Integrated Approach to Vehicle Routing and Congestion Prediction for Real-Time Driver Guidance

ISAM KAYSI, MOSHE BEN-AKIVA, AND HARIS KOUTSOPOULOS

The generation and dissemination of driver guidance that can be used for real-time diversion of traffic are expected to be implemented through the use of real-time traveler information systems. To implement these functions, a system structure consisting of a surveillance module, a congestion prediction module, and a control and routing (CAR) module is proposed, with the focus on the approaches that may be used for congestion prediction and the strategies that may form the basis for routing. It is argued that a congestion prediction capability is critical for the effectiveness of an on-line traveler information system. Such a capability is required to accurately forecast traffic conditions that may exist in the near future. The use of a dynamic traffic assignment model for congestion prediction is suggested. Such a model consists of dynamic driver behavior and network performance modules as well as origin-destination updating capability. Alternatively, statistical time-series methods may be necessary to generate predictions of future traffic conditions. The advantages and difficulties of adopting either approach are discussed. The predicted congestion information is passed to the CAR module to develop diversion strategies to alleviate both recurring and nonrecurring congestion. The role of routing strategies and update frequency in determining guidance effectiveness is discussed.

During the coming decades, efficient operation of existing road networks is expected to be achieved through dynamic traffic management schemes that make use of available and anticipated advanced technologies. Within this context, intelligent vehicle highway systems (IVHS) are currently being developed. These IVHS systems envision the linking of road infrastructure, vehicles, and drivers using advanced communication technology, computers, information display equipment, and traffic control systems.

In this context, advanced traveler information systems (ATIS) that are based on modern information technology may play an important role in reducing traffic congestion and improving traffic flows and safety. It is expected that ATIS will reduce delays caused by both incident and recurrent congestion by providing information to motorists about alternative paths to their destinations using a combination of roadside signals and onboard systems. Such schemes will aim at optimizing driver route selection and making this selection responsive to real-time road and traffic conditions. In this paper a framework for the operation of ATIS is presented and the modeling requirements of its constituent elements are analyzed.

I. Kaysi, Department of Civil Engineering, American University of Beirut, P.O. Box 11-0236, Beirut, Lebanon. M. Ben-Akiva and H. Koutsopoulos, Department of Civil Engineering, Massachusetts Institute of Technology, Cambridge, Mass. 02139.

INTEGRATED FRAMEWORK FOR IMPLEMENTATION OF ATIS

The actual benefits realized from traveler information systems depend heavily on the quality of the traffic information provided to drivers (*1*). This section describes the framework being proposed to provide drivers with guidance that they can have confidence in and that, as a result of improved information, can eliminate the occurrence of adverse impacts [see Ben-Akiva et al. (*2*)]. The discussion that follows describes the system structure and information flow embodied in the proposed framework. Later sections present the principles behind the proposed framework and describe the control and routing (CAR) and congestion prediction (COP) modules.

System Structure

A dynamic network modeling approach is critical to the effectiveness of real-time traveler information systems. Such a modeling approach is needed to accurately assess network performance as well as to forecast traffic conditions that may exist in the near future to develop real-time diversion strategies to alleviate both recurring and nonrecurring congestion conditions.

A proper framework for ATIS implementation should be able to integrate the functional needs referred to above into an operational system. Figure 1 illustrates the system structure and information flow of the framework within which real-time ATIS should be implemented (*3*). The functions performed by each element in the system are briefly described:

- The surveillance system consists of traffic sensors deployed on the various network elements (for example, detectors in the pavement, video cameras, possibly other optical recording equipment). The surveillance system may also include roadside readers that gather information about vehicles that are passing selected points on the network over time. Thus, equipped vehicles themselves may act as elements of the surveillance system by providing information on travel times on specific sections of the network. The collected data may consist of information on flows, speeds, travel times, the numbers of queued and moving vehicles on each link, and incident detection. The actual information gathered will vary from system to system, depending on the particular components included in the system and the coverage of the network by the surveillance system.

- The COP element has the responsibility for providing the control and routing module with the information that is needed to implement routing and guidance strategies. Among the principles adopted in this paper are the need for COP to provide CAR with projected traffic conditions and the fact that COP should be performed by a dynamic traffic assignment (DTA) model that will take into account in its projections the driver response.
- The CAR element generates guidance advice in response to information provided by the surveillance system and by the congestion prediction model. The fact that CAR should maintain projection or guidance consistency constitutes another principle adopted in this paper.

Information Flow

The flow of information from one element in the system to another is described as follows (refer to Figure 1):

- Infrastructure data: Infrastructure data, an umbrella term, includes all network attributes that generally are invariant with time or that change slowly with time. The attributes include the network topology, the geometric attributes of all network elements, the control devices that are installed in the network, any channelization or other type of lane control, circulation restrictions such as one-way streets and prohibited turns, and so on. This information is required by the COP and CAR.
- Historical origin-destination (O-D) data: The historical O-D data generally consist of O-D information obtained by surveys or inferred from traffic counts by assignment models. This historical information is, for the most part, slowly changing over time but should be updated periodically.

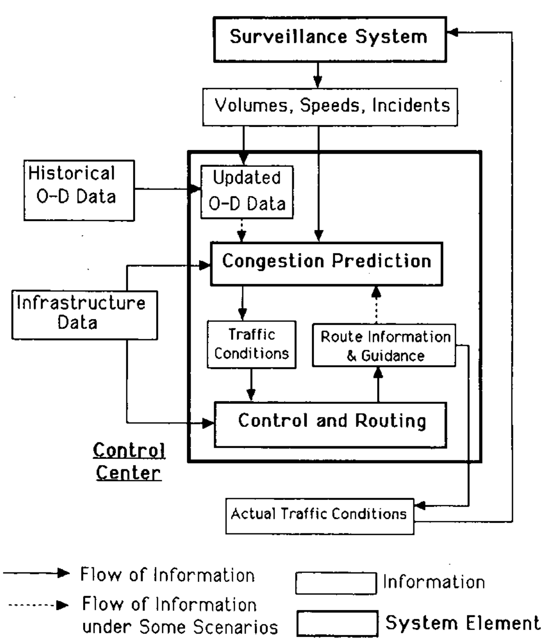

FIGURE 1 Proposed framework.

- The surveillance system: The information provided by the surveillance system consists mainly of direct measurements of volume, speed, occupancy, and the presence of incidents. Eventually the information may also include travel time data from equipped vehicles.
- Updated O-D data: The most recent information from the surveillance system and route information (if available) can be combined with historical O-D data to provide updated three-dimensional (3-D) O-D matrixes for the subsequent time periods.
- COP: Congestion prediction provides estimates of traffic conditions on the network. The updated O-D data are used by some COP scenarios including the proposed DTA. Because route choice modeling and the provision of guidance are sensitive to the destinations associated with flows, O-D data are required to implement the DTA associated with the proposed framework.
- CAR: The traffic conditions identified by the COP are transmitted to the CAR. The routing strategies need this information to develop an optimal response to the developing traffic environment.
- Guidance: The outputs of the CAR generally take the form of route guidance information. Within the proposed framework, guidance data are transmitted to the COP as an input.

PRINCIPLES UNDERLYING THE PROPOSED FRAMEWORK

The following are the major principles underlying the proposed framework:

Principle 1: COP Should Provide CAR with Projected Traffic Conditions

The travel times used for routing purposes by ATIS may be based on historical, current, or predicted traffic conditions. Although the use of historical data may provide a basis for static guidance and navigation, its use alone as a basis for CAR decisions is not expected to be of any value for adaptive routing. The main reason behind this is that historical traffic data are a bad indicator of evolving, day-specific traffic conditions, especially in situations in which traffic patterns display a significant amount of day-to-day variability. An analysis of the performance of various real-time routing strategies by Koutsopoulos and Xu (*4*) confirmed this intuition and indicated that the use of historical data as a basis for real-time routing advice is significantly inferior to the use of, for example, current or predictive information (*5*).

French (*6*), Catling and McQueen (*7*), and Rillings and Betsold (*8*) provide reviews of many existing demonstration projects. In many such projects guidance passed to drivers consists of information regarding current traffic conditions (Smart Corridor, AMTICS, and RACS). Some researchers assert that routing strategies may be formulated on the basis of a control-theoretic approach that requires information on current traffic conditions only. For example, Papageorgiou and Messmer (*9*) use feedback control methods to split traffic between an O-D pair among different routes. They claim that their methodology has low sensitivity with respect to unknown future demand levels and compliance rates that are assumed to be exogenous "disturbances." However, the authors warn that their feedback concept, which is based on observations of current traffic conditions, may not achieve its goal of establishing dynamic user optimum conditions if strong oscillations in the demand levels occur or if the network performance displays strong nonlinearities in case of severe congestion. These remarks by the authors provide further evidence that using current traffic conditions as a basis for guidance will not succeed if current traffic conditions are not good predictors of future conditions.

One of the major principles embodied in the framework being proposed in this paper is that such routing strategies have to be formulated on the basis of a forecast or projection of future traffic conditions on the network (*2*) rather than on instantaneous traffic conditions. The rationale behind using predictive information is that drivers' travel decisions are affected by future traffic conditions expected to be in effect when they reach downstream sections of the network on their way to their destinations. The ALI-SCOUT system uses projected travel times in setting the guidance on the basis of a similar rationale. Therefore, the most useful type of guidance that can be provided to a driver faced with travel decisions would be based on a projection of traffic conditions. In addition, guidance based on traffic information that is predicted using an advanced COP module is most capable of improving the travel time reliability of drivers because its look-ahead capability helps them avoid long future delays. This issue is discussed in more detail under the second principle.

Principle 2: A DTA Model Should Be Used for COP

In all existing traveler information systems and demonstration projects, the guidance passed to drivers is based either on current traffic conditions (Smart Corridor, Travtek, AMTICS, and RACS) or on simple predictions of future traffic conditions (ALI-SCOUT).

The travel time prediction methodology used by the ALI-SCOUT system, for example, constructs a projection ratio of the historical travel time on a specific link to the current travel time, as reported by equipped vehicles (*10*). This ratio is used to predict travel times for vehicles using that link during all future time intervals. Because only a few vehicles are equipped in Berlin, many links are not used by equipped vehicles during particular time intervals; therefore, no estimate of the projection ratio would be available. Consequently, the ratio is modified to reflect current conditions on neighboring links as well as conditions in preceding time intervals. Koutsopoulos and Xu (*4*) note that a problem with this methodology is the fact that the projection ratio is used to predict travel times for all future time intervals, thus implicitly assuming that trends currently observed remain constant for the entire prediction interval. To remedy this particular problem, Koutsopoulos and Xu suggest the use of information discounting. However, the methodology remains heuristic in nature and suffers from other omissions, which are discussed next.

In all these systems and projects, as well as in all analyses being conducted by researchers related to ATIS (*9,11*), there has been no consideration whatsoever of the response of motorists to route guidance in setting such guidance. Such an omission entails a major shortcoming in that the potential concentration of traffic on the recommended routes and the overreaction of drivers in their response to guidance information are ignored. This problem is expected to become more severe as the number of guided vehicles increases. Guidance validity and, as a result, driver compliance would be adversely affected in such schemes.

To overcome this problem it is required that the guidance be based on an advanced COP module that makes its predictions of future congestion in the network on the basis of

- Current traffic conditions (consideration of initial conditions),
- Predicted O-D demand levels (sensitivity to future demand patterns),
- Guidance being provided and anticipated driver response to guidance,
- Traffic control actions to be implemented, and
- Reduction in capacity as a result of incidents that have been detected.

Principle 3: CAR Should Maintain Guidance/ Prediction Consistency

On the basis of the earlier discussion, it becomes clear that none of the guidance systems in existence attempt to anticipate the impact of guidance being provided. The same holds true for analyses conducted by researchers in relation to ATIS. A major principle underlying the proposed framework is that consistency has to be maintained between the guidance being provided to drivers and the predicted traffic conditions. That is, the information system has to check that the guidance being provided will prove to be optimal to guided drivers on the basis of a prediction of the future traffic conditions. This system would result in guidance information that is consistent with anticipated driver behavior and network conditions, would ensure the validity of the guidance information, and would encourage its use by more drivers. In addition, consistency has to be maintained between the traffic conditions as perceived by the COP and the actual traffic conditions so that the COP remains attuned to the real world.

CAR STRATEGIES

When specifying the CAR module, three major issues have to be addressed:

1. What logic should the CAR module adopt to provide drivers with guidance advice? Should information or route directives be provided to drivers? Should route directives be based on shortest-path guidance or on route distributive guidance? How should route distributive guidance be implemented? How can the CAR logic ensure guidance and projection consistency?

2. Given that guidance provided at specific locations of the network can be updated periodically, what is the impact of temporal update frequency on the effectiveness of routing strategies?

3. If drivers receive guidance updates at various locations as they move through the network toward their destinations (as in ALI-SCOUT), what is the impact of spatial update frequency on the effectiveness of routing strategies?

A discussion of each of these issues follows.

CAR Logic

The most basic distinction between various logic types used as part of CAR relates to whether CAR provides drivers with information or route directives.

Information

If information is passed to drivers then CAR would constitute a simple link whereby the traffic data output from COP is interpreted, relayed to drivers, and presented in an understandable form. This represents the case in many demonstration projects, such as AMTICS, RACS, and Smart Corridor, and has served as the basis for ATIS analyses conducted by several researchers [see, for example, Mahmassani and Jayakrishman (*11*)].

Route Directives

Single Route Guidance Versus Route Distributive Guidance When route directives are employed, specific routing strategies may be based on the assignment of all traffic during a guidance interval to one path (usually the shortest path), whereas a different policy may be to distribute traffic over a number of alternate routes leading to the desired destination. Distributing traffic would entail the determination of the optimal fraction of traffic to be guided to the different routes, a process that may require significant computational effort in real time.

Jeffrey et al. (*12*) indicate their concern that guidance provided by AUTOGUIDE may overload a single "best" route to which drivers are guided. To deal with this problem they suggest that the "closeness" of journey times along alternate routes can be used to determine what proportion of vehicles should be sent along each route.

Papageorgiou and Messmer (*9*) used feedback regulation in an attempt to establish a dynamic user optimum whereby various routes used by flow departing at the same time between an O-D pair have the same travel time, and no unused path between the same O-D pair has a shorter travel time. The feedback regulator would observe current traffic conditions and determine the way traffic should be split among various paths connecting O-D pairs, with the aim of establishing the dynamic user optimum conditions. The first feedback law adopted consisted of shortest-route guidance. Vehicles traveling between a specific O-D pair were guided to the route currently having the shortest travel time for the O-D pair under consideration. However, the authors indicated that such logic may lead to strong perturbations of traffic flow, especially if a large fraction of vehicles are equipped. Therefore, they suggested an alternative logic based on a "smooth regulator" using a more advanced feedback law that leads to route-distributive guidance and the specification of optimal rates of splitting traffic among alternate routes.

Difficulties in Implementing Route Distributive Guidance Depending on the specific technology used for providing drivers with route directives, it may not be technically possible to distribute traffic over a number of routes at the same instant of time. For example, if variable message signs are used, the best that can be accomplished is to change the sign within the guidance interval so that the time average of the route directives would correspond to the fractions we wish to achieve for each route. On the other hand, if in-vehicle units are used, it is technically possible to provide different drivers with different route directives at the same time to split drivers among routes according to the optimal fractions. However, this process may not be politically or legally acceptable because of its inequity implications, and a scheme similar to what was suggested for variable message signs may have to be adopted.

In this paper it is proposed that shortest-route guidance be used whenever possible unless it does not succeed in maintaining projection/guidance consistency, in which case it becomes necessary to modify the CAR logic. This issue is discussed next.

Maintaining Guidance/Projection Consistency

It is important to recognize that there is a closed loop within the control center computations in the diagram of Figure 1. The loop involves the following: the COP sends data on projected traffic conditions to CAR, which in turn sends back to the COP information that defines the guidance environment that is in effect at specific times. This guidance information should be taken into account by the COP when projecting traffic conditions for the next forecasting period.

Specific route guidance should be provided only after going through the loop and having the CAR ensure that the guidance is consistent with traffic conditions projected by the COP. The guidance that is provided to drivers should represent a fixed-point solution of the COP/CAR interaction. For example, if shortest-path guidance is being used, consistency would be achieved if traffic conditions and travel times predicted by the COP indicate that vehicles are guided to the route that is predicted to be shortest.

If consistency is not achieved, the guidance advice has to be revised. Future traffic conditions resulting from the revised guidance have to be predicted by the COP and consistency with the provided guidance checked again. This loop may need to be repeated several times until consistency is achieved. If shortest-path guidance is being used then the scheme described above may not converge, and it may not be possible to achieve COP/CAR consistency. This is particularly true if the guidance is not updated frequently and therefore COP predictions take into account the response of a large number of drivers to guidance measures, with possible oscillations in potential messages and route use between subsequent executions of the COP/CAR cycle. Such oscillations and the implied lack of convergence would indicate that the control center is not capable of managing and mitigating the effects of the potential overreaction that will occur as a result of provision of guidance. This necessitates that the CAR logic using shortest-path guidance be modified to control the occurrence of overreaction and to deal with COP/CAR inconsistency.

Consistency Check Whenever the COP/CAR consistency check referred to above indicates a lack of convergence using single-route guidance, the travel time on the alternate routes would not be sufficiently different to warrant sending all guided traffic to the shorter route without causing overreaction. In such a case either the CAR logic should increase the temporal update frequency or guided vehicles should be distributed in such a way that they do not concentrate on any single one of the alternate routes.

Route Directives with Guidance Threshold The practice in existing route guidance experiments and studies of guidance being provided to a specific route whenever the travel time on such a route is evaluated to be shorter than that on alternate routes, even by a very small amount, often causes the guidance/projection inconsistency. As was indicated above, if all guided vehicles are sent to a route that is only marginally shorter than the alternate routes, overreaction is likely to occur. The use of a guidance threshold is suggested when route directives are in effect so that guidance is provided only if the travel time on the route to which drivers are directed is shorter than that on the alternate routes by a specific threshold. In the absence of a route that is shorter than the alternate routes by at least this threshold, guided vehicles have to be distributed over alternate routes. Such a scheme would have the potential to reduce overreaction associated with improved information and may be implemented on the basis of current or projected traffic conditions. Although it does not specifically identify occurrences of inconsistency as the "consistency check" scheme would, this scheme offers a computational advantage over that approach because a full projection/guidance consistency check (which requires a number of iterations and may involve checks over a long projection horizon) is not necessary.

Implementation In the case of collective route guidance technologies such as variable message signs the two schemes described above may implement "vehicle distribution" by simply providing "no guidance" and leaving drivers to distribute themselves as they would under normal traffic conditions. Such "natural" distribution is expected to yield good results because it would not impose excessive loads on any of the alternate routes. In addition, when it is called for, it would eliminate the computational and communication requirements associated with the determination of the optimal splitting fractions associated with a route distributive strategy, and the actual provision of such guidance.

In the case of in-vehicle units, on the other hand, drivers expect to receive route guidance at all times. One way of distributing guided vehicles in this case is to provide guidance in such a way that the time average of the directives relating to specific routes is equal to the average fractions of vehicles that would take each of these routes under normal traffic conditions. This would eliminate the computational effort required to determine the optimal splitting fractions, but guidance still needs to be communicated to the in-vehicle units.

Temporal Update Frequency

The frequency of guidance updates may influence the relative effectiveness of CAR schemes. If updating is infrequent, there is a real danger that a CAR logic such as shortest-path guidance may overload the shortest path and cause overreaction. In this case route distributive guidance might be necessary, at a possibly significant computational cost. On the other hand, if guidance is updated frequently, shortest-path guidance might perform satisfactorily and route distributive guidance might not be called for. The trade-off here is between the extra computational efforts involved in (a) computing the splitting ratios for distributive guidance and (b) providing more frequent updates in the case of single-route guidance. For instance, similar computational efforts may be required to provide distributive route guidance every 2 min or single-route guidance every 1 min for specific ATIS operating conditions. In that case, a comparison of the effectiveness of guidance provided in each of the two cases would be of interest. However, the specifics of this trade-off are not likely to become clear except after some experience with actual ATIS operations.

In addition, the guidance frequency may mitigate the impacts of using a less advanced information basis for CAR. For example, use of current traffic conditions as a basis for CAR would require very frequent updates of route guidance to avoid situations of serious overreaction in which congestion shifts from one location to another (*2*).

Spatial Update Frequency

The temporal update frequency discussed above relates to updates at a specific guidance location. Another consideration in the implementation of CAR strategies is the availability of spatial updates for a specific vehicle that is en route to its destination. Depending on the network structure, it may be possible for such a vehicle to receive guidance at more than one location during its trip. This would provide the CAR with possibilities of revising and correcting guidance that was provided to vehicles at an upstream location but that is no longer valid because of emerging traffic conditions (such as the occurrence of accidents). Such possibilities of spatial updates are potentially significant in improving the effectiveness of various CAR strategies. Koutsopoulos and Xu (*4*) observe that as the spatial update frequency increases, the adverse effects that were observed with low temporal update frequency at high guided fractions are somewhat alleviated. Clearly, this observation is consistent with our a priori expectations. A high degree of spatial update frequency is also likely to make an advanced COP module less necessary.

However, the benefits from a high spatial update frequency depend to a large degree on the network structure. Specifically, the network structure has to provide opportunities for reguidance and route diversion if the spatial update of guidance is to have any effect. For instance, for two routes in parallel with no crossovers, updating guidance along the route when it is not possible to switch to an alternate route would be worthless.

POSSIBLE COP SCHEMES

Types of Information Provided by COP to CAR

As discussed earlier, there exist two main types of information that could be provided by COP to CAR for use by routing strategies associated with ATIS. Some researchers hold the view that proper routing strategies require information on current traffic conditions only, whereas others assert that such routing strategies have to be formulated on the basis of a forecast or projection of future traffic conditions on the network.

COP Models

For the approach that bases its routing strategies on projected traffic conditions, there are the additional questions as to what projection methodology to adopt and the required accuracy of such methodology. These questions will be discussed next.

Conceptually, this scheme is more appealing although its implementation is more complicated. Two possible methodologies are presented below.

1. The first approach consists of a DTA model. This approach is described in more detail later. Briefly, the DTA provides the ability to project traffic conditions while taking into account potential driver response to guidance and the predicted time-dependent O-D matrixes. The DTA involves a prediction of driver behavior given the availability of guidance, an assignment of the predicted time-varying path flows to the network, and the subsequent determination of resulting flows and congestion on the various links of the network by time of day. To provide such capabilities, the DTA requires updated predictions of O-D demands, a dynamic driver behavior module, and a dynamic network performance module. The provision of guidance on the basis of projected traffic conditions is quite difficult in congested networks because such conditions are dependent on the ways in which drivers respond to the information. In other words, the validity of predicted network conditions depends on their consistency with current and future drivers' choices, which depend on the drivers' use of such information. This scheme would also carry with it a simultaneity problem to ensure consistency between provided guidance and projected traffic conditions.

2. Another approach to congestion prediction that involves significantly less computational and hardware requirements is the use of statistical time-series methods. Such an approach, which has been used for adaptive traffic control systems [see, for example, Stephanedes et al. (*13*) and Okutani and Stephanedes (*14*)], would make use of historical as well as recent traffic observations from the surveillance system to come up with congestion predictions. This approach is appealing because it does not require detailed modeling of O-D patterns, network structures, or driver response. However, its validity may be limited to relatively short projection periods. On the other hand, the behavioral models embedded in the DTA are more capable of capturing changes in traffic conditions over longer projection horizons. Both schemes are worthy of further investigation, and it might be that the best approach would be some combination of the two.

CONGESTION PREDICTION USING DTA

Proposed Approach and Elements of DTA

Some of the basic ideas related to the DTA approach proposed here were developed in previous research (*15–17*) that has been extended by Vythoulkas (*18*) to the case of a general network. A similar approach has also been suggested by Cascetta and Cantarella (*19*).

Because the DTA approach is proposed to serve as the COP module that will be used to predict the effects that travel decisions by informed drivers may have on overall traffic conditions, the approach should explicitly treat the distribution of traffic by time of day and drivers' pretrip and en route adjustments. This approach requires significant enhancements to the driver behavior and network performance models suggested in the research mentioned above.

The proposed DTA approach incorporates the following items:

- Dynamic driver behavior modeling,
- Dynamic network performance modeling, and
- O-D updating.

Dynamic Driver Behavior Modeling

A detailed discussion of dynamic driver behavior modeling for the DTA is not within the scope of this paper. At this point it is sufficient to say that the DTA approach has to capture the potential effects of the new information services on the departure time as well as pretrip and en route path choices of individual drivers. Therefore, models of the dynamic choices available to drivers with access to guidance should reflect the fashion in which new information concerning traffic conditions may affect driver behavior [see, for example, Kaysi (*20*) and Lotan (*21*)].

Dynamic Network Performance Modeling

The function of the DTA is to provide CAR with projected travel times on various links in the network. To perform this task, the DTA has to take into account the fact that traffic flows and network capacities vary with time. Temporal variations in traffic flows at various points of the network are caused by

- The time-varying nature of travel demand,
- The response of drivers to anticipated traffic congestion, and

• The delays at upstream bottlenecks that affect the arrival times at downstream facilities.

On the other hand, temporal variations in network capacities include primarily changes in capacities caused by incidents and traffic control actions.

Therefore, in a dynamic traffic assignment model all the following variables are space and time dependent: O-D trips, link capacities, link volumes, and link travel times.

Flow Computation Difficulties

Drivers' travel choices on a particular day (i.e., departure times and routes) translate into specific time-dependent path flows on the network. The dynamic network performance model of the DTA, whose elements are discussed below, has to determine time-dependent link flows and travel times that are consistent with predicted path flows. For simple networks (such as a number of routes in parallel between one O-D pair) this procedure is straightforward. However, for general networks, the correspondence between arc flows and path flows in a dynamic network is not as trivial as in the case of "steady-state" networks in which the two are simply related by a link/path incidence matrix. In dynamic networks, flow on any link during a specific interval is composed of path flows leaving their origin in that interval or in previous intervals and traversing (totally or partially) the link during the specified interval. Moreover, because of congestion, the time required to cross any link during a specific interval depends on the link flow during that interval and, for oversaturated links, on link flows in previous intervals.

Algorithms of varying complexity and levels of detail could be used to deal with this problem. These algorithms range from microsimulation in which individual cars are tracked through the network to macrosimulation algorithms that extend the familiar static network assignment algorithms to cope with time-varying demand and travel times. Many of the algorithms formulated to compute link flows from path flows in the case of within-day dynamic traffic conditions require the solution of a "fixed-point problem." Cascetta and Cantarella (*19*) solved the fixed-point problem in their DTA model by using meso-simulation, in which groups of vehicles are tracked through the network.

Dimensions of Dynamic Network Performance Models

The most important dimensions of network performance models consist of the time representation (continuous or discrete), flow representation (continuous flow, micro level, or macro level), and travel time computation. An analysis of the above items, including the proposed packet approach as well as an illustration of these ideas on a prototypical network and a discussion of the self-calibration aspect of the DTA, may be found elsewhere (*20*).

Updating the 3-D O-D Matrix

One set of inputs required by the dynamic traffic assignment models that are to be constructed are average (historical) O-D flows by time slice. These O-D matrixes differ significantly from the matrixes available for use in transportation planning analyses, for example. The required matrixes (representing dynamic flows) entail a much more refined description of time-varying travel demand between sufficiently small zones that represent the sources and sinks of the network. These dynamic O-D matrixes may be obtained from historically observed time-of-day flows using an off-line DTA.

Data that describe time-of-day variability in link flows or in O-D flows are almost nonexistent at this time. In other words, the 3-D O-D data required to implement the DTA are mostly unavailable, and it probably would take some time before a data base including the required type and level of detail of data were established. This process, however, is likely to be expedited by the substantial improvements in communication and computing devices and the resulting advances in traffic surveillance equipment.

Real-time traffic flow data obtained from the surveillance system are combined with the historical O-D matrix referred to above. The outcome of this process is an updated 3-D O-D matrix that is to be used in the next DTA projection period. The frequency of such updates will depend on the capabilities of the surveillance system to provide new extensive flow data as well as on the computational requirements of the updating process.

The problem of estimating O-D flows in real time on the basis of historical and recent measurements of traffic flows has received increased attention in recent years in view of its applicability for real-time traveler information and traffic control systems. The problem may be applied at the level of the network or individual junctions. Real-time estimation of O-D flows would require measurements of traffic flows at the entrances and exits of intersections. It is also useful to have measurements of travel times between counting locations. These measurements may obviously be obtained by the surveillance system. The aim is to estimate the proportion of traffic that goes from each origin to each destination.

DTA Approach

The dynamic O-D updating procedure would involve an iterative application of an off-line DTA. In this process a historical matrix would be adjusted to reflect the latest measurements from the surveillance system. Thus, the development of a real-time DTA would also benefit the process of dynamic O-D updating. This procedure is analogous to methods of static O-D estimation using a static equilibrium assignment.

Unfortunately, literature is not yet published on dynamic O-D estimation with a DTA. The existing literature on dynamic O-D estimation considers simple situations in which the distributions of travel times between counting stations are known. As such, in addition to unavailability of historical 3-D O-D data required to implement the DTA, the analytical tools required to provide predictions and updates of the O-D matrix have not yet been fully developed. The development of such tools will constitute part of the overall research effort required to implement ATIS.

Methodologies without a DTA that have been proposed to obtain the dynamic O-D estimates fall into two major categories: least-squares estimation and the Kalman filtering procedure.

Least-Squares Estimation

Variations on the least-squares approach (*22–24*) include constrained estimation whereby constraints are imposed on permissible values of O-D flows to ensure that fitted values of O-D flows are consistent with flow conservation and nonnegativity constraints; discounted estimation whereby recent observations are given more weight, a procedure that is necessary when the proportions of flow destined from a specific network entry point to various exits vary over time; and recursive least-squares estimation to improve the estimation efficiency.

Kalman Filtering Procedure

The Kalman filtering formulation is essentially a method developed for signal processing that provides optional estimates of the current state of a dynamic system. It consists of two equations: a transition equation and a measurement equation. The transition equation describes how the set of state variables (the parameters to be estimated) changes over time. The measurement equation provides predicted values of some observed variables on the basis of the (estimated) state variables. A comparison of the predicted and measured values of the observed variables provides a basis for updating estimates of the state variables for the next interval.

The application of the Kalman filtering approach to the problem at hand would have O-D flows represented as state variables that are to be estimated while the observed variables are the traffic counts. Note that this approach may be combined with a DTA as follows: dynamic traffic assignment, using the estimated O-D flows and playing the role of the measurement equation, will provide predicted values of traffic counts at specific stations. These predicted traffic counts are then compared with the actual observed counts at the stations. The prediction error detected from this comparison is used in the transition equation to estimate new values of the O-D flows for the next interval.

CONCLUDING REMARKS

This paper presented a possible system structure for the implementation of ATIS. The system consists of a surveillance system, a congestion prediction module, and a control and routing module. The paper illustrated the functional requirements of these modules and the interactions and information flows among them. The analysis indicated that models different from existing ones are needed to provide the functional capabilities of ATIS.

It was also observed that the occurrence and severity of overreaction may be reduced either by (a) basing the guidance on predictive traffic information that accounts for driver reaction, (b) providing guidance in such a way that traffic is distributed over several reasonable paths, or (c) performing more frequent guidance updates.

The analysis also identified several areas worthy of further research and investigation, including the formulation of more detailed specifications of the component modules and the possible need for algorithms for the implementation of optimal route distributive guidance. Finally, looking further ahead, prototypical tests have to be designed and conducted to examine the actual operation of the proposed system in the real world.

REFERENCES

1. P. W. Bonsall and A. D. May. Evaluation of In-Vehicle Route Guidance Systems. Presented at 5th WCTR, Yokohama, Japan, 1989.
2. M. Ben-Akiva, A. de Palma, and I. Kaysi. Dynamic Network Models and Driver Information Systems. In *Transportation Research*, Vol. 25A, No. 5, 1991, pp. 251–266.
3. M. Ben-Akiva and I. Kaysi. *Congestion Prediction Using Dynamic Traffic Assignment.* Working Paper. Department of Civil Engineering, Massachusetts Institute of Technology, Cambridge, Aug. 1991.
4. H. Koutsopoulos and H. Xu. An Information Discounting Routing Strategy for Advanced Traveller Information Systems (ATIS). *Transportation Research*, in press, 1992.
5. I. Kaysi, M. Ben-Akiva, and A. de Palma. Design Aspects of Advanced Traveller Information Systems. Presented at 2nd International Capri Seminar on Urban Traffic Networks, Capri, Italy, 1992.
6. R. L. French. Intelligent Vehicle/Highway Systems in Action. In *ITE Journal*, Nov. 1990, pp. 23–31.
7. I. Catling and B. McQueen. Road Transport Informatics in Europe—Major Programs and Demonstrations. In *IEEE Transactions on Vehicular Technology*, Vol. 40, No. 1, 1991, pp. 132–140.
8. J. H. Rillings and R. J. Betsold. Advanced Driver Information Systems. In *IEEE Transactions on Vehicular Technology*, Vol. 40, No. 1, 1991, pp. 31–40.
9. M. Papageorgiou and A. Messmer. Dynamic Network Traffic Assignment and Route Guidance Via Feedback Regulation. In *Transportation Research Record 1306*, TRB, National Research Council, Washington, D.C., Jan. 1991.
10. G. Hoffman and J. Janko. Travel Times as a Basic Part of the LISB Guidance Strategy. Presented at IEE Road Traffic Control Conference, London, 1990.
11. H. Mahmassani and R. Jayakrishnan. System Performance and User Response Under Real-Time Information in a Congested Traffic Corridor. Presented at Seminar on Urban Traffic Networks, Capri, Italy, June 1989.
12. D. J. Jeffrey, K. Russam, and D. I. Robertson. Electronic Route Guidance by AUTOGUIDE: The Research Background. In *Traffic Engineering and Control*, Oct. 1987, pp. 525–529.
13. Y. J. Stephanedes, P. G. Michalopoulos, and R. A. Plum. Improved Estimation of Traffic Flow for Real-Time Control. In *Transportation Research Record 795*, TRB, National Research Council, Washington, D.C., 1981, pp. 28–39.
14. I. Okutani and Y. J. Stephanedes. Dynamic Prediction of Traffic Volume Through Kalman Filtering Theory. In *Transportation Research*, Vol. 18B, No. 1, 1984, pp. 1–11.
15. M. Ben-Akiva, M. Cyna, and A. de Palma. Dynamic Model of Peak Period Congestion. In *Transportation Research*, Vol. 18B, No. 4/5, 1984, pp. 339–355.
16. M. Ben-Akiva, A. de Palma, and P. Kanaroglou. Dynamic Model of Peak Period Traffic Congestion with Elastic Arrival Rates. In *Transportation Science*, Vol. 20, No. 2, 1986, pp. 164–181.
17. M. Ben-Akiva and A. de Palma. Dynamical Models of Transportation Networks. *Proc., PTRC Transport and Planning Summer Annual Meeting*, University of Bath, Bath, England, 1987.
18. P. Vythoulkas. A Dynamic Stochastic Assignment Model for the Analysis of General Networks. In *Transportation Research*, Vol. 24B, No. 6, 1990, pp. 453–469.

19. E. Cascetta and G. Cantarella. A Day-To-Day and Within-Day Dynamic Stochastic Assignment Model. In *Transportation Research*, Vol. 25A, No. 5, 1991, pp. 277–291.
20. I. Kaysi. *Framework and Models for the Provision of Real-Time Driver Information*. Ph.D. dissertation. Massachusetts Institute of Technology, Cambridge, Feb. 1992.
21. T. Lotan. *Modelling Route Choice Behavior in the Presence of Information Using Concepts from Fuzzy Set Theory and Approximate Reasoning*. Ph.D. dissertation. Massachusetts Institute of Technology, Cambridge, Aug. 1992.
22. M. Cremer and H. Keller. A New Class of Dynamic Methods for the Identification of Origin-Destination Flows. In *Transportation Research*, Vol. 21B, 1987, pp. 117–132.
23. N. Nihan and G. Davis. Recursive Estimation of Origin-Destination Matrices from Input/Output Counts. In *Transportation Research*, Vol. 21B, 1987, pp. 149–163.
24. A. Kessaci, J. Farges, and J. Henry. On-line Estimation of Turning Movements and Saturation Flows in PRODYN. Proc., *6th IFAC/IFIP/IFORS Symposium on Transportation*, Paris, France, 1989.

Publication of this paper sponsored by Committee on Transportation Supply Analysis.