Development of Prototype Knowledge-Based Expert System for Managing Congestion on Massachusetts Turnpike

Arti Gupta, Victor J. Maslanka, and Gary S. Spring

A prototype knowledge-based expert system has been developed to assist in the management of nonrecurrent congestion. The system encompasses incident detection, verification, and response; it includes a real-time, dynamic network model for motorist diversion. A case-study simulation on the Massachusetts Turnpike illustrates the potential benefits to be derived from the system.

Urban traffic congestion is not a new problem. It antedates the motor vehicle and has been a continuing concern for much of this century. In recent years, several cities, including Phoenix, Atlanta, Houston, San Francisco, and Washington, have identified traffic congestion as their most serious regional problem. In fact, it is a serious national problem and continues to worsen. The reason is not hard to see: the number of cars owned in the United States increased by two-thirds between 1970 and 1987, and total annual vehicle miles traveled (VMT) increased from 1.120 billion to 1.910 billion during the same period. Yet inflation-adjusted expenditures in 1987 were only 6 percent above 1970 levels. Since freeways account for only 3 percent of road mileage in urban areas but carry more than 30 percent of the total VMT, they are of particular interest in addressing the congestion problem.

Congestion may be classified as recurring or nonrecurring. Recurring congestion occurs when demand exceeds supply (usually during peak periods) on a regular basis. Common causes include lane drops, heavy volumes, poor geometrics, weaving sections, and so on. This type of congestion is often seen in cities during peak periods of travel when large numbers of work trips are being made. Recurring congestion is characterized by unanticipated events such as accidents and disabled vehicles that cause a reduction in normal capacity. Given that these events, or incidents, are quasirandom in nature, they are difficult to predict and solutions to the problems they create are difficult to implement.

FHWA sponsored a study in 1986 to quantify the magnitude of the urban freeway congestion problem on a national scale. Estimates have been made for delay, excess fuel consumed, and user costs on the basis of assumed values for user time and wasted fuel, for both recurring and nonrecurring congestion. The results of the study show that there was a 30 percent increase in delay from 1984 to 1985 and predict an estimated fivefold increase by the year 2005 if no improvements are made, of which 70 percent will be due to nonrecurrent congestion. The congestion problem, as well as driver safety, could be greatly improved if these incidents were more efficiently managed.

It has long been acknowledged that we cannot "build" our way free from the congestion problem but must better manage existing facilities using transportation systems management (TSM) techniques. Freeway incident management has been used successfully for the past 30 years as a tool to reduce the impact of incidents in a number of urban areas throughout the United States. Programs of this kind have been in place in several states: California, Arizona, Washington, Illinois, Florida, Texas, and New York are examples.

FREeway INCIDENT MANAGEMENT

This paper explores the application of expert systems to freeway incident management (FIM) and proposes a methodology for developing an expert system to assist in incident management on the Massachusetts Turnpike. The generic incident management process has at its heart a traffic control center that monitors freeway operation for incidents. When an accident occurs, the controller responds appropriately—dispatching emergency vehicles and personnel, notifying appropriate agencies, alerting approaching motorists, and deciding whether to divert traffic and if so along which routes and for how long. From a traffic management viewpoint, freeway incidents should be removed as quickly and efficiently as possible. Additionally, freeway demand should be intercepted and diverted to other routes if the reduced roadway capacity during the incident is insufficient to satisfy demand and if practicable alternative routes exist. Success in achieving these goals results in increased freeway safety and decreased congestion and delay.

Incident management is a continuous process. The system implemented must be always available to detect and respond to incidents. Major components of the FIM process are detection, verification, response, and monitoring or feedback. Incident detection and verification require a freeway surveillance system. This system may be as sophisticated as the automatic detection systems (such as loop detectors and closed-circuit television) used in California or as simple as police...
or traveler call-ins. The latter method is the default surveil-
ance system on most roadways and consists of passing mo-
tors, or police patrols, notifying local police agencies. When
an incident is thought to have been detected, verification is
necessary to screen out false alarms. This is usually accom-
plished through police patrol or closed-circuit television. In-
cident response encompasses a wide range of activities and is
the basic requisite for a successful incident management sys-
tem. Traditional response activities include the provision of
medical services, fire agency response, hazardous material
containment and cleanup, vehicle removal, and traffic control
at the incident scene. Advanced response systems include
motorist information systems, motorist diversion systems, and
interconnections to other traffic control systems such as traffic
signal systems along parallel routes and freeway ramp-metering
systems.

RATIONALITY AND SIGNIFICANCE

The large quantities of information flowing into the traffic
control center, typically all at the same time, make responding
quickly to freeway incidents extremely difficult. Decisions
based on this copious, simultaneous information must be made
quickly and accurately and must be disseminated just as quickly.
This type of response requires traffic system managers who
are well versed in handling emergencies and who are able to
make decisions on the spot. Such managers have access to a
tremendous amount of knowledge derived mainly from work
experience in the area. Thus, the same problem faced in other
application areas requiring special expertise must be faced
here as well. Experts are rare and expensive, and it is often
difficult to retain enough of them long enough to sustain
effective operations. This means that valuable expertise is
often available only sporadically and at significant cost to the
user. It is for these reasons that expert systems offer such
potential. Expert systems are computer programs designed to
solve problems whose solutions require expertise. They at-
tempt to use the knowledge of human experts to solve prob-
lems (6). Perhaps the most compelling reasons for using expert
systems for FIM is their ability to use all available knowledge,
consistently and without error or misjudgment—important
considerations for real-time applications.

Within the context of incident management, the expert sys-
tem is envisioned as a real-time, on-line computer system that
will support the traffic system manager. The traffic system
manager is traditionally a police agency representative re-
sponsible for incident management on a particular portion of
the roadway network. This person is responsible for basic
direction and coordination of all agencies involved in incident
response. Without an expert system, the manager performs
incident management duties, relying on knowledge, memory,
past experience, and written guidelines. The computerized
expert system supports the traffic system manager in the fol-
lowing ways:

1. The expert system can screen large volumes of data,
alerting the manager only of data that appear to be abnormal.
In this manner, the manager is protected from information
overload and can devote his or her time to activities dealing
with those data that suggest the existence of traffic system
abnormalities. If properly programmed to screen traffic data,
the expert system can detect abnormal fluctuations far more
reliably than can a busy and fatigued human being.

2. The expert system provides consistency. The expert sys-
tem will not vary in its response, as might different humans
serving as traffic system managers, or even the same human
under different working conditions. Of primary importance
in this regard, the expert system will not forget important data
or procedures.

3. The expert system provides an automated menu-driven
procedure to guide the traffic system manager through the
tasks of the job. Through interconnection with data bases and
other computerized systems, the traffic system manager can
work much more quickly and therefore effectively.

4. The expert system can be used as a training tool through
its off-line use by both inexperienced and experienced man-
gers through a wide range of hypothesized incidents.

It is important to note that the expert system is not intended
to replace the traffic system manager. Each conclusion reached
by the expert system can be accepted or rejected by the man-
ger, as deemed appropriate. Additionally, the manager may
review the logic that the expert system used to reach its con-
clusion. The nature of incidents that occur is so varied and
unpredictable that the expert system may not be able to re-
pond properly to unforeseen events.

As discussed previously, the response plan is the heart of
incident management. Several automated FIM systems are
currently under development. The FRED system places its
main emphasis on the surveillance, verification, and decision
support for on-site response strategies (5). Lakshminarayanan
and Stephanedes focused almost exclusively on the on-site
aspects of response (7). None of the current efforts, however,
appears to have used network simulation techniques to assess
the suitability of diversion plans.

The incident management process may be broken into two
parts. One part requires judgment regarding appropriate re-
response strategies based on incident severity level, such as
dispatching emergency vehicles, notifying incident manage-
ment teams, and determining response level and appropriate
diversion routes from a large set of preplanned routes for the
facility. This type of decision making (which in this case is
based on type of incident, time of day, location—informa-
tion used to determine expected durations, volume/capacity ratio,
and so on) is a classic expert systems situation (6). The second
component consists of essentially assigning vehicles to a net-
work in real time and determining optimal paths from among
the feasible routes recommended by the first part—clearly
more amenable to a simulation program. Thus, two very dif-
terent computer programming paradigms were used to reflect
the special character of incident management problems: an
expert system was built for the first, and a procedural data
processing and network simulation program was built for the
second.

The expert system developed herein uses a combination of
an “Exsys” shell and a FORTRAN module. The Exsys shell
is a layer of software developed using a generalized expert
system package called EXSYS. This package can be run on
any IBM PC, XT, AT, or compatible computer with 320K
RAM. The expert system selects an appropriate response
strategy and records basic information about an incident. Con-
The case study for this research is a section of the Massachusetts Turnpike. The turnpike is a 132-mi-long highway with 24 interchanges. It caters to both commuter and through traffic. It is well maintained and patrolled. A useful feature of the MassPike is its special emergency access points; highway authorities use them primarily for maintenance, snow removal, and other emergency purposes, and they are occasionally used for traffic diversion purposes also. The MassPike is a tolled facility. Thus, if traffic is diverted off the MassPike during an incident, there is a question of loss of revenue and the ability of toll plazas to handle extra traffic. Although we have identified these issues and recognize that they should eventually form an integral part of a route diversion strategy for a highway such as the MassPike, only toll lane capacities have been included in the present study. These issues need further research in collaboration with the MassPike and other highway authorities.

The section of the MassPike studied in detail is from Exit 9 to Exit 12, with incidents simulated in sections between Exits 10 and 11 and Exits 11 and 11A eastbound. These sections were chosen on the recommendation of MassPike staff, because there have been several serious incidents in the past in this area. For a map of the study section, refer to Figure 1. The study section passes through Sturbridge (Exit 9), Auburn (Exit 10), Millbury (Exit 11), Westborough-Hopkinton (Exit 11A), and Framingham (Exit 12). It intersects with I-84 at Exit 9 and connects to Route 20; with I-290, I-395, and Routes 12 and 20 at Exit 10; with Route 122 at Exit 11; with I-495 at Exit 11A; and with Route 9 at Exit 12. This section provides a variety of alternative routes, making the task of finding the best one challenging and interesting.

The volume and network data for the study network were obtained from the Massachusetts Turnpike Authority and the Massachusetts Department of Public Works. The turnpike authority also provided information on the incident management techniques currently used.

**KNOWLEDGE-BASED INCIDENT MANAGEMENT SYSTEM**

FIM is a multidisciplinary activity. Personnel from various agencies such as highway (traffic and maintenance), police, fire, environmental, and medical agencies need to work together in a coordinated manner. When an incident occurs, it must be decided who should be informed. In the expert system developed here, various incident situations and corresponding responses are included. It was designed as a comprehensive crisis management tool. For example, if there is a spill, depending on whether the spill is hazardous or not, special cleanup forces or regular cleanup forces should be informed. If the spill is a potential fire hazard, a fire crew should be put on alert or dispatched. If there are vehicles to be towed, a towing company should be called on to send trucks. Thus the expert system would prompt the operator for information about the incident and respond with the recommended action. The operator would receive the information from the police or another authorized person at the incident site. The operator would also act on the actions recommended by the expert system. The situations and responses have been formulated on the basis of what highway authorities do now or what they think should ideally be done, on the basis of their knowledge and experience.

**Crisis Management**

Expertise consists of knowledge about a domain, about how to use that knowledge, and about problem characteristics. Therefore, expert systems have three basic components: a knowledge base that contains heuristic knowledge (most often in the form of rules and facts) about the problem domain, an interpreter that contains reasoning methods (i.e., ways to process and use domain knowledge), and a data base that contains problem characteristics.

**Knowledge Base**

The process of coding the knowledge base consists of implementing a much more detailed version of the decision tree shown in Figures 2, 3, and 4 into a set of if-then-else combinations. The decision tree provides the conceptual framework of the problem into which details may be placed. Details, such as appropriate responses and incident severity level determinations, were taken from interviews with Massachusetts Turnpike Authority personnel. The system currently contains 36 if-then-else rules of thumb, some of which are necessarily site-specific. For simplicity, all site-specific information is stored in the body of the rules. However, the expert system shell used does provide a facility by which to separate site-specific information—thus allowing for system transferability. The rules for alternative route preplanning are based on flow and capacity constraints by time of day and location and the in-
Incident's expected duration (in turn based on severity). The rules choose a set of alternative routes that are feasible, taking into account the dynamics of the parameters involved. The knowledge base has three primary parts:

- Incident detection and verification—This requires a freeway surveillance system. The nature of this system may vary considerably. In the system developed here, the surveillance consists of notification by either passing motorists or police patrol. This method exists on many roadways. A more sophisticated system may rely on vehicle occupancy measurements, volume or speed data, or closed-circuit television. After an incident is detected, it needs to be verified to screen out false alarms. This is accomplished through police patrol or closed-circuit television, if the incident has been reported by a motorist or similar observer.

- Classification of the incident—This consists of information provided by police patrol on the incident characteristics, such as its time, location, and severity.

- Notification of the incident—This consists of notifying all agencies required to clear and manage the incident site, after the occurrence of the incident has been verified.

After these three tasks have been accomplished, expert system control is passed to the diversion module with appropriate data to decide if diversion is required and to recommend diversion routes. Figure 2 depicts this process. The diversion module provides a diversion strategy.

**Interpreter**

The rule-based off-the-shelf interpreter EXSYS was chosen for this system because it provides a very simple programming environment while offering many useful utilities for the developer and the end user. The shell has all of the basic features necessary for effective implementation. It is relatively inexpensive and has a fairly friendly user interface. It provides a rudimentary explanation facility, allows what-if scenarios, interacts well with other external programs, and allows specification of uncertainty values. These qualities, along with the ability of the user to review the logic employed in arriving at a decision, make EXSYS an excellent prototyping tool.

**Data Base**

The system data base has several components: remotely sensed volume data, capacity data stored internally, and an active, constantly changing component that contains the current state of the problem—that is, which rules have been fired, which facts are true, and so on. Volume and capacity data for each link in the network are stored in ASCII data files containing temporal and spatial volume and capacity information. This is meant to simulate a real-time situation in which volume data are continuously fed into the computer system from field detectors. For each incident there exists a "best" diversion route to be followed depending on the location, severity level, and the time of day of the incident. Criteria used to determine incident severity level were expected incident duration and the volume/capacity (v/c) ratio of the affected freeway link.

At present, the freeway v/c ratio at any point is computed by the system using volume and free-flow capacity values.
retrieved from the volume data file just described. However, the system is structured so that it supports traffic data acquisition hardware. It does not differentiate between data captured from a data file and data captured from, say, a loop detector. Similarly, system outputs could just as easily be displayed on roadside changeable message signs as on a computer screen.

Example

Consider an example to illustrate the operation of the expert system. We will first describe an incident and then go through the sequence of questions asked by the expert system, the responses selected by an operator, and the action recommended by the expert system.

Suppose a trailer truck containing combustible material overturns, colliding with a car near Mile Marker 95 in the eastbound direction on the MassPike at 9:00 a.m. Assume that the incident is reported by a driver on the scene. Let us further suppose that several people are hurt, some with serious injuries; that all MassPike eastbound lanes are closed; and that it is estimated that it will take approximately 2 hr to clear up the incident and restore the normal flow of traffic.

The sequence of questions and response is given in the following.

Incident Verification

- **Question**: Who is reporting the incident?
  1. Police
  2. Observer
- **Response**: 2 (observer)
- **Recommended action**: Dispatch police cruiser to verify the incident.

- Incident is verified.

Figure 2 depicts the steps in this phase.

Incident Classification

- **Question**: What is the approximate location of the incident (mile marker) and direction?
  - **Operator**: 95, eastbound
- **Question**: What is the expected duration in minutes?
  - **Operator**: 120
- **Question**: What is the extent of the blockage?
  1. Shoulder disabled
  2. Shoulder blocked
  3. One lane blocked
  4. Two lanes blocked
  5. Three lanes blocked
  - **Operator**: 5 (three lanes blocked)

- Incident is classified.

Figure 3 shows the steps in this phase.

Incident Notification

- **Question**: What kind of injuries?
  1. Serious
  2. Moderate
  3. Mild
  4. None
- **Operator**: 1 (serious)
- **Recommended action**: Dispatch ambulance to Mile Marker 95 immediately.

- Emergency medical technician (EMT) is notified.

- **Question**: There is . . .
  1. Fire
  2. Spill
  3. Neither
- **Operator**: 2 (spill)
- **Question**: What is the classification of spilled material?
  1. Combustible
  2. Other hazardous
  3. Nonhazardous
- **Operator**: 1 and 2 (combustible and other hazardous)
- **Recommended action**: Alert fire department to stand by and inform hazardous material cleanup team.

- Fire department and cleanup forces are notified.

- **Question**: What types of vehicle are to be towed?
  1. Car
  2. Small truck
  3. Trailer truck
  4. None
- **Operator**: 1 and 3 (car and trailer truck)
- **Recommended action**: Dispatch one small and one large tow truck to Mile Marker 95 immediately.

- Towing company is notified.

This completes the notification phase, because all agencies have been notified (Figure 4). At this point, control and all incident information is passed to the FORTRAN module.

Traffic Diversion

When an incident reduces freeway capacity and causes congestion, the main concern of the transportation manager is to confine the problems due to the incident to that area itself, clear the incident as soon as possible, and return traffic flow to a normal condition. One of the main concerns is whether traffic should be diverted from the incident area. If it should be, How much traffic should be diverted? What alternative routes are appropriate? and When should the diversion end? The method for determining the appropriate alternative routes consists of two procedures: alternative route preplanning and real-time route diversion.

In the choice of alternative paths, several criteria can be applied. Travel time is one of the most widely used criteria, since it can be easily quantified and is of utmost concern to a motorist stuck in traffic. Here one needs to make a dis-
distinction between system- and user-optimal states. In the system-optimal state, travel time is minimized for the entire network; that is, the overall travel time for all the users is optimized. The term “user optimal” implies that each user tries to optimize travel time, which will not generally result in the system-optimal solution for transportation networks.

Alternative Route Preplanning

This step consists of developing preincident, detailed alternative route contingency plans for any location on the freeway system. As a first step, the freeway is divided into different sections and capacities are estimated. Then an inventory of freeways and all roadways that might serve as alternative routes for every section of freeway is completed. Information such as street widths, curvature, grades, pavement conditions, adjacent land uses, and weight restrictions is used to estimate the capacity of each link and to determine the suitability of each route. Special information, such as the presence of schools and special events, is taken into account in determining alternative routes. For example, during school opening and closing times, it may be desirable to avoid diverting traffic to school routes; during special events, such as sporting events, there may be no reserve capacity for diverted vehicles. This information can easily be stored in the knowledge base. Thus, we have a set of alternative routes for each section of the freeway with information about them for all hours of the day. Figure 1 depicts the MassPike study section divided into different links. There are in all 50 links and 38 alternative routes. The map shows 37 links; the rest of the links correspond to the MassPike links between the exits and the MassPike on- and off-ramps.

Real-Time Route Diversion

When a congestion-causing incident occurs, the expert system will act as an evaluator of the situation and help the system operator make decisions about traffic diversion. Real-time diversion implies assigning traffic to different routes by considering the prechosen alternative paths in a dynamic assignment modeling process. In this study, we have developed an algorithm to divert traffic from the MassPike by considering the conditions on the MassPike and on alternative routes and the subsequent effects of diversion strategy. Static and dynamic network models were developed to assign traffic to alternative paths during an incident. These network models for route diversion were tested extensively for incidents occurring at different times of the day, for various levels of lane closure, and for different incident durations. The purpose of the simulation was to test the output of the models for reasonableness and to increase insight into the problem. The test results show that MassPike exist ramps are the major bottlenecks when traffic is diverted off the MassPike. This can be attributed to the limited capacity of toll booths.

It was concluded that static models are not appropriate for modeling dynamic traffic events, such as incidents, because static models assume constant network characteristics over time. The use of static models to model incidents leads to impractical solutions, because of the inability of these models to include changes in volumes, demand, and capacities over time. In contrast, dynamic models can incorporate these changes and, therefore, provide better, more practical solutions.

The main input to the dynamic model is incident data and network data. Incident data include the time, duration, severity, and location of the incident; network data include volume, capacity, number of lanes, lane width, length, and free speed for each link in the entire network. The model operates on an IBM-compatible PC. The model output includes measures of effectiveness such as systemwide vehicle miles, vehicle hours, and queue size. The output also includes link flows, v/c ratios, and travel times at the end of each simulation period.

The optimization algorithm developed to select alternative routes is a heuristic procedure that has three objectives: minimization of travel time, reduction of congestion, and safety—that is, reduction of secondary incidents. Traffic is never diverted from the MassPike unless it reaches a certain level of congestion and extra capacity is available on alternative routes. The selection criterion for choosing alternative paths is a combination of v/c ratio and travel time. Briefly stated, the goal is to divert traffic to alternative routes only if all three of the following criteria are met:

1. MassPike will experience congestion if traffic is not diverted;
2. Uncongested alternative routes are available; and
3. Travel times on the alternative routes are less than the travel time on MassPike.

All alternative routes start from the MassPike exit (an exit or two upstream of the incident site) and return to the MassPike at an exit or two downstream of the incident site. During an incident, capacity and volume of the incident link are reduced. The volumes on the MassPike links downstream from the incident link are reduced by the difference between the MassPike demand volume and the reduced capacity of the incident link. Capacity on the incident link is reduced only for the anticipated duration of the incident. The model can also handle capacity changes at other times on other routes. The simulation is repeated once every minute until the queue at the incident time site becomes zero and the incident is cleared.

More than 100 simulations were conducted. Incidents were modeled between Exits 10 and 11A, with alternative routes extending from Exit 9 to Exit 12. Two- and three-lane closure incidents of various durations (30, 60, and 120 min) at different times of the day (7:00 a.m., noon, and 3:00 p.m.) were considered.

An analysis of the results obtained from the simulations shows that these results are reasonable and are consistent with the optimization objectives of the traffic diversion algorithm. For example, queue size increases as the duration of the incident increases, provided everything else remains constant. An incident of 30-min duration at 7:00 a.m. between Exit 10 and 11 would result in a queue of 222 vehicles if traffic were diverted from both Exit 9 and Exit 10. A similar incident of 60-min duration would lead to a queue of 455 vehicles. An incident of 120 vehicles would also lead to a queue of 455 vehicles, because the peak hour finishes at 8:00 a.m.; after
that, volume on the MassPike drops. As more vehicles are diverted off the MassPike, queue size decreases, but it takes longer for the system to return to normal. This insight into the system performance is very important if the system is to be extended for more than one incident in a given time period.

The use of “expert” algorithms in diverting traffic results in a reduction in congestion on the MassPike and systemwide savings in vehicle hours. Consider a two-lane closure incident of 30-min duration between Exits 10 and 11 at 7:00 a.m. If no diversion were exercised, a queue of 455 vehicles would be formed at the MassPike. On the other hand, if diversion were performed using the expert algorithm, queue length would reduce to 221 vehicles, and 8 vehicle-hr would be saved. A similar incident at noon would result in a queue of 111 vehicles. The diversion would reduce the queue length to zero and result in a savings of 20 vehicle-hr.

The algorithm never diverts traffic from more than one exit unless there is enough traffic at the incident link to suggest queue formation. For example, for a two-lane closure incident at noon, traffic is never diverted from Exit 9 because no queues are anticipated. Thus, the algorithm works consistently with the objective of not diverting traffic on longer paths, unless necessary.

CONCLUSION

Given the nature of the incident management problem, which involves many interacting agencies and the utilization of pre-selected incident response plans, the use of a knowledge-based expert system as a support tool for the traffic system manager is highly recommended. The application to the Massachusetts Turnpike test case provides support for the suitability of this approach. For the system to become operational, more testing and research are required. It would be useful to test the system on an actual section of the Massachusetts Turnpike rather than in simulation. This system could be used as a starting point for a turnpike-wide incident management system. The application of such an expert system would require a sophisticated communication system to disseminate information; the acquisition of such equipment should be considered. The dynamic network model for motorist diversion was tested successfully without requiring the input of information that is not readily available, such as origin-destination data. Using the discretization and incremental assignment of diverted motorists, the model can provide results in real time. However, as with any effective alternative route information or diversion system, effective implementation presumes real-time information about traffic conditions on the alternative routes. Therefore, an effort should also be made to collect such data.

The challenges in this area of research include modeling motorist response to information and considering more than one incident at a time. In the first area, it is important to realize that many of the modeling assumptions generally accepted by the transportation community as part of metropolitan transportation planning do not apply. We cannot assume that motorists will perform as we would like. The importance of motorist response has long been recognized as an issue that requires considerably more research before it can be incorporated as a meaningful parameter in any route diversion model. In the second area, the expert system approach appears promising because of its flexibility and adaptation to the problem at hand.

REFERENCES


Publication of this paper sponsored by Committee on Freeway Operations.