

Research Article

Development and Evaluation of a Control System for Regional Traffic Management

John L. McLin¹ and William T. Scherer²

¹ MITRE Corporation, F063—Airspace and Airport Analysis, McLean, VA 22102, USA

² Department of Systems and Information Engineering, University of Virginia, Charlottesville, VA 22904, USA

Correspondence should be addressed to John L. McLin, jlm3xy@virginia.edu

Received 24 October 2010; Accepted 12 February 2011

Academic Editor: Ram M. Pendyala

Copyright © 2011 J. L. McLin and W. T. Scherer. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Traffic congestion is a worsening problem in metropolitan areas which will require integrated regional traffic control systems to improve traffic conditions. This paper presents a regional traffic control system which can detect incident conditions and provide integrated traffic management during nonrecurrent congestion events. The system combines advanced artificial intelligence techniques with a traffic performance model based on HCM equations. Preliminary evaluation of the control system using traffic microsimulation demonstrates that it has the potential to improve system conditions during traffic incidents. In addition, several enhancements were identified which will make the system more robust in a real traffic control setting. An assessment of the control system elements indicates that there are no substantial technical barriers in implementing this system in a large traffic network.

1. Introduction

Traffic congestion occurs when there is insufficient capacity to meet the prevailing demand for the transportation infrastructure and incurs consequences that include increased delay, pollution, and wasted fuel. Addressing the congestion problem is a challenge for transportation planners because of increasing traffic demand and the limited opportunities to build additional capacity. As a result, transportation planners have placed considerable focus on traffic management solutions which more effectively utilize the existing transportation infrastructure. These efforts are aided by Intelligent Transportation Systems (ITS) technologies which are defined as “advanced electronics and communication technologies to enhance the capacity and efficiency of the surface transportation system” [1].

Regional traffic control can be considered a macroscopic approach to traffic management and is a means to deal with both recurrent (e.g., daily congestion) and nonrecurrent congestion events (e.g., incidents, bad weather). The concept of regional traffic control centers on the coordination of traffic management efforts between jurisdictions during normal traffic operations as well as during traffic incidents and

emergency events. Coordination is realized by exchanging data, implementing compatible timing plans, and developing coordinated emergency management plans. Tactical regional traffic management has been implemented in some metropolitan areas by establishing regional traffic control organizations which provide real-time traffic management throughout the region (e.g., Houston’s Transtar). In the future regional coordination may become required traffic management in metropolitan areas because “as the spatial extent of recurrent or nonrecurrent traffic congestion increases, control strategies by which each subnetwork operates in virtual isolation become increasingly ineffective” [2].

Coordinated traffic diversions are an effective method that can be employed in regional traffic control during nonrecurrent events to help improve network conditions. Spare network capacity is utilized to minimize the delay that drivers’ experience. In addition, coordinated efforts minimize the opportunities for antagonistic control actions. For example, the diversion of vehicles from an interstate onto an arterial can result in secondary congestion without the proper assessment of the capacity of the diversion route as well as the proper adjustment of timing plans on the diversion route.

One of the significant challenges in regional traffic control is the sheer number of control decisions that are required. Traffic managers have several choices during nonrecurrent events. First, they can implement predefined control plans, or they can develop traffic control plans on an ad hoc basis. Predefined plans are likely a set of broad actions that are applied to a large number of incident scenarios and may not represent the best management solution to the current problem. Ad hoc plan development can be difficult because of network dynamics under congestion conditions and the complex interactions between control measures [3]. In this situation, the effectiveness of traffic management becomes highly dependent on the traffic manager's experience.

The discussion above highlights the need for a regional traffic control system that can rapidly select and implement the best possible control decisions during nonrecurrent events. This paper discusses the development and evaluation of a control system designed to provide regional traffic control. The paper also discusses several key elements of the control system including the collection of real-time traffic data in large traffic networks and the identification of nonrecurrent conditions. Results show that the control system has the potential to improve system performance during nonrecurrent events.

2. Introduction to the Regional Control System

Figure 1 displays the operational environment for the control system that is designed to provide integrated traffic control decisions during nonrecurrent events. The main control measures that can influence traffic behavior include arterial traffic signals, ramp meters, and variable message signs (VMS). It is also possible to include more exotic control measures such as reversible flow lane indicators, variable speed limits, and in-vehicle information systems. Signal timing plan changes and VMS-induced diversions are the control measures explored in this research.

The control system is based on a partition of the entire traffic network into individual subnetworks and furthermore into sections. Subnetworks are loosely defined extended road segments such as a long arterial roadway with multiple signalized intersections or an interstate road segment that may stretch several miles. A section is a much shorter road segment and the grouping of consecutive sections comprises a subnetwork. Obtaining section data is important for the control system because it enables the rapid detection of nonrecurrent conditions as well as provides the ability to pinpoint where the bottleneck has occurred.

A nonrecurrent event is the key condition that is required for the system to initiate a control action; otherwise, the background traffic control plans remain in effect. Current traffic data for each section is obtained from probe vehicles, as well as system sensors, and are analyzed for abnormal conditions. Probe vehicles are vehicles that traverse the network and transmit position data using various means such as cell phone GPS or a transponder. The probe data is used to determine vehicle speeds and other systems metrics.

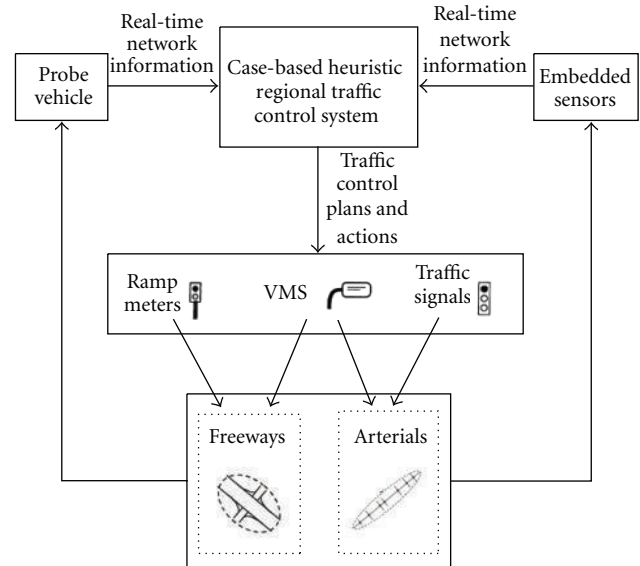


FIGURE 1: Operational environment for the regional control system.

System sensors can be represented by a mix of technologies; however, in the tests performed here are assumed to be in-loop ground sensors. The system sensors are vital in collecting traffic volumes at key entrance locations into the regional network. A traffic incident (i.e., a nonrecurrent event) is declared if smoothed flow and speed values fall below statistically established thresholds.

The control system combines several artificial intelligence techniques including Case-Based Reasoning (CBR) and Genetic Algorithms (GAs). CBR systems operate under the philosophy "that similar problems can be solved by similar solutions and that problems tend to recur" [4]. Case-based reasoning has been applied within the domain of traffic control including real-time driver routing [5] and traffic operator decision support [6]. GAs are a metaheuristic search method based on evolution theory, and have a history of finding good solutions to large optimization problems. In particular, GAs have been used to develop signal timing plans for various traffic conditions in signalized traffic networks [7–9].

The control system uses CBR as the mechanism to select proper traffic management actions during the incident. The case in a case base that is most similar in features to the current incident scenario is selected and the control actions associated with that case are then implemented. It is also conceivable to introduce real-time control plan development within the CBR control system (when there isn't a suitable case that describes the current scenario). A GA helps populate the case base by finding the best combination of traffic management actions for all of the subnetworks and utilizes a traffic flow model to assess the "fitness" of control actions combinations. The case base is populated by iteratively solving for a number of nonrecurrent scenarios. The system is able to accommodate a greater variety of nonrecurrent events as the number of cases in the case base increases.

Several feedback loops are incorporated into the control system. The first feedback loop is used to ensure that the best case from the case base is selected for the current conditions at the incident location. In general, the conditions at the traffic incident will continue to deteriorate until an equilibrium point is reached. Therefore, once the incident has been identified, the system will continue to assess system conditions and implement the best control plan for several iterations. The second feedback loop is used to determine when the control algorithm is no longer needed. If the speed at the road section where abnormal conditions were first identified returns to values above a predefined threshold speed, then the control actions are terminated.

3. Elements of the Regional Traffic Control System

In this section, the specific elements of the regional control system are described in detail including the structure of the case base and the process in which the case base is populated. The traffic flow model used by the GA is also presented here. The section concludes by describing how nonrecurrent incident detection is accomplished.

3.1. Case-Based Reasoning (CBR). CBR is an efficient method to rapidly find solutions to a given problem in a limited time window. This is especially important for real-time traffic management because of the rapid system degradation that is experienced when a nonrecurrent event occurs. A case base can be defined as a collection of a large number of cases, each containing descriptions and solutions to specific problems. There are three major parts of a case including the problem description, the solution to the problem, and the outcome (state of the system). The problem description in each case is characterized by a set of features which contain enough information to completely encapsulate the relevant facets of the problem.

The case base reasoning process shown in Figure 2 includes the sequential steps of: retrieving, reusing, revising, and retaining solutions to problems. The CBR process begins after a new problem has been identified and quantified so that it can be compared to the problem features in each case. The retrieval process consists of selecting the case which has the highest similarity to the current problem. Subsequently, the solution for the selected problem in the case base is then extracted during the reuse process. Revision entails modifying the proposed solution to more adequately reflect the current problem. Finally, the revised solution is implemented and added to the case base if it has been determined to be a good solution.

3.2. The Structure of the Traffic Control Case Base. The structure for the case base used by the traffic control system is illustrated in Table 1. The case base consists of n cases and each case describes a specific nonrecurrent scenario and includes the solution for the scenario in terms of an action vector. The description of each feature in the regional control case base is given below.

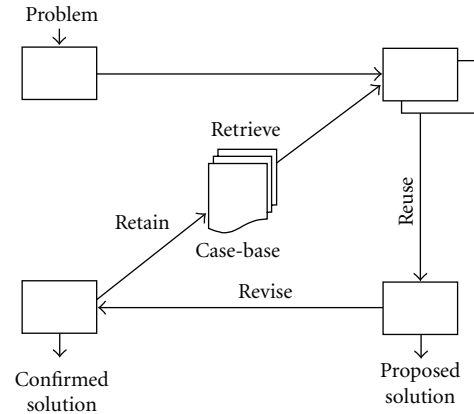


FIGURE 2: Case based reasoning cycle [10].

- (i) S_i is the numeric identification of the subnetwork where the incident occurred in the i th case.
- (ii) P_i is the section position, which is the source location for the incident in the i th case.
- (iii) R_i is the remaining capacity at the source section during the incident scenario in the i th case. In most circumstances the remaining capacity is equal to the incident flow.
- (iv) \mathbf{O}_i is a vector of the previous flow values for the past m minutes which represent the major inflows into the traffic region. The elements of \mathbf{O}_i are the flow values for each of the j inflow locations.
- (v) \mathbf{E}_i is the vector of traffic flows into the region which will occur during the nonrecurrent event. The elements of \mathbf{E}_i are the flow values for each of the j inflow locations.
- (vi) \mathbf{A}_i is a vector that represents the best traffic management actions for the i th case. The elements of vector $\mathbf{A}_i = [a_{i1}, a_{i2}, \dots, a_{ik}]$ correspond to a specific traffic timing plan, management action (diversion), or combination of both for each of the k subnetworks.

3.3. Populating the Regional Traffic Control Case Base. Figure 3 illustrates the process of how the case base for the traffic control system is populated and focuses on evaluating multiple incident scenarios within a traffic flow model. The primary inputs for the process are a list of incident scenarios, a set of traffic control timing plans (for arterial roads and ramps), and the available traffic management actions.

Individual nonrecurrent incident scenarios are defined by (1) the traffic flows which entered the network prior to the incident, (2) the flow that enters the network during the incident time interval, and (3) the remaining capacity at the source traffic section for the nonrecurrent event. Traffic control plans for the signalized subnetworks are developed using SYNCHRO which is a software package used by traffic engineers to develop timing plans for a given set of network traffic flows [11]. The other input is the traffic management (i.e., congestion management) actions, which

TABLE 1: Representation of the structure of the case base.

Case number	Sub network	Incident position	Incident flow	Previous flow vector	Expected flow vector	Action vector
1	S₁	P₁	R₁	O₁ [<i>o</i> ₁₁ , <i>o</i> ₁₂ , <i>o</i> _{1j}]	E₁ [<i>e</i> ₁₁ , <i>e</i> ₁₂ , <i>e</i> _{1j}]	A₁ [<i>a</i> ₁₁ , <i>a</i> ₁₂ , <i>a</i> _{1k}]
—	—	—	—	—	—	—
—	—	—	—	—	—	—
—	—	—	—	—	—	—
n	S_n	P_n	R_n	O_n [<i>o</i> _{n1} , <i>o</i> _{n2} , <i>o</i> _{nj}]	E_n [<i>e</i> _{n1} , <i>e</i> _{n2} , <i>e</i> _{nj}]	A_n [<i>a</i> _{n1} , <i>a</i> _{n2} , <i>a</i> _{nk}]

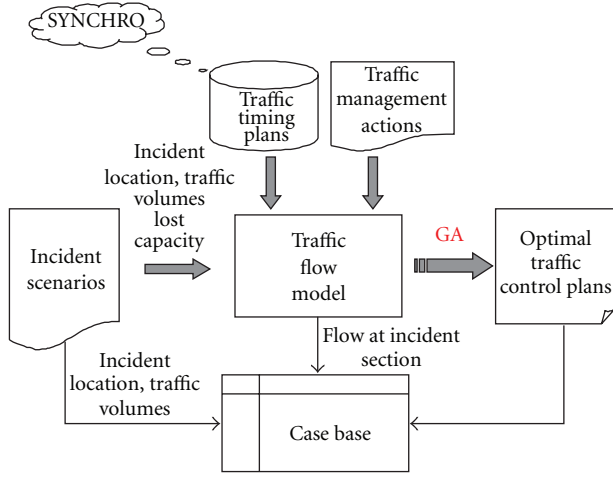


FIGURE 3: Illustration of the process for populating the case base.

identify diversion actions that occur at different points in the network. A diversion action also specifies a diversion percentage of the through volume on roadways.

The population of the case base begins once the timing plans, management actions, and incident scenarios have been compiled. The traffic volumes (prior to the incident and during the incident) are used as the input volumes for the traffic flow model. In addition, the remaining capacity at the incident section given in the scenario is used to adjust the capacity at the incident location in the traffic model. The GA is then executed to determine the best combination of subnetwork timing plans and diversion control actions that produces the least delay. The traffic flow model acts as a “fitness function” within the GA and provides the delay associated with a selected combination of timing plans and control actions. The elements of the incident scenario as well as the solution found by the GA are entered into the case base, and the process is repeated for each incident specified in the incident scenario list.

3.4. Matching Traffic Incident Conditions to the Case Base. Matching the current problem to the most relevant case in the case base is an important process within the CBR system. When an incident (i.e., the current problem) has been identified, it is quantified into a vector of features (problem description vector), which are used to compare against the features in each case in the case base. The method used here to determine the best case for the current problem

is a combination of template retrieval (TR) and a nearest neighbor (NN) algorithm.

Template retrieval limits the case search to only certain cases which contain a specific value for a feature, or range of values for a feature, and helps reduce the search space. In this research, the retrieval limit is set such that only cases pertaining to the subnetwork associated with the problem section are considered. This is important because the selection of a case from another subnetwork can lead to the implementation of poor traffic management solutions.

Nearest neighbor algorithms are a method used to find the similarity between two cases and can be determined by calculating the weighted Euclidean distance between a case vector and problem description vector. The similarity $S(x, y)$ between the two vectors is found using (1), where x represents the case vector for a specific case in the case base and y represents the vector which describes the current problem. K is the total number of features to be compared and consists of the values in columns 2–6 in Table 1,

$$S(x, y) = \sqrt{\sum_{f=1}^K w_f \times D(x_f, y_f)^2}. \quad (1)$$

In addition, w_f is the weight value assigned to feature f and the difference function D is defined as

$$D(x_f, y_f) = (x_f - y_f). \quad (2)$$

The differences in magnitude between feature elements (e.g., flow values, section position) necessitate the normalization of these values so that the weighting scheme given in (1) provides the desired outcome.

3.5. The Traffic Flow Model Used by the CBR Traffic Control System. The traffic flow model is the heart of the case-based system because the selection of timing plans and actions by the GA are based on their ability to minimize delay within the flow model. At its core, the flow model propagates input flows throughout the network and determines delays based on Highway Capacity Manual (HCM) delay equations at traffic intersections and uses Bureau of Public Roads (BPR) equations to determine travel delay on roadways. Queue delays and queue lengths are also based on HCM equations and are calculated for road sections when the through volumes on road sections exceed the capacity of the section. The total delay is determined after the GA has selected a candidate set of plans for all of the subnetworks and the plans which yield the lowest delay are used as the best operational management actions for the given incident.

The traffic flow model uses a series of nodes and links to model the traffic network. Nodes represent either an intersection or a point where traffic from different streams converge or diverge. Links connect the nodes together and represent the roadways in the traffic networks. Traffic flows are propagated through the network by introducing input flows at certain nodes and the values are obtained from the incident scenario description. At each node there are links that feed into and out of the node. The traffic on links that feed into a node are split based on time-dependent percentages obtained from historical data and propagated on the links emanating out of the node.

A subset of the nodes in the traffic flow model represent signalized intersections and the calculated delays are based on the method outlined in the HCM. The total control delay for a specific movement consists of the uniform delay, the incremental delay, and the initial queue delay.

The uniform delay provides the control delay for vehicles, which are assumed to arrive uniformly throughout the analysis period. The incremental delay estimates the delay for vehicles that have additional delay due to over-saturated conditions and cause vehicles to remain at a traffic light for more than one signal cycle. The initial queue delay is used to model vehicle queuing that is present when the system first detects an incident. This addresses the reality that current network conditions are a function of traffic flows in prior/previous time periods. The initial queue delay is calculated by using the previous flow vector (column 5 in Table 1) to determine whether the system capacity can meet the previous flow demand at each intersection. If the capacity is not sufficient then an initial queue delay is calculated for those intersections with insufficient capacity.

The travel time delay along a road segment is obtained by determining the difference between the free flow travel time and the travel time obtained using the Bureau of Public Roads (BPR) equation. Originally developed in the 1950s, the BPR equation relates the travel time on a road section to the volume present on the road, the road capacity, and the free flow travel time [12]. The free flow travel time for a road segments is defined as the time necessary to traverse a road segment at the prevailing speed limit.

The queuing delay that results when the current road capacity does not meet the current demand is also calculated to improve the accuracy of the traffic flow model. Note that this delay is different from the queuing delay caused by traffic signalization. If the capacity of a section does not satisfy the traffic demand, the queue length is determined using HCM calculations. A queue delay is calculated using (3) for all links whose sum length is less than the queue length. The queue delay is found by finding the difference between the section demand (v) and section capacity (c) and then multiplying by one half the square of the incident duration (T),

$$QD = [(v - c)] * \left(\frac{T^2}{2} \right). \quad (3)$$

3.6. Incident Detection by the Control System. The control system must have the capability to recognize when control actions are needed and also when to terminate them. This

capability can be achieved by using automatic incident detection (AID) algorithms, which were first introduced in the 1970s to help traffic operators learn of traffic incidents on interstates [13]. AID algorithms are evaluated primarily on two criteria including the false alarm rate (FAR) and mean time to detect (MTTD) [14]. There generally exists a trade-off between these measures in that an AID algorithm with a low FAR will have a high MTTD (and vice versa).

The standard normal deviate (SND) is a statistical approach to incident detection first developed by the Texas Transportation Institute (TTI) and is based upon the calculation of the SND for a number of traffic parameters [15]. The SND is calculated by finding the difference between the observed traffic parameter and the mean parameter value, and then dividing by the standard deviation of the traffic parameter. An incident is declared if the SND for the traffic parameter(s) exceeds a threshold, which is typically a predefined number of standard deviations from the mean.

The application of the SND on interstate roadways is straightforward since there are minimal perturbations and the data (e.g., speed) is generally Gaussian distributed. Signalization as well as other effects on arterials introduces a great deal of variability within the traffic parameter data. To minimize variability, the traffic data is smoothed by averaging it with the data from previous time periods. Another problem is that the SND approach assumes a Gaussian distribution and arterial traffic data (smoothed) might not follow this statistical distribution. However, the implication of Chebychev's inequality allows the reasonable application of this method to nonnormal data distributions [16].

Incident detection by the control system is achieved by calculating the smoothed flow and speed for each section in the network. Empirical experimentation indicates that the combination of flow and speed results in a very low FAR and reasonable MTTD. At each time interval, the control system collects flow and speed data (via probe vehicles) for each section and smoothens them with the most recent observations. An incident is declared if the SND of both parameters falls below their thresholds.

4. Evaluation of Traffic Control System

The performance of the regional control system was evaluated by comparing the system and subnetwork performance measures against other control strategies during varying nonrecurrent events. In addition, the evaluation provides an illustrative example of the control method as well as the potential for the control system to improve conditions during congestion events.

4.1. Simulation Environment for Testing the Regional Traffic Control Algorithm. Typically, it is not possible to test prospective traffic control systems in a field setting and traffic simulation is the next best alternative to evaluate them. AIMSUN (Advanced Interactive Microscopic Simulator of Urban and Non-Urban Networks) is a microscopic simulation program that has been used as an analysis tool in a number of transportation studies including urban traffic

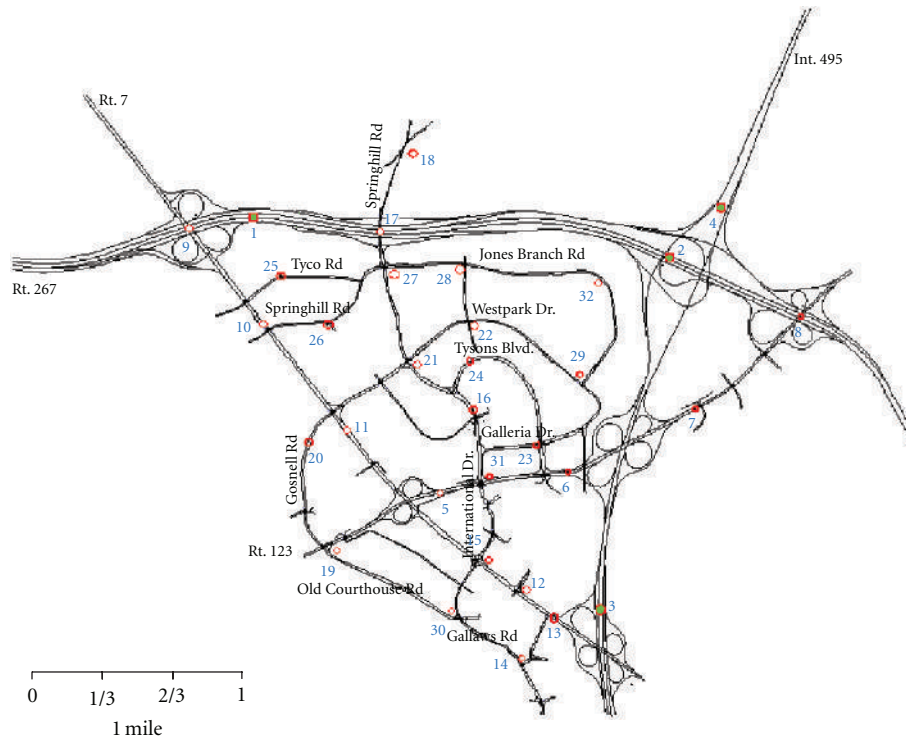


FIGURE 4: Tysons Corner simulation model.

network control [17], regional traffic management [18], and ramp metering [19].

Testing of the control system was performed using an AIMSUN simulation model of Tysons Corner, VA, which is located 8 miles west of Washington DC (Figure 4). Tysons Corner is ideal for analyzing integrated regional traffic control because it contains several major roadways including the Dulles Toll Road/Route 267 (highway), I-495 (interstate), Route 7/Leesburg Hwy (arterial), Route 123/Chain Bridge Road (arterial), and many local roads.

AIMSUN enables the evaluation of new traffic control strategies through the use of the AIMSUN API, which allows for external changes during simulation runs to signal timing parameters and ramp meter rates, as well as the initiation of vehicle diversions. In addition, the API allows the control system to access simulation statistics as they are being generated during the simulation run. The regional control algorithm is compiled as a dynamic linked library and this allows it to communicate with the simulator through the API (Figure 5). Throughout the experiments that were performed, the use of probe vehicles was mimicked by using data collected by the traffic simulator for each link in the network.

4.2. Populating the Case Base and Establishing Control Boundaries. The Tysons Corner network was partitioned into 7 subnetworks including three arterial subnetworks and 4 unsignalized subnetworks (Figure 6). Subnetworks (1–3) are arterial subnetworks and (4–7) are interstate subnetworks. The number of partitions was based on significant traffic

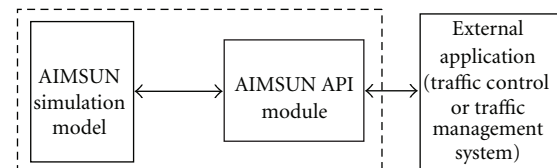


FIGURE 5: Controlling the simulation network using an API.

flow movements and also by the artificial boundaries created by crossing roadways. For example, subnetwork 1 (SN1) is bounded by I-495 to the south and by the Dulles Toll road to the north. The major flows of traffic on SN1 are introduced by traffic exiting off of I-495 and travelling north on SN1 as well as traffic exiting the Dulles Toll Road (SN6) onto SN1 and travelling south.

In the experiments performed, the control system provides control for a subset of all the subnetworks and includes subnetworks 1 through 4 (SN1–SN4). This provides over 12.3 linear miles of control coverage and also provides signalized traffic control for 11 intersections on SN1, 9 intersections on SN2, and 7 signalized intersections on SN3.

The case base was populated by evaluating 123 incident scenarios, each with an assumed incident duration of 60 minutes. The cases were generated by parametrically adjusting the volumes introduced into the regional network and by adjusting the lost capacity at a specific section (used to model incident severity). A set of traffic signal timing plans had previously been developed in SNYCHRO for each of the signalized subnetworks and included 14 timing plans for SN1,

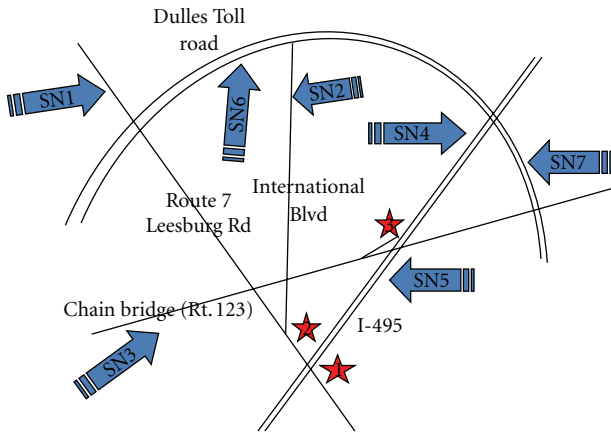


FIGURE 6: Tysons Corner subnetworks.

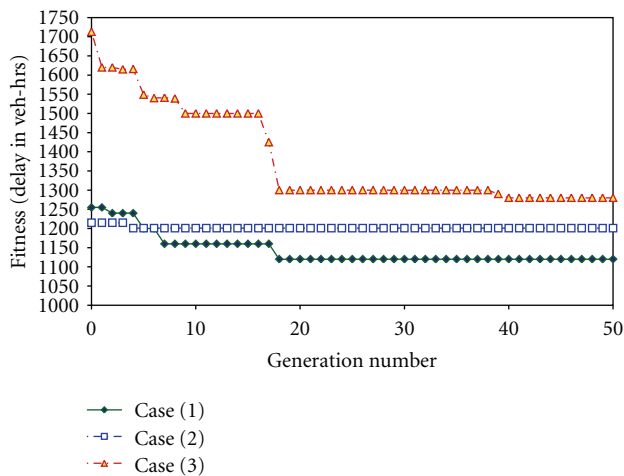


FIGURE 7: Fitness value as a function of GA generations.

10 timing plans for SN2, and 22 timing plans for SN3. The timing plans are based on various volume scenarios input into SYNCHRO which then calculated the optimal plan for each scenario. Diversion traffic management actions (diversion levels) were defined for each of the diversion points in the network and are described in detail in Section 4.4.

The best control plan (i.e., timing plan & diversion level) for each scenario was found by running the GA for 50 generations. Preliminary analysis indicated that the number of generations was sufficient to reach convergence, which is defined by the lack of change in fitness value after a certain number of iterations (Figure 7). The GA was structured such that it only selected control plans that were feasible and this helped reduce the time required to reach convergence. The best control plan for each scenario was then used to populate the case base.

4.3. Performance Measures to Evaluate the Control System.

In general, most performance measures for large networks are derived from a set of core measures including speed, travel time, and delay [20–22]. AIMSUN is able to calculate and record these measures along with a number of other

measures for the entire system as well as for individual sections and subnetworks. The simulator collects performance statistics into a database at user-defined intervals for post simulation analysis. For example, system level statistics can be recorded every 2 minutes. Delay is the most relevant measure because the GA selects the timing plans based on the combination that produces the lowest delay in the traffic flow model.

4.4. Describing the Experimental Design. The performance of the regional control system was evaluated by assessing its ability to improve traffic conditions during nonrecurrent incidents, which were varied by location and duration.

An arterial incident was introduced on Leesburg Road affecting the northwest bound traffic on SN1 and consisted of one full lane closure along with approximately 80% of a second lane at section 19 (see Figure 8(a)). The location of the incident (indicated by the star in the figure) makes it feasible for vehicles to divert onto another roadway (International Dr.).

Figure 8(b) shows the incident location on an interstate roadway (I-495) which blocks 3 full lanes at section 9 on SN4. The incident location also enables drivers to divert upstream from the incident onto Chain Bridge road (SN3). The arrows in the figures indicate the route taken by drivers when a diversion is either indicated by the control system via VMS or through expected driver behavior.

Four scenarios were evaluated using AIMSUN including:

- (i) Scenario 1: 50 minute incident duration on an arterial road (Leesburg Road).
- (ii) Scenario 2: 90 minute incident duration on an arterial road (Leesburg Road).
- (iii) Scenario 3: 60 minute incident duration on an interstate road (Interstate I-495).
- (iv) Scenario 4: 90 minute incident duration on an interstate road (Interstate I-495).

For all scenarios, the simulations were run beyond the time when the incident ended which allowed all vehicles affected by the incident (e.g., queued vehicles) to exit the system. This aided in the proper calculation of system-wide performance results. The following three control strategies were evaluated for their effectiveness:

Strategy A. Baseline traffic control where the default PM timing plan is used throughout the simulation. Drivers are assumed to stay on the congested roadway.

Strategy B. Traffic control where the default PM timing plan is used throughout the simulation; however, it is assumed that a certain percentage of drivers on the affected incident road divert.

Strategy C. The regional traffic control system is used which selects a control plan based on current conditions once an incident has been confirmed.

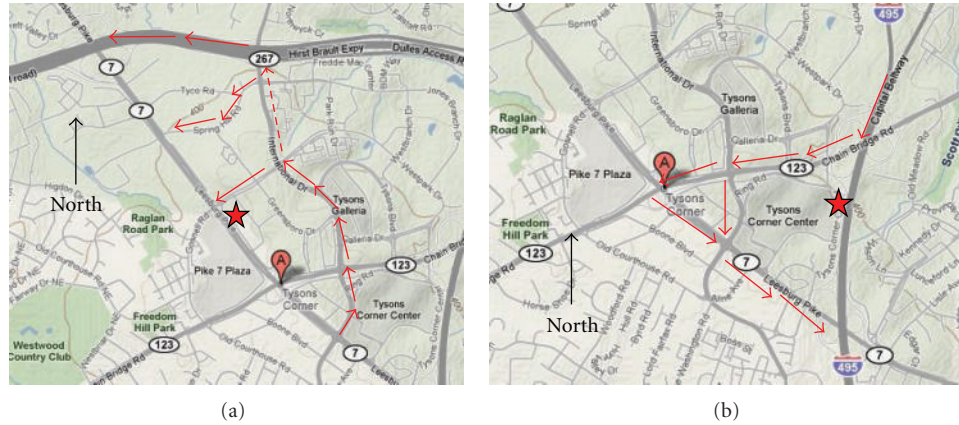


FIGURE 8: (a) Incident on Arterial Roadway; (b) Interstate Roadway.

The true diversion rate for drivers is an unknown random variable, and therefore when strategy B was evaluated it was assumed that the driver's diversion rate varied. During each simulation, a different diversion rate was used and the results for Scenario B are presented as an average for all the diversion rates. For scenarios 1 and 2, the diversion rate was between 10% and 50% of the through volume on Leesburg Road and in scenarios 3 and 4 the diversion rate ranged between 5% and 15%. Diversion rates higher than the prescribed range resulted in excessive queuing at the entrance to the diversion route and it seemed unrealistic to assume that drivers would opt to select this action.

4.5. Simulation Results. Simulations of the Tysons Corner model were run for all scenario/strategy combinations during the PM peak rush hour (5 PM) and each scenario/strategy combination was executed using 10 replications. The numbers of replications was limited because of the scenario/strategy combinations as well as the time required to run a replication. Nevertheless, the number of replications was sufficient to perform statistical analysis on the results. Because AIMSUN is a stochastic simulator, the control system was exposed to varying traffic volumes and vehicle behaviors.

On average, the arterial incidents were detected 6.2 minutes after the incident occurred while interstate incidents were detected 5.2 minutes after they occurred. The control system terminated control actions 3.8 minutes longer during scenario 2 versus scenario 1 (15.6 min versus 11.8 min) after the incident had been removed. This was due to the additional congestion build-up on the arterial behind the incident location during scenario 2. Interstate incident management ended 10 minutes after the incident was removed in scenarios 3 and 4 and indicates that incident conditions were isolated to the vicinity of the incident with negligible queue growth.

Case 30 was the most frequent case selected by the control system when Scenarios 1 and 2 were evaluated using Strategy C, and calls for a diversion of 75% of the through volume from Leesburg Pike onto International Blvd.

The selection appears reasonable since case 30 describes a capacity loss of 1.8 lanes on section 18, which is behind (upstream) the actual incident location. During Scenarios 3 & 4 the control system selected case 94, and also appears to be reasonable because it describes an incident where capacity is reduced by 2 lanes at section 8 on interstate I-495. Case 94 includes a diversion of 15% of the through (southbound) volume onto Chain Bridge Road. Both cases include signal timing plans for each of the signalized subnetworks under control.

Table 2 provides system-wide results for each of the strategies when tested during the incident scenarios. The results show that strategy C is able to provide the best overall reduction in system delay for all scenarios. In addition, the difference between strategy C and the other strategies increases as the incident duration increases. When compared to the base control case (i.e., strategy A) the regional control system improves upon system delay by almost 13% for scenario 1, 26% for scenario 2, 5% for scenario 3, and 8% for scenario 4. The performance measures for speed and travel time also are better when strategy C is compared to the other control strategies for each of the scenarios.

The impact of strategy B on system performance during the evaluation of each scenario is interesting. In scenarios 1 and 2, strategy B improved upon the baseline traffic management (strategy A) and confirms that the diversion of drivers helps improve system conditions. However, strategy B appears to perform the same, or slightly worse than strategy A during scenarios 3 & 4 (interstate incident). This is likely due to the large volume of highway vehicles that divert onto the arterial and suffer delays because there are no signal timing plan adjustments on the diversion roadways.

The improvement in total system delay for strategy C over the baseline strategy during scenarios 3 and 4 is not as large when compared to the improvement during scenarios 1 and 2. This can be attributed to several factors. First, vehicles that divert from the interstate to the arterial face signalization delays, which are comparable in duration to the incident delay on the mainline. Second, the AIMSUN API does not allow changes in signal offsets during simulation runs, even though SYNCHRO provides offset values as a

TABLE 2: System-wide results for each scenario and control strategy.

Scenario	Strategy	Speed mph	Travel time sec/mile	Delay time sec/mile
1	Strategy A	22.7	60.6	29.6
	Strategy B	23.6	58.4	27.2
	Strategy C	24.1	57.3	25.8
2	Strategy A	19.9	69.3	38.1
	Strategy B	21.0	65.7	34.6
	Strategy C	23.3	59.6	28.1
3	Strategy A	24.8	55.4	24.1
	Strategy B	24.6	56.0	24.8
	Strategy C	25.2	54.8	23.0
4	Strategy A	24.0	57.3	26.0
	Strategy B	23.9	57.7	26.1
	Strategy C	24.8	55.5	24.0

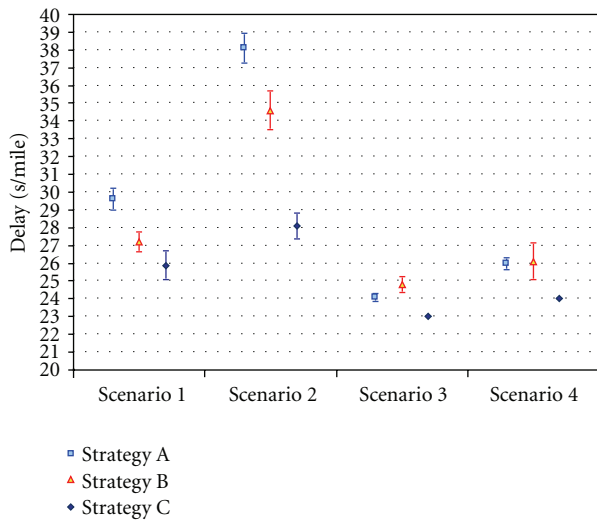


FIGURE 9: Confidence intervals on system delay for scenarios.

component of the timing plans that it creates. Observations during simulation runs for scenarios 3 and 4 indicated that the progression of vehicles was poor when strategy C was evaluated. This resulted in heavy delays for vehicles located on the diversion routes. Some preliminary experimentation indicated that modification to the timing plans resulted in better progression and improved the system performance measures.

The 95% confidence interval (CI) for the mean system delay values are provided in Figure 9 and show that strategy C is disjoint from all the other strategies during each scenario and implies that the delay improvement is statistically significant. In fact, hypothesis testing confirms that the mean delay value for strategy C were statistically different from strategy A at significance level $\alpha = 0.05$ for all of the scenarios. There are no intervals associated with strategy C during scenario 3 and 4 because the system delay value obtained during each replication was the same. The differing performance levels obtained by strategy B during

each scenario are highlighted by the calculated intervals. In scenarios 1 & 2 the CI for strategy B is disjoint from strategy A and clearly shows improved performance. However, in scenario 3 the interval delay values for strategy A are better than strategy B. In scenario 4 the CI for strategy B covers the interval for strategy A and is longer in length, thus strategy A appears to provide more reliability during this scenario.

4.5.1. Assessing the Extent of Congestion. The extent of congestion was analyzed by using recorded simulation speed data. A segment was classified as a congested segment if the speed value for a section was found to be lower than its statistical incident speed threshold. Table 3 provides the average and maximum congestion lengths that occur between the baseline and the regional control system. The results show that the congestion length is minimized for all scenarios when the control system is utilized. The congestion length found in scenario 2 when control strategy A was used would have been longer than 13,820 ft., except that the extent of the network model did not extend further out, and thus could not be calculated.

The distance between the incident location on Leesburg Road (Route 7) and the diversion point is approximately 5200 feet. Using strategy C, the maximum congestion extent reaches over 5800 ft for scenarios 1 and 2, although the actual physical queue never reaches the diversion location. The discrepancy between the actual queue length and the congestion extent for scenarios 1 and 2 is explained by the congestion which forms behind the diversion point and is counted as a part of the congestion extent.

The potential encroachment of the queue onto the diversion point is a concern because it limits the means to improve the traffic situation. Given an incident of longer duration it is possible that the access point could succumb to larger queue formations. This indicates that the control system should have additional feedback during an incident event to allow adjustments of the diversion rate, perhaps based on the growth of the queue. It is not entirely surprising that the queue formations extended so far in scenario 2 since

TABLE 3: Extent of congestion on incident roadway.

Scenario	Strategy A		Strategy C	
	Average congestion length (ft)	Maximum congestion length (ft)	Average congestion length (ft)	Maximum congestion length (ft)
Scenario 1	5793	9042	3381	5840
Scenario 2	10007	13820	4539	6565
Scenario 3	2246	5644	1207	2194
Scenario 4	3875	7261	1715	2859

TABLE 4: Effect of driver compliance on system performance.

Compliance rate	Scenario 1	Scenario 2
100%	13%	26%
75%	12%	23%
50%	10%	19%
25%	9%	15%

the best control plan selected by the GA is based on 60-minute incident durations.

4.5.2. Effect of Driver Compliance Rate on Control System Performance. The driver compliance rate is defined as the percentage of drivers who actually follow the given diversion information. The true level of compliance is unknown and varies greatly. In [23], several studies were discussed each of which provided widely different estimates of the actual driver compliance rate. For example, one study estimated the compliance rate to be between 27% and 44%, while another study mentioned estimated the compliance rate to be between 70% and 90%. Up to this point a compliance rate of 100% had been assumed for diversion actions and it was deemed important to assess the sensitivity of the control system performance to different levels of compliance.

The effect of different compliance rates on the system delay was tested for scenarios 1 and 2 by altering the percent of drivers that comply with the diversion advice. Three diversion compliance levels were tested (75%, 50%, and 25%), and implemented by multiplying a given control action diversion percentage by the reduced compliance percentage. Table 4 provides the results for the different compliance levels and includes the default 100% compliance to compare the reduced compliance rates. The results show the percentage change between the baseline control strategy (strategy A) and the regional control strategy (strategy C).

The results show that a decrease in the compliance rate corresponds to an increase in the system delay. It can be seen that the effect of lower compliance is more severe for the longer incident duration. For example, a reduction in compliance from 100% to 25% results in a 30% drop in the performance of the control algorithm during scenario 1, while the same reduction results in a 42% drop in the performance during scenario 2. However, it should be understood that the system performance comparison between strategy A and strategy C reflects both the diversion action and the change in timing plan. Therefore, the 9%

change in system delay at a 25% diversion rate is due to both the diversion and the timing plan.

5. Conclusions

Empirical testing of the control system shows that it may improve traffic conditions during nonrecurrent incidents. The algorithm was also able to demonstrate that it can operate as a true automated control system including incident detection, congestion management, and knowing when to terminate control actions. Results indicate that the system is more beneficial during incidents of longer duration, and in addition, the effectiveness of the algorithm is a function of the incident location and network topology. The level of driver compliance can also greatly affect system performance, especially for more severe incidents.

Analysis of the simulation results indicates that the control system can be improved in several areas. First, the progression parameters determined by SYNCHRO need to be included in simulation testing so that the control system can be evaluated more thoroughly. Second, the need for additional feedback response during traffic incidents was identified because of the danger in queue growth, which can block diversion passages. Third, adding incident duration as a feature of each case in the case base and populating the case base with incidents of different durations may lead to better control management actions. Finally, the traffic flow model can be made more accurate by improving the modelling of queuing phenomenon during traffic incidents or investigating other traffic models (e.g., mesoscopic simulation) to be used as the evaluation function within the GA.

It is important to assess the transferability of the control method presented here to larger traffic networks. The development of signal timing plans for each subnetwork is independent of each other, and therefore this component of the control system is quite scalable. Results show that the GA is able to converge quite rapidly to the best selection of timing plan for a specific incident scenario. For a larger network, the number of decision variables will increase substantially; however GAs are quite amenable to parallel processing and this may reduce the time required to find the best control solution.

One of the tenets of CBR is that similar problems are solved by similar solutions. Under this assumption, a case base with a moderately sized, yet diverse, case base can be applied effectively for a large number of incident scenarios. Furthermore, there exists the potential for finding the best

solution for individual incidents on an ad hoc manner (i.e., real time) in those situations where no case in the case base is similar enough. In such circumstances, the current network state could be used as input into the traffic flow model along with a parallel implementation of the GA to determine the best traffic management actions in a reasonable time window.

Perhaps the largest obstacle to implementation may be the institutional barriers which exist and prevent the implementation of regional traffic control. As traffic performance in metropolitan areas continues to degrade, these barriers will likely fall in response to the need to improve traffic conditions.

References

- [1] Cambridge Systematics, *Performance Review of U.S. DOT Innovative Finance Initiatives*, Federal Highway Administration, 2002.
- [2] Y. Pavlis and W. Recker, "A methodological framework for integrated control in corridor networks," in *Proceedings of IEEE Intelligent Transportation Systems Proceedings*, pp. 637–642, August 2001.
- [3] S. Hoogendoorn, B. De Schutter, and H. Schuurman, "Decision support in dynamic traffic management. Real-time scenario evaluation," *European Journal of Transport and Infrastructure Research*, vol. 3, no. 1, pp. 21–38, 2003.
- [4] D. B. Leake, *Case-Based Reasoning: Experiences, Lessons, and Future Directions*, MIT Press, 1996.
- [5] A. W. Sadek, M. J. Demetsky, and B. L. Smith, "Case-based reasoning for real-time traffic flow management," *Computer-Aided Civil and Infrastructure Engineering*, vol. 14, no. 5, pp. 347–356, 1999.
- [6] B. De Schutter, S. P. Hoogendoorn, H. Schuurman, and S. Stramigioli, "A multi-agent case-based traffic control scenario evaluation system," in *Proceedings of the 6th IEEE International Conference on Intelligent Transportation Systems (ITSC '03)*, p. 678, Shanghai, China, 2003.
- [7] H. K. Lo and A. H. F. Chow, "Control strategies for oversaturated traffic," *Journal of Transportation Engineering*, vol. 130, no. 4, pp. 466–478, 2004.
- [8] B. Park, C. J. Messer, and T. Urbanik, "Traffic signal optimization program for oversaturated conditions: genetic algorithm approach," *Transportation Research Record*, no. 1683, pp. 133–142, 1999.
- [9] G. Abu-Lebdeh and R. F. Benekohal, "Design and evaluation of dynamic traffic management strategies for congested conditions," *Transportation Research Part A*, vol. 37, no. 2, pp. 109–127, 2003.
- [10] I. Watson, "Case-based reasoning is a methodology not a technology," *Knowledge-Based Systems*, vol. 12, no. 5-6, pp. 303–308, 1999.
- [11] D. Husch and J. Albeck, *SYNCHRO User Guide*, Trafficware Ltd, 2001.
- [12] G. A. Davis and H. Xiong, *Access to Destinations: Travel Time Estimation on Arterials*, Minnesota Department of Transportation, Office of Research Services, 2007.
- [13] M. J. Cullip and F. L. Hall, "Incident detection on an arterial roadway," *Transportation Research Record*, no. 1603, pp. 112–118, 1997.
- [14] J. A. Khoury, C. T. Haas, H. Mahmassani, H. Logman, and T. Rioux, "Performance comparison of automatic vehicle identification and inductive loop traffic detectors for incident detection," *Journal of Transportation Engineering*, vol. 129, no. 6, pp. 600–607, 2003.
- [15] C. L. Dudek, C. J. Messer, and N. B. Nuckles, "Incident detection algorithm evaluation," *Transportation Research Record*, vol. 495, 1974.
- [16] S. M. Ross, *Introduction to Probability and Statistics for Engineers and Scientists*, Elsevier/Academic Press, New York, NY, USA, 3rd edition, 2004.
- [17] C. Diakaki, V. Dinopoulou, K. Aboudolas, and M. Papageorgiou, "Extensions and new applications of the Traffic Signal Control Strategy TUC," in *Proceedings of the Transportation Research Board Annual Meeting*, 2003.
- [18] J. Barceló and D. García, "Scenario analysis: a simulation based tool for regional strategic traffic management," in *Proceedings of the 9th World Congress on Intelligent Transport Systems ITS America*, 2002.
- [19] J. Hourdakis and P. G. Michalopoulos, "Evaluation of ramp control effectiveness in two twin cities freeways," in *Proceedings of the Transportation Research Board Annual Meeting*, 2002.
- [20] T. Lomax and D. Schrank, *The 2007 Urban Mobility Report*, Texas Transportation Institute, 2007.
- [21] S. B. Medley and M. J. Demetsky, *Development of Congestion Performance Measures Using ITS Information*, University of Virginia, 2003.
- [22] T. Lomax and S. Turner, *Quantifying Congestion Final Report*, 1997.
- [23] Q. Meng, H. L. Khoo, and R. L. Cheu, "An ATIS-based online dynamic control strategy for the urban expressway-arterial corridor systems," *Transportation Research Record*, vol. 2000, pp. 44–50, 2007.

