Arterial Incident Detection Integrating Data from Multiple Sources

Nikhil Bhandari, Frank S. Koppelman, Joseph L. Schofer, Vaneet Sethi, and John N. Ivan

An integrated incident detection system for an arterial street network being implemented for the ADVANCE project, an advanced traveler information system demonstration in the northwest suburbs of Chicago, Illinois, is described. Incidents will be detected using three distinct data sources: loop detectors, probe vehicles, and anecdotal sources. Specialized incident detection algorithms will process each of these data types separately. The outputs from the fixed detector, probe vehicle, and anecdotal source algorithms will be integrated by a data fusion process to determine the overall likelihood that an incident has occurred at any particular location. The incident detection system will also estimate the expected duration of the incidents and their effects on link travel times as a function of the type of incident.

Incidents are unexpected events that disrupt the flow of traffic on a segment of a roadway link and have significant effects on link travel times; examples are stalled vehicles, collisions, and materials spills. The effect of an incident is to reduce the capacity of the segment; if demand volume is high enough, this can result in queues, delays, and increased travel time on the link. Early detection of incidents can help traffic management agencies respond quickly, dispatch emergency vehicles to the incident site, and perhaps divert traffic to reduce delay. Detection of incidents also helps agencies warn the oncoming traffic and thereby reduce the danger of secondary incidents (1).

Recently, there has been much interest in increasing the efficiency of existing roadways by developing intelligent transportation systems (ITS) and particularly advanced traveler information systems (ATIS). One aim of these systems is to provide road users with real-time information on travel times and roadway status to reduce their individual travel times. An important component of these systems will be the ability to detect traffic flow disruptions on the road network and alert and divert potential users of the affected links.

The ITS field tests in the United States and Europe (e.g., ADVANCE (2), Pathfinder (3), TravTek (4), ALL-SCOUT (5), EURO-SCOUT (6)) used or will use multiple data sources such as traffic sensors, probe vehicles, video cameras, and anecdotal sources to collect real-time traffic information. The largest of these demonstrations is the ADVANCE project in suburban Chicago. ADVANCE will provide approximately 3,500 participants with real-time route planning information based on up-to-date travel times and incident information in the test area. ADVANCE drivers will be local residents, and information about recurrent congestion or navigational guidance will be of limited value as they will have considerable experience with the network. Incidents—unexpected nonrecurrent events on the network—can have significant effects on link travel times. Real-time information about an occurrence and its impact on travel time will be valuable even to travelers who are familiar with the network structure under normal conditions. Evidence from recruitment studies conducted for ADVANCE suggests that incident detection will be important for attracting participants to the project and sustaining their interest over the period of the demonstration (7). In general, knowledge about incident locations and their travel time effects will enable the ADVANCE project to give drivers more accurate estimates of link travel times for route planning and also to provide information on the reasons for increases in travel time.

ADVANCE will integrate information from three distinct data sources to detect incidents:

- **Fixed detectors**, which provide occupancy and volume data averaged over a fixed time interval (e.g., 5 min) for a specific section of selected network links;
- **Probe vehicles** participating in the demonstration project, which travel freely on the network and automatically report link travel times by radio; and
- **Anecdotal sources**, reports of particular events affecting traffic flow provided by people traveling on or monitoring the road network, including emergency services workers.

On the ADVANCE network, made up primarily of suburban arterial streets, fixed detectors providing volume and occupancy data are located approximately 350 ft upstream from selected signalized intersections on major arterials. A fraction of these will be connected by telephone line to the ADVANCE traffic information center (TIC) to support incident detection. Probe vehicles and anecdotal sources will provide data intermittently at locations determined by the location of the probe-equipped vehicle and the reporting source, respectively. Specifically, probe vehicles, driven by drivers volunteering to participate in the operational test for up to 2 years, will automatically report travel times by radio to the TIC each time they complete link traversals. During any time interval, there may be no probe report for many links in the network because of the relatively small number of probe vehicles. Similarly, anecdotal data will be available only when emergency personnel or motorists report a traffic incident on the link. These reports will be collected from a centralized emergency services dispatch center responsible for police, fire, and ambulance services in the test area; from the cellular telephone emergency reporting center; and from the state department of transportation's emergency patrol fleet.

To obtain the best determination of incident versus nonincident conditions under different conditions of data availability, the
authors adopt a hierarchical structure for the incident detection system in which data are first processed by specialized algorithms for each data source (fixed detector, probe vehicle, and anecdotal algorithms), then data fusion processes integrate all the available data to determine the overall likelihood that an incident has occurred at any particular location. This approach provides the flexibility to use the identification (ID) system when information on ADVANCE links is available from only one data source by employing the relevant algorithm in isolation; for a majority of incidents data from only one (or two) sources will be available because only a small fraction of links in the test area have fixed detectors and the small probe vehicle fleet can provide information from only a limited number of links during any short time interval. When more than one data source is available, this system will extract the most useful information from all available sources to make a more accurate determination of incident presence rather than select a single best source for each incident detection (3).

This paper presents a fully integrated incident detection system that is being implemented for the ADVANCE project and describes the individual components of the ID system, the input and output requirements of each component, and the relationships among the components. The development, refinement, and calibration of the individual algorithms used by the system are documented elsewhere (8–12).

OVERVIEW OF ADVANCE INCIDENT DETECTION SYSTEM

Figure 1 shows the relationships among the different components of the ADVANCE incident detection system; the components are as follows:

- **Fixed detector algorithm** uses the real-time and historical data provided by fixed detectors located on major arterials to classify conditions on the detectorized streets as incidents or nonincidents.
- **Probe vehicle algorithm** uses travel time reports by probe vehicle and historical travel times on these links to interpret traffic conditions as incident or nonincident.
- **Anecdotal algorithm** uses information provided by emergency personnel and other motorists on the network to detect incidents in real time.
- **Data fusion algorithms** combine the output from the fixed detector, probe vehicle, anecdotal algorithms, and the operator interface.

![Diagram of ADVANCE incident detection system](image-url)

**FIGURE 1** ADVANCE incident detection system.
• **Duration and impacts module** determines the expected duration of the incident and the impacts on the incident link travel times.

• **Operator interface** allows the TIC personnel to view the output from the data algorithms and to key in incident reports from other sources. The output from the duration and impacts module will be available to the operator for review.

The automatic data algorithms (the fixed detector and probe vehicle algorithms) classify conditions on links for which current data are available at the end of a prespecified period (e.g., 5 min); “current data” refers to probe reports or detector output received by the TIC during the current period. If both probe vehicle and fixed detector data are available for a link, the output from both algorithms is combined in the data fusion (Step 1) module; the fusion process is bypassed if data are available from only one source, and the output from the corresponding algorithm is used alone. When all the links with current data have been processed, the classification results are saved in the automatic detected incident file containing an unique tag for each link and a flag indicating incident presence or absence.

A new algorithm for incident detection using fixed detector data was developed instead of adapting existing pattern recognition (13–15) or time series methods (1,16). The primary reasons for this were (a) the existing algorithms were developed primarily for freeway environments whereas the ADVANCE network consists mostly of arterial streets (it is important to recognize that traffic flow characteristics on arterials are very different from freeways because of the presence of traffic signals, parking, and such, which results in greater variability in traffic flow measures such as occupancy and travel time), and thus the freeway algorithms are not readily transferable to arterials; (b) they require loop detector data for short intervals of time (approximately 30 sec to 2 min), and it will not be possible to get these data for less than 5- or 15-min intervals for ADVANCE; and (c) many of the existing algorithms use data from adjacent pairs of detectors, which generally are not available on ADVANCE arterials in the test area. Further, no well-established methods are available for using probe vehicle and anecdotal data and data fusion processes for incident detection.

The fixed detector, probe vehicle, and data fusion algorithms were estimated using simulated data because no actual field data and corresponding incident confirmations were yet available. When it is deployed, the ADVANCE operational test will generate field data that will be used to recalibrate the algorithms. Each of these models was estimated using discriminant analysis (17), which produced a function of the traffic flow parameters, the value of which is used to identify incident or nonincident conditions. Discriminant analysis uses prespecified values of prior probabilities of incidents (priors) to control the classification output and develop more realistic models. Incident conditions will exist during only a small fraction of time periods on any given link; this low probability of incidents in the real world is taken into account by adopting incident priors of 0.0001 (i.e., in the absence of any other information, any particular report has a probability of 0.0001 of being an incident report). This ensures that the number of false alarms (incident reports generated in nonincident conditions) will be small.

The anecdotal algorithm uses data from anecdotal sources (e.g., computerized emergency dispatch systems) and produces a link-specific output that is expected to be more detailed than the automatic classification output. The operator will be able to key in incident reports from other sources using a menu-based interface. The link-specific anecdotal algorithm output and operator inputs will be stored in the TIC operator/anecdotal report file containing a link identification tag, incident indicator, incident type, and a vector of variables representing the incident intensity.

At the end of every period, the incident information from TIC operator/anecdotal report file and the automatic detected incident files will be matched by link; the data fusion (Step 2) will be performed for links that have classification information available from both the files. Fusion will be bypassed for links having incident information from only one of the files.

The duration and impacts module uses the output from the data fusion (Step 2) process to estimate the duration and impacts of the incidents based on the incident type and intensity. For links with missing values for the incident type and intensity information, default estimates of duration and impacts will be used. For many links and periods, no data will be available; for these links the default conditions will be assumed to be nonincident. Some incidents will last for more than one period; when such incidents are detected in consecutive periods, the incident duration will be updated in the duration and impacts module for each period that the incident is detected.

The TIC operator will be able to review the final output before it is passed to other ADVANCE processes. The operator will be given a limited time window for review to avoid a backlog of detected incidents not reported to participating drivers. Eventually, reports of incidents and their effects on travel time will be transmitted by radio to route planning computers in participating vehicles.

### ALGORITHM COMPONENTS

#### Fixed Detector Algorithm

The fixed detector algorithm compares current and historical volume and occupancy data from fixed detectors at the end of every period (8); Figure 2 shows the flow diagram for the algorithm. The historical (nonincident) volume and occupancy data are aggregated over a fixed time interval for each detector location by day type (weekday, weekend, etc.) and time of day.

The algorithm uses current and corresponding historical volume and occupancy data to compute two variables:

- **Occupancy deviation** = \( occupancy_{\text{observed}} - occupancy_{\text{historical}} \)
- **Volume/occupancy deviation** = \((volume/occupancy)_{\text{observed}} - (volume/occupancy)_{\text{historical}}\)

A discriminant score is then computed using Equation 1, which determines incident presence in the proximity of each detector:

\[
\text{Discriminant score} = -14.880 + 0.0192 \times \text{occupancy deviation} - 4.088 \times \text{volume/occupancy deviation} \tag{1}
\]

If the discriminant score is greater than 0, an incident is flagged for the link associated with the detector for the corresponding time period; if the discriminant score is less than 0, normal conditions are assumed. Incident classification, discriminant scores, and values of the deviation variables are provided to the data fusion module for every link that is processed by the algorithm.

Both the fixed detector and probe vehicle algorithms identify incidents that occur either on the link from which the detector output or probe report is obtained or on the upstream portion of the adjacent downstream link. However, because of the typical traffic flow impacts of arterial street incidents, the incidents detected are
Read current volume and occupancy for the detector.

Historic volume and occupancy database

Compute - occupancy deviation,
- (volume / occupancy) deviation.

Compute discriminant score and classify link as incident or non-incident.

Output to data fusion (step 1) module - classification results,
- discriminant scores,
- occupancy deviation, and
- (volume / occupancy) deviation.

**FIGURE 2** Fixed detector algorithm.

The current versions of both fixed detector and probe vehicle algorithms do not differentiate between these cases; they will attribute all such incidents to the reporting link. This will be the correct assignment in most cases, and it will tend to have the correct impact on drivers in other cases, diverting them from the incident links.

**Probe Vehicle Algorithm**

The probe vehicle algorithm requires current and historical (nonincident) probe travel time data; it operates in two stages as shown in Figures 3 and 4 (9). In the first stage (Figure 3), average link travel times are computed by aggregating individual probe reports at the end of the period. Aggregation of probe reports results in a more accurate representation of the traffic conditions by averaging out aberrant nonincident probe reports. The aggregation procedure will depend on the number of reports received in the current time period. If only one report is available in the current period, probe reports from the previous period are averaged with the current reports. If more than one report is available, the average travel time is computed using all the probe reports received during the current period. If no report is received in the current period, or if in two consecutive periods only one report is received from a link, that link is not processed because of the unreliability of individual reports.

The second stage of the probe algorithm is the application of the different models to classify conditions on the links as incident or normal; Figure 4 shows this procedure. Travel time ratio and speed ratio are computed using the observed travel time on the links and the corresponding historical travel times stored in a data base. Incident presence is determined by computing the discriminant score; the effective cut-off travel time ratios for declaring an incident is dependent on the number of reports received during the detection interval to recognize the increased reliability of the average link travel time with increasing number of probe reports (9). The cut-off points are given in Table 1. The output from the probe vehicle algorithm consists of the classification results, discriminant scores, travel time ratios, and speed ratios for every link processed by the algorithm.

**Data Fusion—Step 1**

The data fusion algorithm reviews each link in the network once in every period and determines incident presence or absence for each link for which reports from both probe vehicle and fixed detector algorithms are available (11). The classification result for those links will be stored in the automatic detected incident file. If a report is available from only one of the algorithms, it will be used directly as the output of the automatic data fusion. Once all the links are evaluated for that time period, the automatic detected incident file will be combined with the TIC operator/anecdotal report file.

Two approaches, discriminant analysis and artificial neural networks, were tested for this fusion task. Comparison of incident detection results with discriminant analysis and neural network data fusion models showed that the model estimated using discriminant analysis had better detection rates for less extreme priors, whereas the neural network model performed better for more extreme priors (11). It is useful to evaluate both of these models with real data before selecting one over the other. For the current version of the ADVANCE incident detection system, the discriminant analysis model will be adopted because of its simpler structure. However, all the data required by the neural network model will be saved for offline comparative testing of the two models.

The best data fusion model using the discriminant analysis approach uses occupancy deviation and volume/occupancy deviation from fixed detector data, and travel time ratio and speed ratio from probe vehicle data (Figure 5) for computation of the discriminant scores as follows:
FIGURE 3 Computation of average travel time using probe reports.

\[
\text{Discriminant score} = 3.005 - 0.255 \times \text{occupancy deviation} - 4.523 \\
\times \text{(volume/occupancy) deviation} - 24.573 \times \\
\text{speed ratio} + 1.834 \times \text{travel time ratio} \quad (2)
\]

An incident is flagged if the discriminant score is greater than 0. The performance of this model was substantially better than either the fixed detector or probe vehicle algorithm when applied to the same data set (11), showing that overall detection ability can be improved by using detector and probe vehicle data together (when available).

**Anecdotal Information Algorithm**

The anecdotal incident detection algorithm uses a qualitative description of incidents reported by field observers, both trained and untrained, to detect incidents in real time (10). The two primary sources of anecdotal data will be:

1. The Northwest Central Dispatch System (NWCD), a computer-aided emergency service dispatch agency serving six communities in the center of the ADVANCE test area; and
2. The *999 center, which receives toll-free calls from cellular telephone users voluntarily reporting roadway incidents and other problems.

Reports from the *999 center, operated for the Illinois State Toll Highway Authority, originate primarily from lay citizens; they will be in the form of qualitative descriptions of events and nominal location references. A simple manual data connection to the *999 center is planned for late in the ADVANCE operational test. The rest of this section focuses on the use of NWCD anecdotal inputs and inputs from other sources through the TIC operator.

**TABLE 1 Cut-Off Point for Declaring Incidents Used by Probe Vehicle Algorithm**

<table>
<thead>
<tr>
<th>Number of Reports</th>
<th>Travel Time Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>3.45</td>
</tr>
<tr>
<td>3, 4</td>
<td>2.80</td>
</tr>
<tr>
<td>5, 6, 7</td>
<td>2.60</td>
</tr>
<tr>
<td>8, ..., 15</td>
<td>2.40</td>
</tr>
<tr>
<td>15, 16, ...</td>
<td>1.45</td>
</tr>
</tbody>
</table>
Reports from NWCD will be captured from the computer system in the dispatch center, where all incoming calls, emergency services dispatches, and other communications are entered into a data base. The ADVANCE anecdotal algorithm will use descriptions of emergency vehicle dispatches to roadway incidents, explicit or implicit incident confirmation and clearance reports provided by on-scene emergency service personnel, incident type descriptions by standard codes (accident with property damage, accident with personal injuries, hazardous material spills, motorist assist, etc.), and (in some cases) incident intensity as reflected by number of service units on the scene or other qualitative descriptions. NWCD receives incident location information from callers in various forms, including street addresses, intersections, and landmark names; this information is geocoded to street addresses within the NWCD computer system.

Figure 6 shows the main components of the anecdotal ID algorithm. Initially, the only source of anecdotal data will be NWCD. Data from NWCD will be preprocessed at NWCD before transmission to the ADVANCE TIC to extract only roadway incidents, and only those descriptor variables of interest to ADVANCE. The NWCD preprocessor will

- Distinguish new incidents from update reports on incidents already identified,
• Separate incidents of interest to ADVANCE (i.e., roadway blocking incidents) from others, and
• Maintain a list of active incidents; format messages to be sent by telephone to the TIC.

Data from other anecdotal sources will have a separate preprocessor and translation module that will be similar to the TIC/NWCD preprocessor and translation module. Further, a procedure for matching reports from different sources for the same incident will be required. These will be incorporated in the anecdotal ID system when data from other sources become available.

The anecdotal reports will be received by the TIC/NWCD preprocessor and translation module that assigns the information from the report into several fields (i.e., information source, incident history number, incident location, incident type, unit update report, and incident intensity). The incident history number is defined by NWCD and used as a basis for maintaining unduplicated files of data for each active incident. Location will be in the form of street addresses. Incident type codes will be a short list of types of interest to ADVANCE. Mobile unit update reports will indicate emergency service unit radio call numbers and a status code (enroute, on-scene, clear). This information will be used to confirm incidents and determine clearance:

• Incidents will be confirmed 3 min after the first emergency responder arrives on the scene unless that unit reports clear,
• Incidents will be reported clear 5 min after the last emergency unit leaves the scene.

These criteria were established because responding units usually do not formally confirm incident presence or signal final clearance.

Next, each incident will be assigned to a link by converting the address to a segment ID using a file with coordinates of all links and intersections in the test area and finding the street name. Once the street name is found, the address will be used to identify the segment. The link-specific anecdotal reports (link identification, incident indicator, incident type, and incident intensity) on confirmed and cleared incidents will then be saved in the TIC operator and anecdotal report file.

Intensity data will be available for use in enhancing estimates of incident duration and traffic impacts; initial intensity data will be a simple count of emergency units on the scene. It may prove useful and feasible to parse free-form text reports from NWCD, and particularly from *999 (which will be rich in qualitative data from untrained observers) for use in more complex impact estimation efforts. Development of these capabilities must await the availability of field data from the ADVANCE implementation.

FIGURE 6 Different components of anecdotal incident detection algorithm.
The TIC operator will also be able to enter anecdotal reports of incidents from telephone calls and other sources into the TIC computer system for use in the incident detection system. Inputs will match the data types used by the anecdotal algorithm: information source, location by address or intersection, incident type, and status (confirmed or clear). When the operator receives a report, he or she will open an incident report window on the TIC terminal. This window will display an input form, allowing data items to be keyed in or selected from short menus with a pointing device.

The operator inputs would bypass the anecdotal algorithm and go directly to the TIC operator/anecdotal report file. In this way, the operator will be able to clear an incident already detected or define a new incident not previously recognized by the ID system. The operator interface will also allow the operator to display a list of all active incidents by location, type, and start time.

Data Fusion—Step 2

The data fusion (Step 2) process combines the output from the automatic data fusion algorithm and TIC operator/anecdotal algorithm; Figure 7 shows the process. This data fusion is accomplished using a rule-based approach. For the current version of the ADVANCE ID system, the output from the anecdotal algorithm and TIC operator will override the output from the automatic algorithms (i.e., fixed detector and probe vehicle algorithms). Initial anecdotal algorithm output will come only from NWCD, a highly reliable source originating with emergency services professionals. When other anecdotal sources are brought on-line, a fusion procedure that blends algorithm inputs along the lines of Step 1 fusion approach described earlier will be developed. The output from the data fusion (Step 2) process will be saved in the final classification file that will be passed on to the duration and impacts module.

Duration and Impacts

The impact of an incident is its effect on the travel time. The duration of an incident is the length of time during which that travel time impact occurs. The clearance time of an incident is the time from detection until the blocking event is removed from the roadway. The duration may last beyond the clearance time if a major queue must be dissipated. In other cases, the vehicle(s) involved in the incident can be moved off the road and the traffic conditions returned to normal, but the anecdotal algorithm will not declare the incident as cleared until the last emergency unit has left the site; in such cases the duration is less than the clearance time. The duration is more important to ADVANCE than the clearance time, but only the latter is typically available from anecdotal sources and may be used to estimate the former. Since current ID algorithms cannot distinguish between the two, the authors adopt a conservative perspective and set the incident duration to be 10 percent more than the average clearance time.

The duration and impacts module will receive the final classification file from the data fusion (Step 2) process. The incident type and intensity variables (number of police and fire units on scene), when available, will be used to compute the expected duration and impact of the incident. Figure 8 shows the procedure for determining incident duration and impact for different categories of incidents (12); some related incident types have been grouped together because they have similar characteristics. These incident types are a subset of incident categories used by NWCD; when data for other types of incidents become available, they will be incorporated in the algorithm.

The duration of incidents can be expressed in terms of minutes or algorithm cycle (5-min periods). Figure 8 reports duration in minutes; this can be converted to algorithm cycle units by dividing duration by cycle length (i.e., 5 min) and increasing fractional time period values to the next highest integer. Since the algorithms oper-
ate at the end of each period, it will be simpler to use periods as the unit of duration, especially for the duration updating process. Finally, when a specific clearance message is received from NWCD or other valid sources, the incident will be terminated.

Initially the incident impact, formulated as a travel time ratio, will be based on default values derived from the simulation studies (12). These values are based on the distribution of travel time ratios for identified incidents. The median of the travel time distribution is used for motorist assist and traffic stops, and the 75th percentile of the travel time distribution, for all other incidents except severe incidents and in cases when the incident type is unknown. For severe incidents (such as accidents with entrapments, accidents involving hazardous materials, and fire-related incidents), 1.5 times the values used for all other incidents (excluding motorist assist and traffic stops) are used. Default values of duration will be used for incident links for which the incident type is not known, and the links with no data from any of the sources will be assumed to be nonincident.

For incidents that last longer than one period, the duration will be updated every period by the time elapsed since the incident was first detected. If the incident lasts longer than the expected duration, then for every additional period that it is detected, the duration will be set to one more period until the incident is cleared or a nonincident message is generated by the algorithms that identified the incident. The procedural details for duration updating are provided elsewhere (18).
Operator Review
The operator review module will allow the operator to review the output of the ID system before it is passed to other ADVANCE processes. Through this option the operator can override the algorithm recommendations on the presence of incidents on the links. Initially, all the ID output will be confirmed manually by the operator. If the operator does not confirm the output or does not take any action, the links will be assigned nonincident status. In all cases (i.e., when operator takes an action or does not respond to ID output), the ID output will be saved with the corresponding operator response; the operator responses will be evaluated to make this feature more effective.

CONCLUSIONS
This paper describes the incident detection system being implemented for the ADVANCE project. The ID system will use information from three distinct data sources: fixed detectors, probe vehicles, and anecdotal sources processed through specialized algorithms. The output from these algorithms will be integrated using a two-stage data fusion process to determine the overall likelihood that an incident has occurred at any particular location. On the basis of type of incident, the expected duration and travel time impacts of the incidents will be determined.

In contrast with other incident detection methods, this approach is designed to integrate information from multiple data sources. It is based on the concept that effective integration will result in an enhanced detection capability, making use of the special characteristics of each data source. Evaluation with field data during the ADVANCE operational test will provide an opportunity to verify this design concept.

ACKNOWLEDGMENTS
The information in this paper was partially funded by and developed for the ADVANCE project, a joint ITS undertaking by FHWA, the Illinois Department of Transportation, the University of Illinois at Chicago, and Northwestern University operating under the auspices of the Illinois Universities Transportation Research Consortium, Motorola, Inc., and other participating organizations.

REFERENCES

The contents of this paper reflect the views of the authors, who are responsible for the information and accuracy of the data presented. It does not constitute a standard, specification, or regulation. It does not necessarily reflect the official policy of any of the ADVANCE parties, nor does any mention of manufacturers or products constitute an endorsement of the ADVANCE parties.