A KNOWLEDGE-BASED SYSTEM FOR CONTROLLING AUTOMOBILE TRAFFIC

Alexander Maravas* and Robert F. Stengel**

Department of Mechanical and Aerospace Engineering
Princeton University
Princeton, NJ 08544

Abstract

Transportation network capacity variations arising from accidents, roadway maintenance activity, and special events, as well as fluctuations in commuters’ travel demands complicate traffic management. Artificial intelligence concepts and expert systems can be useful in framing policies for incident detection, congestion anticipation, and optimal traffic management. This paper examines the applicability of intelligent route guidance and control as decision aids for traffic management. Basic requirements for managing traffic are reviewed, concepts for studying traffic flow are introduced, and mathematical models for modeling traffic flow are examined. Measures for quantifying transportation network performance levels are chosen, and surveillance and control strategies are evaluated. It can be concluded that automated decision support holds great promise for aiding the efficient flow of automobile traffic over limited-access roadways, bridges, and tunnels.

Introduction

U.S. automobile traffic has been growing by 4 percent a year to its current level of 2 trillion vehicle-miles, and it is expected to double to 4 trillion vehicle-miles by 2020. According to Federal Highway Administration, if no significant improvements are made in the highway system, congestion delays will increase by as much as 400 percent [1]. According to IVHS America, the annual cost of congestion to the U.S. in lost productivity is estimated at over $100 billion [2]. In many areas there is very little that can be done to increase road capacity. There is not adequate right-of-way next to existing roads, and in many cases the cost of a new highway is prohibitively expensive. It is therefore imperative that new ways be sought to make better use of existing infrastructure.

In 1987 the Federal Highway Administration formed Mobility 2000, a joint effort between the government, industry and academia. This led to the formation of an organization called the Intelligent Vehicle Highway Society of America, or IVHS America [1]. IVHS America aims at improving the level of transportation services that are currently available to the public by integrated systems of surveillance, communications, computer and control process technologies [4].

IVHS technologies have been grouped into four generic elements: Advanced Transportation Management Systems (ATMS), Advanced Driver Information Systems (ADIS), Automated Vehicle Control (AVC), and Commercial Operations [3]. This paper concentrates on ATMS, which involves the management of a transportation network. Implementation of such systems requires development of real-time traffic monitoring and data collection techniques. More precisely, an Advanced Traffic Management System should have the following characteristics, as specified by the proceedings of the Mobility 2000 conference [3,4]:

- real time operation
- responsiveness to changes in traffic flow
- surveillance and detection
- integrated systems
- collaboration of jurisdictions involved
- effective incident control strategies

Effective incident control strategies will be a crucial part of this project. Contrary to widespread belief, not all congestion is due to rush hour traffic; 56% of costs incurred by congestion are due to non-recurrent events or incidents. Incidents include vehicle accidents, unfavorable weather conditions, highway maintenance, and road reconstruction [3]. It is essential to determine the nature and scope of an incident as quickly as possible. The control center should be informed about the incident either through
the police or other jurisdictional agencies. A more advanced approach would be visual validation of incidents with the use of camera surveillance systems. Effective detection and verification of incidents will lead to lower disruption of traffic flow [3]. Drivers could be routed to alternate paths to avoid unnecessary tie-ups and frustration due to long delays [3]. Traffic control centers would need real-time information about the network condition. An intelligent vehicle/highway system would have to monitor traffic throughout the day. For the near future, sensors will include inductive loops buried just below the surface and ultrasonic sensors mounted overhead. These devices will be able to count the number of vehicles passing a certain point and will gauge their speed. Another alternative would be for image-processing computers to extract traffic data from television pictures from cameras on the arterial network. Recent tests have provided very promising results about the accuracy of inductive loop detectors [1]. Ultimately, improvement of network surveillance technologies and link-time estimation techniques will be crucial in the implementation of an intelligent vehicle/highway system [5]. In the future, vehicles equipped with navigational systems could communicate directly with the control center, giving information about the drivers' locations, speeds, and destinations.

Advanced Traffic Management Systems will have to be integrated with Advanced Traveler Information Systems (ATIS) to ensure higher efficiency of the control system. Drivers will be informed about congestion, roadway conditions, and alternate routes through audio-visual means in the vehicle and through variable message signs at strategic points of the network. Information provided might include incident locations, fog or snow on the road, restrictive speeds, and lane conditions. Two-way real-time communication between vehicles and the control center could be facilitated by radio communications, cellular systems, and satellite communications [4].

Among the benefits of IVHS will be reduction in traffic congestion, reduction in the number of accidents, improved transit service, less fuel wasted, and fewer emissions from idling engines [4]. Fully integrated ATMS/ATIS combinations could reduce congestion in urban areas from 25 to 40%. Unchecked traffic congestion is the largest contributor to poor air quality and wasted fuel consumption. IVHS will not solve all problems in transportation, but it will increase the level of services rendered.

Fundamentals of Traffic Flow Modeling

Evaluating traffic performance requires a thorough understanding of traffic flow characteristics and analytical techniques. The most important macroscopic flow characteristics are flow rate, density, and speed [7]. The flow rate $q$ past a point is expressed as [8]:

$$ q = \frac{n}{T} \quad (1) $$

where $q$: flow rate past a point

$n$: number of vehicles passing point in time interval $T$

$T$: time interval of observation

It is important to recognize that $q$ is sensitive to the selected time interval $T$ during which the measurement began and ended. Whatever the value of $T$, the most common units for $q$ are vehicles/hr.

Density (or concentration) can be expressed as [7,8]:

$$ k = \frac{n}{L} \quad (2) $$

where $k$: density

$n$: number of vehicles on road

$L$: length of road

Density is an instantaneous traffic measurement, and its units are usually vehicles/lane-mile. In most cases, one mile and a single line of vehicles are considered. Traffic densities vary from zero to values that represent vehicles that are completely stopped. The upper limit of $k$ is called the jam density and is on the order of 185 to 250 vehicles per lane-mile, depending on the length of the vehicles and the average distance between vehicles.

Speed is another primary flow variable. Space-mean speed is the average speed of the vehicles obtained by dividing the total distance traveled by total time required, and it is expressed as [8]:

53
where \( u \): space-mean speed
\( s_i \): distance traveled by vehicle \( i \) on roadway
\( m_i \): time spent by vehicle \( i \) on roadway

A very important relationship exists between the three fundamental stream flow variables that have been defined above. This relationship is [7]:

\[
q = ku
\]  \hspace{1cm} (4)

A linear speed-density relation has been assumed to simplify the presentation. The relation can be expressed as [7]:

\[
u = u_o - \left( \frac{u_o}{k_o} \right) k
\]  \hspace{1cm} (5)

This relationship indicates that as speed approaches free flow, speed density and flow approach zero. An increase in density causes a decrease in speed until flow is maximized at \( q_m \) and speed and density reach their optimum values \((u_o, k_o)\). Further increases in density cause density to reach its maximum value \( (k_i) \) and speed and flow to approach zero [7]. A relationship between flow and density can be obtained by substituting equation (5) into (4). This yields [7]:

\[
q = u_o k - \left( \frac{u_o}{k_i} \right) k^2
\]  \hspace{1cm} (6)

Under low-density conditions, flow approaches zero and speed approaches free-flow speed \( u_o \). At optimum density, flow is maximized, and speed attains an optimum value. Maximizing the objective function (6) by setting its derivative equal to zero \( dq / dk = 0 \), we find that optimum density occurs at half the jam density \( (k_o = k_i / 2) \). This is true only for a linear speed density relationship, but the same reasoning can be applied to other non-linear relationships [7].

The optimum speed is half the free flow speed \( (u_o = u_f / 2) \). Because \( q_m = k_o u_o \), it is evident that \( q_m = u_o k_o / 4 \). Once again, it is important to remember that these values are only true for a linear speed-density relation [7]. The flow-density relationship often serves as a basis for highway control. Density is used as the control parameter, and flow is the objective function. At low densities, demand is being satisfied and level of service is satisfactory, but as density increases control is needed to keep densities below or near the optimum density value [7]. Other models have been proposed by transportation planners. This is an optimization problem with traffic flow as the objective function and traffic density as the control parameter. The technology to measure traffic density exists, and its knowledge can help us estimate traffic flow, average speed, travel time, and level of service.

Assumptions of Surveillance and Control Procedures

Every morning there is a large pool of people east of the Hudson river who want to cross it. Every member of this group of "intelligent agents" is part of a continuously changing environment, holding an individual behavioral response to the evolution of the dynamic system. It has been observed that commuters usually choose the route with the shortest travel time. Consequently, drivers tend to divide themselves between two routes in such a way that the travel times between them are identical. Finally, an equilibrium point is reached in which commuters do not have much choice between two routes since the user cost (travel time) of traversing the two links has stabilized [8].

In many cases, an incident can disturb the user equilibrium. In such instances transportation planners can help establish a new equilibrium where no route is underutilized. By dispensing traffic advisories to the right people, one can be assured that misallocation of transportation resources can be avoided. Computer simulations have proven that the use of route guidance systems can reduce travel time to all drivers by up to 20% [9]. Route guidance was found to be more helpful as the duration of incidents increased, in which cases more drivers had to be routed to achieve an optimal traffic assignment [9]. A corresponding surveillance procedure is illustrated in Figure 1.

In selecting between alternate routes, a user-optimal system chooses a route that minimizes the travel time of the individual driver.
However, a system-optimal system selects a set of routes that minimizes the overall travel time of all drivers [11]. Routing decisions should not be made independently because every decision effects the whole network: if many drivers choose a certain route at the same time, that route may become congested and non-optimal [10].

Figure 1 Proposed Surveillance Procedure.

A Declarative Framework for Traffic Control

Traffic control can be partitioned into reflexive, procedural, and declarative functions. Reflexive functions operate at the subconscious level of human thought; they are instantaneous reactions to external stimuli. Procedural actions also operate on a subconscious level, but they are more complex than reflexive functions, following a set of basic rules and actions that represent skilled behavior. Declarative functions operate at the conscious or preconscious level of thought. On the conscious level they require attention; on the preconscious level they require intuition and conceptual formulations. They involve decision making and provide models for system monitoring, goal planning and system/scenario identification [12,13].

A traffic management system requires goal planning and system monitoring. At every instant, all alternatives must be considered and decisions must be made through a deductive process [13]. A declarative model makes such decisions in a process that is similar to human reasoning [6]. All our knowledge, beliefs, and experience of traffic modeling are placed in a declarative framework to provide a system that reasons in an intelligent manner.

This paper focuses on declarative traffic controls. Rules that reflect the controllers knowledge of the response of the control system are established. Programming is implemented as a CLIPS expert system that supports various programming paradigms. CLIPS was chosen because it can be easily integrated with C programs. The rule-based programming paradigm is useful in modeling traffic incidents. In procedural programming, the order in which all the commands are executed is pre-specified. In reality, traffic incidents are often unpredictable. By establishing the appropriate heuristics and rules, the system becomes "intelligent." When it is faced with a certain problem, it uses pattern matching that is appropriate to the existing facts.

Computer systems often are organized in a modular structure to increase computational efficiency. Time can be saved by looking at rules and facts that are relevant at that instant. Modular representation allows partitioning of the knowledge base into easily manageable segments, thus making the system easily expandable. CLIPS incorporation of modules is similar to the blackboard architecture of other systems. Knowledge sources are kept separate and independent, and different knowledge representation techniques can be used. Communication of all sources takes place through the blackboard [15].

Task definition is an important factor in the development and design of such rule-based systems. The ultimate goal is to develop an expert system of expert systems, which is a hierarchical structure that reasons and communicates like a team of cooperating people might [13].

A knowledge-based system called the Traffic Information Collator (TIC) has been developed at the University of Sussex. The TIC receives police reports on traffic incidents and automatically generates appropriate warning messages for motorists. The system operates in real-time and is entirely automatic. The TIC is comprised of five processing modules each holding its own knowledge base [16,17].

Research in air traffic control has been conducted in a similar manner. Cengeloglu integrated an Air Traffic Control Simulator (written in C) with a traffic control decision framework (implemented in CLIPS). The simulator creates a virtual reality of airspace and continuously updates the information
Using a Knowledge-Based System for Decision Aiding in Traffic Control

The aim of this project is to design an expert system that evaluates traffic conditions and dispenses travel advisories to commuters and traffic control centers. A decision-support system has been written in the CLIPS programming environment to illustrate several traffic control procedures (Fig. 2). The program is geared toward illustrating a method of traffic control rather than solving a particular problem. Some traffic parameters have been approximated and certain simplifying assumptions have been made to avoid unnecessary computational complexity.

A small network was constructed for initial program development. The network consists of a section of the New Jersey Turnpike and the routes connecting Exits 14, 16, and 18 to the Holland Tunnel, Lincoln Tunnel, and George Washington Bridge. Node and link data for all the roads in New Jersey and New York has been accumulated by the Civil Engineering Department at Princeton University. Future research could be geared toward applying the declarative control procedures developed in this project to a network of a larger scale.

Actual implementation of this system is based on it receiving real-time information about traffic conditions of a network using sensor measurements at different points on the network. The current program uses scenario files, with hypothetical traffic densities from all roads in the network. The program reads these values (as well as the pertinent time period of the day) and creates the appropriate data constructs, which are asserted as facts. The system then matches these facts based on pre-specified heuristics. Once it has concluded a declarative search, it gives certain advisories and recommendations at the CLIPS command prompt. All advisories are made with the idea that they would be broadcast on changeable signs at several roadway locations; they could also be displayed at traffic management centers or on the displays of suitably equipped vehicles. Alternative scenario files test the system under several hypothetical operating conditions.
Implementation of the Traffic Management System

The decision-support system, programmed with CLIPS 6.01, supports several procedural functions, that are necessary for the computation of relevant information. Once the density is known, the average travel speed, flow rate, service rate, and level of service can be calculated from the appropriate models and equations.

Historical data of expected traffic demand for the Hudson river crossings have been stored in a function that returns the expected number of vehicles at a certain time period. The use of an appropriate filter would make this function redundant. It provides information about the probable evolution of traffic and is sufficient for the preliminary development of the system. It makes the system "forward looking" and capable of predicting congestion buildup. Historical data could be stored as facts; however, this clogs up the facts list, and the system is forced to look at irrelevant facts.

After reading initial data from the scenario file, the roadway densities are asserted as facts, and they are put onto the facts list. The initialization rules create the appropriate data constructs based on these basic facts and the appropriate procedural functions. Most of the program's knowledge is stored in templates, which are similar to structures in C. Thus every link of the network has its own template, containing information such as speed, density, flow, operational lanes, service rate, and accident-status. Templates are convenient for storing data and can be readily accessed and modified by the user. After all the values of the slots of the templates have been calculated, the templates are asserted as facts. Thereafter, decisions are made by pattern matching on the values of the templates' slots. Data storage is compact so as not to overflow the CLIPS fact list.

The system is readily adjustable to lane closures due to maintenance. For instance, if one of the lanes of the Lincoln Tunnel is closed due to maintenance, then when the knowledge base is initialized, it takes this fact into account. In the case of bad weather, such as a snowstorm, the values of the free-flow speed and the jam density should be reevaluated. Since these deviations from normal operating conditions are incorporated into the system once it is initialized, all decisions made thereafter are adjusted accordingly. System initialization at every time increment ensures that the link-node data is augmented to reflect current conditions. Thus the system is very flexible and can include all foreseeable traffic situations.

Broadcasting travel advisories is a challenging part of the program. For instance, the fact that there is an accident on the route from Exit 14 to the Holland Tunnel is useful for people approaching that exit, but not for people traveling away from it. Thus the system is faced with the decision of whether to make a broadcast and to whom it should be made. Due to the geometry of the network, the broadcast heuristics are different for every decision node. Even in such a small network, there are numerous broadcasting possibilities.

Automated communication between commuters and the control center can be achieved by the appropriate use of advisory rules. The data templates also contain information such as the presence and severity of accidents on the links of the network. Every link's template has a slot that is called "accident status," which can be either TRUE or FALSE. The default value is FALSE, but once an accident occurs it is switched to TRUE. Thereafter, the system ensures that the appropriate people are informed of the incident. Templates also store the
estimated time to restore the link to normal operating conditions and the degree of lane blockage at the accident site.

While radio advisories provide commuters with information about traffic delays and adverse traffic conditions, it is doubtful that they give them all the information they need at the right time. Since not all drivers are tuned to the same station and radio advisories may lack central coordination, the result of broadcasting may not be system-optimal. Changeable message signs ensure that the appropriate people get the right message at the right time. The system is more helpful under such circumstances, since it can help bring the network back to user equilibrium by ensuring that all links are properly utilized.

In general the factors that should be considered in any routing decision are: a) traffic density and velocity profile of main and alternate route, b) length of main and alternate route, c) percentage of divertible traffic volume, and d) demands at on-ramps on both routes [20]. Once the density of the link is known, an average speed can be computed using one of the traffic models described earlier in this paper. Dividing the length of the link by the average speed yields the current (experienced) travel time.

Routing between two alternate routes can be achieved by the use of a tolerance-level measurement. The difference between experienced travel times (or flow) between two routes can be calculated. If it is higher than some prespecified level, then commuters should be routed to the underutilized roadway. The use of individual travel times is user-optimal, whereas the use of traffic flow is system-optimal. Optimizing commuters travel time can probably be done better with the use of shortest-path algorithms, whereas equilibrating flows can be incorporated in an expert system.

Traffic flow is a measure of highway productivity. Ensuring maximum utilization is essentially an optimization problem, with flow as the objective function and density acting as the control parameter. The flow of a link is maximized at the optimal density \( k_o \). For a linear model, the optimal operating density is half the jam density \( k_j \). The system has the expertise and knowledge to recognize how far the actual density measurements are from ideal conditions. If the traffic density of a road is less than optimal, then the road is under-utilized. If the density is higher then optimal, then it is over-utilized. An advanced intelligent highway system should be able to monitor and compare traffic flow in both directions of a link to see if the decision for a lane-direction change is warranted. The ultimate goal is to ensure that all routes are utilized properly. This section of the system is subject to a lot of development and improvement since the domain knowledge is uncertain. Currently there is no clear way to route traffic optimally.

The system searches its historical data to see if traffic demand for a link is expected to rise or fall in the next time period. If travel demand is expected to fall and the road is being underutilized, the system suggests that more vehicles be routed to that link. Diversion of traffic flow is an effective method of improving traffic performance and can be achieved by rerouting drivers with specific direction and destinations [20]. Advisories on the New Jersey Turnpike are different for people traveling North than for those traveling South.

The issue of how long a message should be broadcast is significant. Suppose that congestion can be avoided if 30% of drivers respond to the diversion sign. If only 15% of the drivers follow the diversion recommendation, congestion will not improve as expected. A feedback control system could take this into account by deciding to broadcast the diversion message for a longer time period until the desired utilization levels are reached. This approach compensates for all uncertainties in the percentage of divertible traffic flow [20]. The critical design issue is assuring that the system reaches stability quickly.

Knowledge base initialization at frequent time intervals, through the use of sensor measurements, makes this method of broadcasting a closed-loop feedback control process. Since the real-time implementation of this system would rely on cyclic search, a message would be sent every time the system decides that traffic should be diverted. Currently the program does not have real time measurements or simulated traffic demands, so it does not execute a cyclic search. It considers 24 one-hour periods over the span of a day. Real time implementation would require that the CLIPS knowledge be reinitialized at every time increment with the state measurements.
Conclusion

This project has made a step towards emulating an automated decision process for an Advanced Transportation Management System. This non-conventional approach to transportation modeling has examined the applicability of intelligent control techniques in management. Since a fully operational Intelligent Vehicle Highway System will require full integration of symbolic and numerical knowledge, declarative rules have been embedded with procedural code. A framework for modeling traffic incidents has been provided. The link information of the system is augmented on a real-time basis, and advisories are issued based on the current state of the system.

Operation of this system requires that the traffic control center has knowledge of the traffic densities of all links in the transportation network. The technology for making such measurements exists and has been described earlier in this paper. Implementation of this system will require research into how programming software will be integrated with all the system sensors.

Before fully implementing such a control process, it is necessary to choose what degree of automation the system should have. The system should be allowed to run on its own, without human intervention, only when all possible errors have been removed from it.

It is difficult to foresee all the incidents that could happen. Even for such a small network there are many possibilities for issuing travel advisories. As the area that is monitored by sensors becomes larger, it is evident that human operators cannot check everything that is going on. However, the heuristic rules and frequent knowledge-base initialization make the system adaptive to many situations. The size of the knowledge base and the number of advisories are limited only by the available computer memory. Initially the system can be tested in a small area. Thereafter, additional rules for broadcasting to other locations can be added incrementally. Additional details of this research can be found in Ref. 21.

Future Work

Future work can be geared towards expanding the knowledge base by using the object-oriented programming paradigm offered by CLIPS. Node and link data of large networks could be stored compactly in an object-oriented format. Different classes of roads could be defined (e.g. arterial, expressway, intersection). Every link would then be an instance of these classes and it would inherit some of its properties from them.

The cyclic search must be implemented with the use of actual or simulated data, requiring the knowledge base to be reset (initialized) at frequent time intervals. It would be interesting to link CLIPS with a traffic simulator written in C. The simulator could generate traffic demands, and the CLIPS program could issue appropriate advisories.

This system knowledge base is subject to refinement. Additional rules must be added so the system knows what to do in the absence of certain density measurements, which could arise from malfunctioning sensors. Since not all transportation engineers use the same traffic models, it would be desirable to allow the user to use different traffic models or to be able to create his own model with an equation parser. The traffic routing technique and its relevant objective function needs to be reevaluated. Backlogs due to ramp-metering also should be examined. The use of estimators and prediction algorithms would enhance system performance significantly.

A learning control system would be able to learn from its daily experiences. Initially, it could be "trained" on a set of simulated data. Data archiving, on a daily basis, would increase the size of the system's knowledge base. Thereafter, it could evaluate how well it handled a previous accident. Hence, if it was in a similar situation it would use its acquired expertise to issue the appropriate advisories. An intelligent system could detect traffic incidents from unusual sensor readings. Eventually the system would be able to distinguish between weekday and weekend traffic patterns. Considering the recent technological advances in intelligent control systems, the era of the fully automated highway might not be very far away.

Acknowledgement

This work was supported by the National Science Foundation, Grant No. ECS-9216450.
References


