths, and coordination strategies) that affect delay also would help. Using base-year delays as a starting point can introduce some beneficial inertia. Since real highway systems are adaptive, it is certainly a mistake to use a calibration process that would rigidly fix (either explicitly or implicitly) operational characteristics that are known to be variable.

LEVELS OF ADAPTATION

As adaptive travel forecasting gains acceptance, planners will need to seriously consider the appropriate amount of adaptation for their networks. The HCM, of course, does not discuss adaptive travel forecasting, but it does indicate how adaptation can occur. The following levels of adaptation could be invoked, to various degrees, for any given network.

Level 0—no adaptation. Capacity is rigidly fixed on all streets and intersection approaches.

Level 1—low-cost traffic engineering improvements for isolated intersections without changing the type of traffic control. Capacity varies with the amount and nature of conflicting and opposing traffic. (Examples are signal timing and conversion of a through lane to an exclusive lane.)

Level 2—major traffic engineering improvements for isolated intersections. Capacity varies with the amount of and nature of conflicting, opposing, and subject approach traffic. (Examples are installation of signals and relocation of bus stops.)

Level 3—traffic engineering improvements involving a system of intersections. Capacity and delay vary with the nature of surrounding intersections. (An example is signal coordination.)

Level 4—minor geometric changes at isolated intersections. Capacity varies principally with volume on the subject approach. (Examples are adding exclusive lanes, removing on-street parking, and increasing curb radii.)

If all levels of adaptation were fully included in the network, the assignment would be constrained only by cost or by operational limitations, making it similar to the network design problem. With the procedures described here, planners can try to handle Levels 2, 3, and 4 subjectively.

There is no scientific way to determine the levels of adaptation for any given forecast. However, it is reasonable to expect all long-term forecasts to be adaptive to the extent that obvious design flaws or operational deficiencies in the highway system are eliminated. A good working assumption is that continuing efforts will be made to eliminate bottlenecks due to poor geometry, especially those with low-cost solutions. An important implication of adaptation is that planners may be able to ignore many small and isolated reductions in capacity when building their future-year networks.

CONCLUSIONS

Adaptation can provide additional realism to travel forecasts. An adaptive travel forecast requires (*a*) delay relationships that are properly sensitive to traffic controls and (*b*) the ability to modify the way in which the traffic controls are operated. We currently possess sufficient knowledge of traffic flow to allow a network to adapt its operational characteristics to forecast volumes. But we do not yet possess a complete understanding of the effect of adaptation on equilibrium traffic assignments.

An adaptive travel forecast estimates delays at intersection approaches through complex intersection simulations. The HCM contains some well-accepted relationships for this purpose. The implied delay/volume functions are discontinuous and nonmonotonic. Nonetheless, it is still possible (at least some of the time) to obtain acceptable equilibrium solutions with a form of incremental traffic assignment. This technique also permits the network to have elastic demands. A simple test is available to determine whether an equilibrium solution has been reached.

Because an adaptive network can reallocate resources among competing facilities, there need not be a unique solution.

Furthermore, there appears to be a tendency for an adaptive network to generate fewer used paths than a nonadaptive one.

Comparisons of small networks suggest that adaptive networks do not require much more computational effort, appreciably more computer memory, or much more user expertise. There are additional data requirements, but these are tolerable.

Adaptation, including better intersection delay relationships, promises to reduce the independence on network calibration. Adaptation can help avoid the locking of base-year operational characteristics into future-year forecasts.

Further research is needed to determine the prevalence of multiple equilibrium solutions in full-sized networks and the extent to which unwanted solutions can be avoided by setting starting delays consistent with existing traffic. Additional future research should compare the computed intersection delays with those experienced on actual networks.

ACKNOWLEDGMENTS

The author is indebted to Roger Tobin of GTE Laboratories for sharing the proof of the test of user-optimal equilibrium and his knowledge of asymmetric network problems. Portions of this research were supported by the Federal Highway Administration. The author expresses gratitude to the many QRS II users who provided their networks and insights.

REFERENCES

1. M. J. Smith. Traffic Signals in Assignment. *Transportation Research B*, Vol. 19B, No. 2, 1985, pp. 155–160.
2. R. W. Bentley and T. A. Lambe. Assignment of Traffic to a Network of Signalized City Streets. *Transportation Research A*, Vol. 14A, 1980, pp. 57–65.
3. M. J. Smith and M. Ghali. The Dynamics of Traffic Assignment and Traffic Control: A Theoretical Study. *Transportation Research B*, Vol. 24B, No. 6, 1990, pp. 409–422.
4. P. Marcotte. Network Optimization with Continuous Control Parameters. *Transportation Science*, Vol. 17, No. 2, May 1983, pp. 181–197.
5. D. R. Leonard and P. Gower. *User Guide to CONTRAM Version 4*. Transportation and Road Research Laboratory, Supplementary Report 735, 1982.
6. T. L. Magnanti and R. T. Wong. Network Design and Transportation Planning: Models and Algorithms. *Transportation Science*, Vol. 18, No. 1, Feb. 1984, pp. 1–55.
7. C. S. Fisk. A Conceptual Framework for Optimal Transportation Systems Planning with Integrated Supply and Demand Models. *Transportation Science*, Vol. 20, No. 1, 1986, pp. 37–47.
8. *Special Report 209: Highway Capacity Manual*. TRB, National Research Council, Washington, D.C., 1985.
9. A. J. Richardson. A Delay Model for Multiway Stop-Sign Intersections. In *Transportation Research Record 1112*, TRB, National Research Council, Washington, D.C., 1987, pp. 107–112.
10. M. Kyte. Estimating Capacity of an All-Way Stop-Controlled Intersection. In *Transportation Research Record 1287*, TRB, National Research Council, Washington, D.C., 1990.
11. *Transportation Research Circular 373: Interim Materials on Unsignalized Intersection Capacity*. TRB, National Research Council, Washington, D.C., July 1991.
12. A. Hansson. Swedish Highway Capacity Manual: Part 2, Capacity of Unsignalized Intersections. In *Transportation Research Record 667*, TRB, National Research Council, Washington, D.C., 1978, pp. 4–11.
13. K. G. Baass. The Potential Capacity of Unsignalized Intersections. *ITE Journal*, Oct. 1987, pp. 43–46.
14. L. LeBlanc, E. Morlock, and W. Pierskella. An Efficient Approach To Solving the Road Network Equilibrium Traffic Assignment Problem. *Transportation Research*, Vol. 9, 1975, pp. 309–318.
15. S. P. Evans. Derivation and Analysis of Some Models for Combining Trip Distribution and Assignment. *Transportation Research*, Vol. 10, 1976, pp. 37–57.
16. W. B. Powell and Y. Sheffi. The Convergence of Equilibrium Algorithms and Predetermined Step Sizes. *Transportation Science*, Vol. 16, 1982, pp. 45–55.
17. S. Dafermos. Traffic Equilibrium and Variational Inequalities. *Transportation Science*, Vol. 14, No. 1, 1980, pp. 42–54.
18. A. J. Horowitz. Convergence Properties of Some Iterative Traffic Assignment Algorithms. In *Transportation Research Record 1220*, TRB, National Research Council, Washington, D.C., 1990, pp. 21–27.
19. A. J. Horowitz. Tests of an Ad Hoc Algorithm for Elastic-Demand Equilibrium Traffic Assignment. *Transportation Research B*, Vol. 23, No. 4, 1989, pp. 309–313.
20. A. J. Horowitz. Convergence of Certain Traffic and Land-Use Equilibrium Assignment Models. *Environment and Planning A*, 1991, in press.
21. S. Lewis and S. McNeil. Developments in Microcomputer Network Analysis Tools for Transportation Planning. *ITE Journal*, Oct. 1986, pp. 31–35.
22. B. B. Heydecker. Some Consequences of Detailed Junction Modeling in Road Traffic Assignment. *Transportation Science*, Vol. 17, No. 3, Aug. 1983, pp. 263–281.

Publication of this paper sponsored by Committee on Highway Capacity and Quality of Service.