

# UC Berkeley

## Working Papers

### Title

A New Methodology for Evaluating Incident Detection Algorithms

### Permalink

<https://escholarship.org/uc/item/35j0g3tm>

### Authors

Petty, Karl  
Bickel, Peter J.  
Kwon, Jaimyoung  
[et al.](#)

### Publication Date

2000-08-01

**This paper uses Postscript Type 3 fonts.  
Although reading it on the screen is difficult  
it will print out just fine.**

CALIFORNIA PATH PROGRAM  
INSTITUTE OF TRANSPORTATION STUDIES  
UNIVERSITY OF CALIFORNIA, BERKELEY

## **A New Methodology for Evaluating Incident Detection Algorithms**

**Karl Petty, Peter J. Bickel, Jaimyoung Kwon,  
Michael Ostland, John Rice**  
*University of California, Berkeley*

**California PATH Working Paper  
UCB-ITS-PWP-2000-11**

This work was performed as part of the California PATH Program of the University of California, in cooperation with the State of California Business, Transportation, and Housing Agency, Department of Transportation; and the United States Department Transportation, Federal Highway Administration.

The contents of this report reflect the views of the authors who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the State of California. This report does not constitute a standard, specification, or regulation.

Report for MOU 353

August 2000

ISSN 1055-1417

# A New Methodology for Evaluating Incident Detection Algorithms\*

Karl F. Petty, Peter J. Bickel, Jaimyoung Kwon, Michael Ostland, John Rice  
Department of Statistics  
University of California, Berkeley  
{pettyk,bickel,kwon,ostland,rice}@stat.berkeley.edu

June 2, 1999

## Abstract

We present a novel, off-line approach for evaluating incident detection algorithms. Previous evaluations have focused on determining the detection rate versus false alarm rate curve — a process which we argue is inherently fraught with difficulties. Instead, we propose a cost–benefit analysis where cost mimics the real costs of implementing the algorithm and benefit is in terms of reduction in congestion. We argue that these quantities are of more practical interest than the traditional rates. Moreover, these costs, estimated on training data, can be used both as a mechanism to fine–tune a single algorithm as well as a meaningful quantity for direct comparisons between different types of incident detection algorithms. We demonstrate our approach with a detailed example.

**Keywords:** Incident detection

## Acknowledgements

The authors gratefully acknowledge the support of the National Institute of Statistical Sciences (NISS) and the California PATH Program of the University of California.

## 1 Introduction

The cost of delay on freeways caused by non-recurring incidents is significant. Some estimate that the cost will be \$35 billion/year by the year 2005 [J. A. Lindley, 1986]. To reduce the impact of an incident, a traffic management center (TMC) needs to quickly detect and remove it from the freeway. A large literature of automatic incident detection algorithms (AID’s) has emerged to address the problem of quick detection. AID’s which can operate on data collected by ubiquitous inductance loop detectors include simple filtering [Stephanedes and Chassiakos, 1993], pattern recognition [Payne et al., 1976, Payne and Tignor, 1978, Tsai and Case, 1979], catastrophe theory [Persaud and Hall, 1989], and more recently neural networks [Ritchie and Cheu, 1993, Hsiao et al., 1994] and genetic algorithms [Abdulhai and Ritchie, 1997].

---

\*Funding for this research was provided in part by NSF-DMS 9313013 to NISS and California PATH MOU-353.

Faced with such a large literature, practitioners have a sizeable task in determining which — if any — AID to implement. Adding to the difficulty is the fact that most algorithms are presented alone in the literature and in such a way that meaningful, literature-based comparisons between AID’s are nearly impossible. Further, most AID’s depend on a set of parameters that must be set (“tuned”) by the practitioner. Unfortunately, setting parameters can be very difficult in practice, and performance can be very sensitive to these settings. Since poor performance usually translates into an unacceptably high rate of false alarms, many practitioners find AID’s too problematic for implementation in large urban environments [Balke, 1993]. We present a methodology that will enable practitioners to both tune algorithms and make meaningful, direct comparisons between AID’s.

Standard evaluation of AID’s is in terms of detection rate (DR) versus false alarm rate (FAR) curves — where the higher the curve, the better the algorithm. We feel that this approach has several major difficulties, of which we now list but two. First, to discuss these rates requires an *a priori* judgement by the researcher as to what constitutes an incident and what is detectable. This subjectiveness can lead to separate studies classifying different events as incidents on the same data set. Consequently, fair comparisons between different algorithms are nearly impossible from the literature alone.

A second problem is that DR-FAR curves treat all incidents with equal importance. That is, failing to detect a low-impact breakdown on the shoulder contributes equally towards the false alarm rate as missing a major accident that causes hours of congestion. From a practical point of view, this is a fundamental flaw. Similarly, all detections, regardless of the delay in the time to detection, count equally in the detection rate. Practically speaking, this too is nonsense. Since average time to detection is usually reported separately from the DR-FAR curve (if it is reported at all), the trade-off between the DR-FAR curve and the average time to detection is seldom clear.

Our approach to evaluating AID’s attempts to solve these problems by abandoning DR-FAR curves in favor of a more natural cost function. Namely, to evaluate a specific AID we estimate the cost that would be incurred from delay experienced if that AID were actually implemented. To this cost we add the cost of implementing the AID (dispatching tow trucks, etc.). Both of these costs are based on some assumptions and estimated from training data. By holding these assumptions fixed, however, the costs are directly comparable across different algorithms and can be used in tuning an individual algorithm. Moreover, by dealing with congestion/implementation costs instead of DR-FAR curves, the severity of the incidents and time to response are automatically factored into the analysis.

Reducing congestion is by no means the sole purpose of incident detection. Representing an AID by a single number based on delay and the cost of responding can not possibly account for all such purposes. Some important considerations may not be quantifiable at all. However, congestion and implementation costs are important criteria, which can be evaluated to inform decision making. A meaningful univariate criterion permits systematic exploration of tuning parameters within algorithms and comparisons between algorithms — important, practical goals that seem extremely difficult, if not impossible, to achieve from the traditional methods of evaluation.

Finally, our approach allows flexible formulations of cost; for example, the cost function could reflect increasing benefit for responding promptly to injury accidents. Our method requires only that the cost can be estimated from training data. Thus, the practitioner wishing to use his or her own measures of cost can do so based on our generic approach.

In the next section we present our methodology for a generic cost function. This is done intentionally as we feel that practitioners are best qualified for making the assumptions linking congestion and utility which constitute a loss function. However, for concreteness we give an example in Section 3 in which it is necessary to make some specific assumptions. Section 4 contains some

general remarks about cost functions and possible extensions.

## 2 Methodology

In this section we describe our approach in some generality. At this level the ideas are quite simple. In Section 3 a specific example is presented which illuminates the details behind the following general discussion.

We adopt an AID cost structure with two parts corresponding to delay and implementation. As an AID is tuned to be more sensitive, the more it reduces delay through the quick detection of incidents. On the other hand, greater sensitivity also translates into increased implementation cost, since the AID is calling for more interventions. The goal is then to find a fully-tuned AID that strikes a balance between these conflicting aims.

Let  $S$  be an AID which depends on a vector of parameters,  $\theta_s$ . We seek a pair  $(S, \theta_s)$  with low total cost for a given freeway system. To estimate total cost, we use training data,  $D$ , which is collected during a period in which AID interventions are not being made.  $D$  must contain two things: a list of the locations and durations of all incidents<sup>1</sup> (usually from probe vehicles) and the data sufficient to implement the AID (usually inductance loop detector data). However,  $D$  could also incorporate such data as highway patrol logs, cellular phone data, video data, and travel times from vehicles with automatic toll collection tags. Such additional data would increase the precision of the cost function estimates.

Let the functions  $f$  and  $g$  denote the delay and implementation cost functions, respectively. The evaluation of  $f$  requires estimation of the delay if a certain  $(S, \theta_s)$  were used. This is the subject of Section 3.1. Also present in  $f$  is a conversion from delay (in vehicle-hours) to cost (in dollars).  $g$  takes actions prescribed by  $(S, \theta_s)$  (eg. “dispatch tow truck to location  $x$  at time  $t$ ”) and assigns a cost to these actions. These functions can be anywhere from very simple (our example in Section 3) to quite complex (see Section 4).

We fit parameters for a particular  $S$  by finding

$$\theta_s^* = \arg \min_{\theta_s} f(S, \theta_s, D) + g(S, \theta_s, D). \tag{1}$$

If the dimension of the parameter space is not too large, an exhaustive search over a grid of values can be used to find the minimum cost  $\theta_s^*$ . Otherwise, gradient search methods, like the Downhill Simplex [Press et al., 1992], are applied. In the latter case the usual caveats of local minima apply.

Define  $h(S, D) = f(S, \theta_s^*, D) + g(S, \theta_s^*, D)$ , the lowest possible cost of  $S$  for data  $D$ . To compare AID’s  $S_1$  and  $S_2$  we can then compare  $h(S_1, D)$  to  $h(S_2, D)$ . To avoid over-fitting to the training data, which would make an AID with a high-dimensional  $\theta_s$  appear better than it otherwise would, it is preferable to find  $\theta_{s_i}^*$  with one set of training data and  $h(S_i, D)$  with another.

An important AID for comparison is the “do nothing” algorithm (ie. never dispatching a tow truck), which we henceforth denote  $S_0$ . Of course,  $S_0$  depends on no parameters, and  $g(S_0, D) = 0$  for all  $D$ . A minimal requirement for the viability of  $S$  is  $h(S, D) < h(S_0, D)$ . Sensitivity to the underlying cost assumptions can be addressed in part by how persistent the relationship between  $h(S, D)$  and  $h(S_0, D)$  remains as the assumptions are altered.

---

<sup>1</sup>Our list differs importantly from that of the DR-FAR approach in that we do not pre-screen incidents for detectability. Any abnormal event can be added to the list as a possible incident. Of course, the definition of “abnormal” is subjective, but the same subjective list is used for each AID under evaluation. Moreover, if a particular event causes no delay, then an AID will not be penalized for failing to detect it.

### 3 An Example

In this section we demonstrate our method using data from the Freeway Service Patrol Evaluation Project [Skabardonis et al., 1995, Skabardonis et al., 1997, Skabardonis et al., 1996, Petty et al., 1996] collected from a seven mile section of freeway on I-880 in Hayward, California. This section of freeway was instrumented with type 170 loop controllers spaced approximately 1/3 of a mile apart. Each loop controller monitored eight to ten mainline, double-loop detectors. Flow, speed, and occupancy data were aggregated over 30-second intervals for our analysis. The project also used rotating probe vehicles to collect incident data. The interested reader can find more details in the cited references.

We compare three AID's: The "do-nothing" AID,  $S_0$ , the Basic California AID [Payne et al., 1976, Payne and Tignor, 1978],  $S_1$ , and the Minnesota Algorithm [Stephanedes and Chassiakos, 1993],  $S_2$ . Brief descriptions of  $S_1$  and  $S_2$  are given in the Appendix.

Evaluating equation (1) requires three tasks for each algorithm and our training data: (i) estimation of the delay that would have occurred using that algorithm and a given set of parameters (ii) conversion of this delay to cost, and (iii) calculation of implementation cost. The next three sub-sections address these points in order. The final sub-sections present results for our I-880 analysis.

#### 3.1 Estimating Incident Delay

We now discuss the estimation of the counter-factual incident delay that would have been experienced under AID  $(S, \theta_s)$ . The key assumptions will provide a link between a reduction in incident duration and a reduction in delay. We begin by determining the actual delay associated with each incident in our training data. This requires a list of incidents with corresponding time-space extent as well as flow and speed data.<sup>2</sup> We first determine the delay for a segment (link) of freeway based on the difference in travel times between normal and incident conditions [Epps et al., 1994]. The definition of delay is simply the extra time it takes to cross a freeway segment multiplied by the number of vehicles on the segment:

$$d_i = f_i \left( \frac{l_i}{V_i} - \frac{l_i}{V_T} \right), \quad (2)$$

where  $i$  indexes link,  $d_i$  is the delay in vehicle-hours,  $l_i$  is the length of the link in miles,  $f_i$  is the link flow,  $V_i$  is the current link speed (mph), and  $V_T$  is the typical link speed under incident-free conditions for that time of day. Note that it is possible to determine this delay for each freeway link at each time period from only loop detector data. Of course, if more precise data are available, one should incorporate them to get better estimates in (2). Next, we compile this link-by-link delay data into contour plots of delay. This allows us to associate buildups of delay with specific incidents [Petty, 1997]. From this we can determine the actual delay for each incident. Denote the delay from the  $k$ th incident by  $d_0^k$ .

We now turn to estimating the delay of incident  $k$  if  $(S, \theta_s)$  were implemented. We denote this hypothetical delay by  $d_s^k$ . For  $S_0$ , we already have  $d_{s_0}^k = d_0^k$  since  $D$  was collected under  $S_0$ . For other AID's, we first estimate the reduction in incident duration, and then we convert this reduction in duration into a reduction in delay. If the incident is not detected, then it is again natural to take  $d_s^k = d_0^k$ .

To express delay as a function of incident duration we make the assumption that one can model the incidents with a standard queuing model. Alternative assumptions are briefly discussed in Section (4). Figure 1 is a standard cumulative delay plot for a general incident [May, 1990] which

---

<sup>2</sup>For our analysis, an object (broken down vehicle, debris, etc.) appears on the incident list if it is witnessed by two or more probe vehicles.

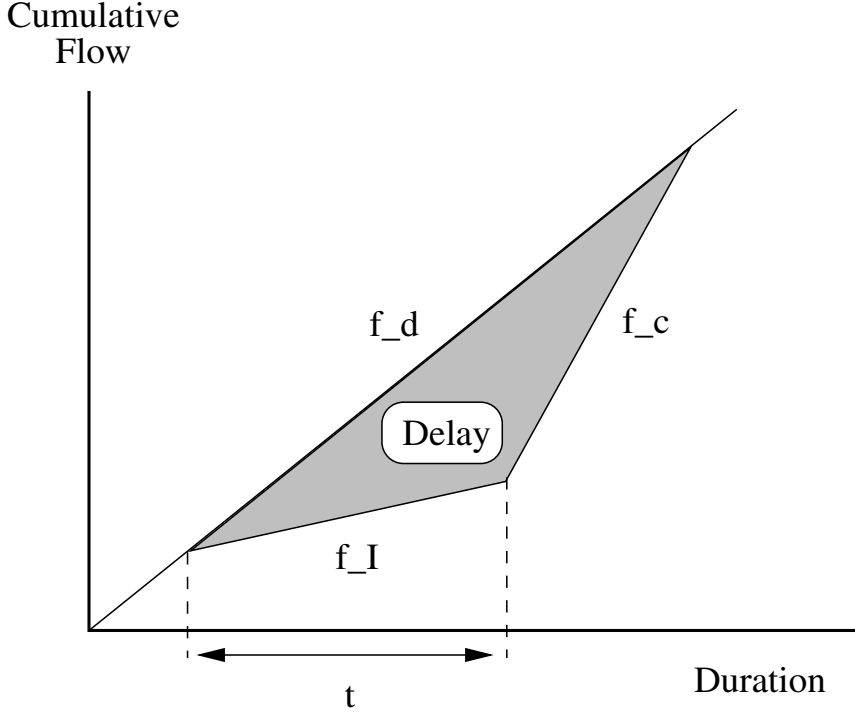


Figure 1: Standard calculation of incident delay using a cumulative flow diagram (which is derived from the standard queuing model). The line labeled  $f_I$  is the capacity of the freeway during the incident,  $f_c$  is the discharge capacity of the freeway,  $f_d$  is the demand on the freeway, and  $t$  is the duration of the incident. The shaded area is delay.

corresponds to the standard queuing model. By this model  $d_0^k$  is the shaded area inside the triangle, and is given by:

$$d_0^k = \frac{t^2 (f_d - f_I)(f_c - f_I)}{2 (f_c - f_d)}. \quad (3)$$

In equation (3) the actual incident duration,  $t$ , is known from our training data, and  $d_0^k$  is known from our previous calculations. The constants  $f_d$ ,  $f_c$ , and  $f_I$  are the demand, discharge capacity and capacity during the incident, respectively. Suppose our AID calls for an intervention that reduces the duration to  $t' \leq t$ . Then, by (3)

$$d_s^k = d_0^k \left( \frac{t'}{t} \right)^2 \quad (4)$$

is the new delay. Thus, no assumptions on the values of  $f_d$ ,  $f_c$ , or  $f_I$  are necessary.

A final step to evaluate (4) is to express the new incident duration,  $t'$ , as a function of the time of intervention — denoted by  $u$  — following a successful detection. Two components are needed: the time,  $T$ , required for the tow truck to reach the incident, and the time  $\Delta$  required to clear the incident. Then,

$$t'_0 = \min(t_0, u + T + \Delta) \quad (5)$$

In this example we make the simple assumption that  $T = \Delta = 10$  minutes for all incidents. See Section 4 for other possibilities.

We note that researchers without probe vehicle data may benefit from simulated loop detector and incident data. To do so, one generates incidents of various types and calculates the delay for



each incident by plugging baseline values for  $f_I$ ,  $f_d$ , and  $f_c$  into (3). The incident capacity,  $f_I$ , should depend on the incident type. Using highway patrol incident data is another alternative to deploying probe vehicles.

### 3.2 Converting Delay into Cost

Having estimated the  $k$ th incident’s hypothetical delay under  $(S, \theta_s)$ , we now convert this delay into cost. One simple assumption is that each vehicle-hour of delay costs  $K_D$  dollars. Thus,

$$f(S, \theta_s, D) = K_D \sum_k d_s^k \tag{6}$$

where the sum is over incidents. Section 4 discusses other possibilities.

### 3.3 Cost of implementation

Next, we assign an implementation cost to  $(S, \theta_s)$  based on our training data. A significant portion of this cost is the cost of responding to the incidents detected by the AID. A reasonable simplification is to charge  $K_T$  dollars for every tow truck dispatched by the AID. Let  $D(x, t) = 1$  if a tow truck is dispatched to link  $x$  at time interval  $t$ , and  $D(x, t) = 0$  otherwise. Then

$$g(S, \theta_s, D) = K_T \sum_{x,t} D(x, t) \tag{7}$$

Again, generalizations are discussed in Section 4.

### 3.4 Dispatching Tow Trucks

A final detail for evaluating (7) is the translation of the output of  $(S, \theta_s)$  into tow truck dispatches,  $D(x, t)$ . Recall,  $(S, \theta_s)$  produces a sequence of points (or “calls”) in space–time for which an incident is believed to exist. Naively taking  $D(x, t) = 1$  at every call can result in sending several (sometimes dozens) of tow trucks to the same point in space at consecutive time intervals. Therefore, it is necessary to incorporate some common sense into the final dispatching decision.

We refer to this map from the calls of an AID to tow truck dispatches as the *dispatcher*. After some initial trial–and–error, we settled upon a dispatcher that follows the following rule: If  $D(x_0, t_0) = 1$ , then  $D(x, t) = 0$  for all  $(x, t) \in [x_0 - b_1, x_0 + b_2] \times (t_0, t_0 + b_3]$ . This space–time black–out region discourages an AID from sending out multiple tow trucks to the same incident. In principal the  $b$ ’s could be added as parameters to the vector  $\theta_s$ , embedding the dispatcher inside  $(S, \theta_s)$  completely. However, to reduce the dimension of the search in (1), we fix  $b_1 = b_2 = 1$  link and  $b_3 = 10$  minutes. We attach this same proxy for a human dispatcher to all algorithms.

### 3.5 Tuning the California Algorithm

In this section, for expository purposes, we tune the California algorithm,  $S_1$ , on data from a single day, February 18, 1993. This turns out to be a good day for both detection and expository purposes since we can clearly see incidents taking place that have well defined congestion. In Section 3.7 we make a more appropriate analysis based on a collection of days.

Figure 2 is a contour plot of delay (cf. equation (2)) for this day. The boxes on the contour plot represent the location and duration of the incidents as witnessed by the probe vehicles. The number in the lower left-hand corner of each box is the incident number in our database. So, for example, incident #117 occurred near loop detector #4 around 8:50am and lasted until around 9:15am. The

Southbound AM contour: 2/18/93

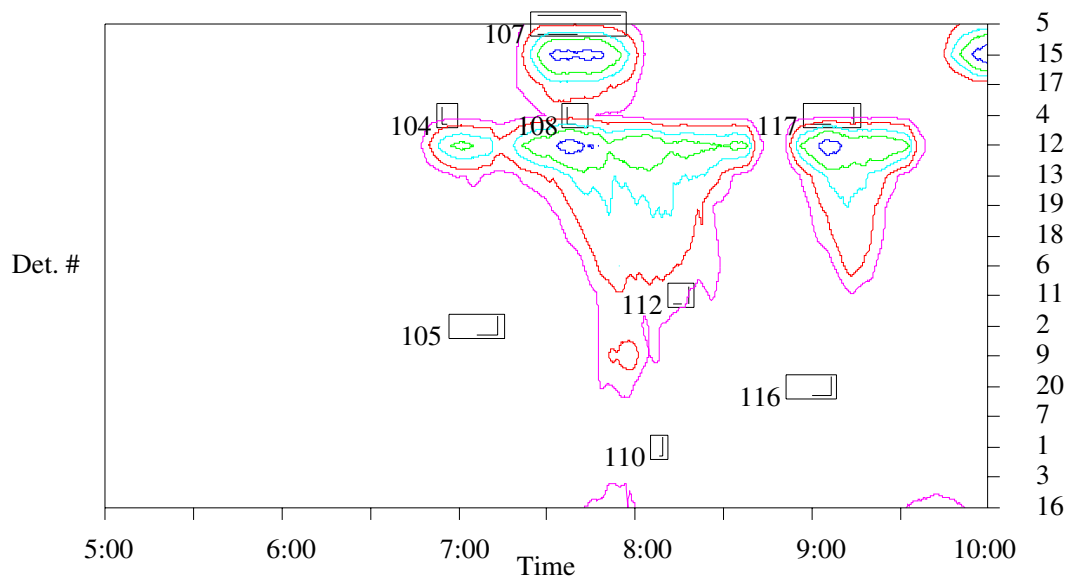


Figure 2: Contour plot of delay with incidents. Traffic flows upwards in the picture (the vertical axis is the loop detector number). Incidents are shown by boxes. Within a box, the position of the bar(s) code(s) the type and location of that incident: bottom-left means accident, bottom-right means breakdown, top means in-lane, and left (right) means it took place on the left- (right-) hand side of the freeway. For example, incident #116 was a breakdown on the right-hand side of the freeway.

congestion from this incident finally cleared around 9:40am. The widths of the incident boxes have no meaning. Of the eight incidents on this day, half had no visible impact on delay.

$\theta_{s_1}$  is three dimensional, so (1) can be found by evaluating the objective function on a 1000 point grid with ten points per dimension. Evaluating, of course, means running  $(S_1, \theta_{s_1})$  on the loop detector data to produce the incident detection calls and then following the steps in sections 3.1 – 3.4 to compute the costs.

Figure 3 superimposes the incident detection calls for a certain value of  $\theta_{s_1}$  onto the delay contour plot. Incident #117 is quickly and consistently detected. Incidents #104 and #108 are also detected. However, #107 is completely missed, and we also have a large number of erroneous detection points between loop detectors 16 and 3. A closer inspection of the data reveals that loop detector 16 has extremely high occupancies — much higher than any of the other loops in this section of freeway. This loop detector station appears to have been mis-tuned and was erroneously reporting high occupancy values. This is especially damaging to the California Algorithm which is specifically looking for differences in occupancy values. Nevertheless, to simulate realistic conditions we leave these erroneous readings in our data.

These detection calls are then passed to our dispatcher algorithm (section 3.4). The resulting dispatches are shown as boxes on Figure 4. The eleven dispatches at the bottom of the plot are false alarms. It is clear that a better dispatching strategy might recognize this lengthy pattern of false alarms and eventually ignore the calls between these faulty detectors. However, our naive dispatcher leads to a conservative estimate of the cost of the AID.

Finally, the tow truck dispatch information is processed according to sections 3.1 – 3.3 with  $K_D = 10$  and  $K_T = 70$  to arrive at the costs  $f(S_1, \theta_{s_1}, D)$  and  $g(S_1, \theta_{s_1}, D)$  for our particular value of  $\theta_{s_1}$ . Repeating this process for all 1000 parameters on our grid, we arrive at a four dimensional

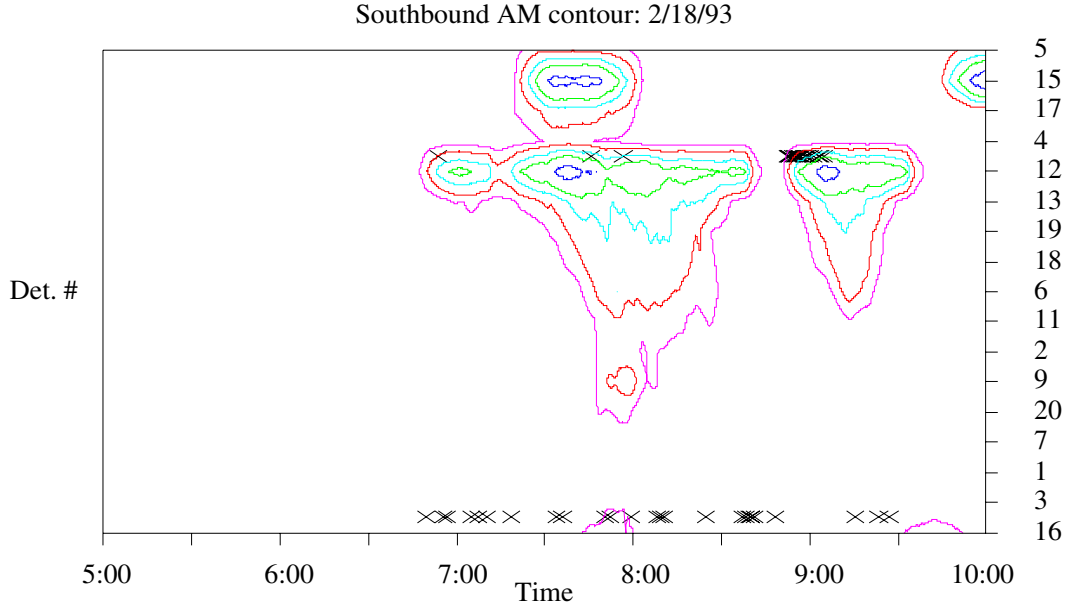


Figure 3: Delay contour plot with detection points (marked by “X”’s) for  $S_1$  and a particular  $T_{s_1}$ . The detection points near the bottom result from faulty loop detector readings.

cost surface as a function of  $\theta_{s_1} = (\theta_1, \theta_2, \theta_3)$ . Figure 5 is a graph of this surface holding  $\theta_2 = 0.4$  and letting  $\theta_1$  and  $\theta_3$  vary.

The corner of Figure 5 around  $\theta_1 = 35$  and  $\theta_3 = 0.5$  represents the region where the AID is not dispatching any tow trucks at all — i.e. the do-nothing cost,  $h(S_0, D)$ . The parameters are set so high that nothing is ever detected. As  $\theta_3$  and  $\theta_1$  decrease, the cost rises initially. This results from the algorithm beginning to respond to the false alarms. The dips in the cost surface as the parameters decrease are due to the algorithm detecting incidents and responding in time to reap benefit from the reduced delay. For the extreme lowest settings of the parameters,  $(S_1, \theta_{s_1})$  produces a huge number of false alarms, blowing the cost up to about twice  $h(S_0, D)$ . The minimum total cost of  $h(S_1, D) = \$3505$  occurs at  $\theta_{s_1}^* = [19, 0.4, 0.3]$ .

That  $h(S_1, D)$  is considerably less than  $h(S_0, D)$  — even using the faulty detector data — suggests that  $\theta_{s_1}^*$  is a good set of parameters for this particular freeway and day. Of course, one day is far from adequate to settle on parameters. We note that it is entirely plausible for a family of detection algorithms to yield a cost that is no better than the do-nothing cost over the entire range of reasonable parameters. In this case, the particular AID family may not be beneficial (in terms of our costs) for the given freeway.

### 3.6 Varying Cost Assumptions

The ratio of the cost of dispatching a tow truck,  $K_T$ , to the cost per vehicle-hour of delay,  $K_D$ , is crucial in determining how  $S_1$  compares to  $S_0$ . A low ratio means  $S$  has less to lose from dispatching tow trucks. In the previous section we used  $K_T/K_D = 7$ , which may not be appropriate for all regions. A natural performance curve is to plot  $h(S, D)/h(S_0, D)$  as function of  $K_T/K_D$ . Figure 6 is an example of such a plot for  $S_1$  on our single day. Note that  $h(S_1, D)$  occurs at a potentially different  $\theta_{s_1}^*$  for each value of  $K_T/K_D$ . From Figure 6 we learn that  $h(S_1, D) < h(S_0, D)$  whenever  $K_T/K_D$  is below about 15.5. At  $K_T/K_D = 7$ ,  $S_1$  reduces the total cost to approximately 80% of the cost of  $S_0$ . The curve is capped at 1.0 because the parameters can always be set high enough so

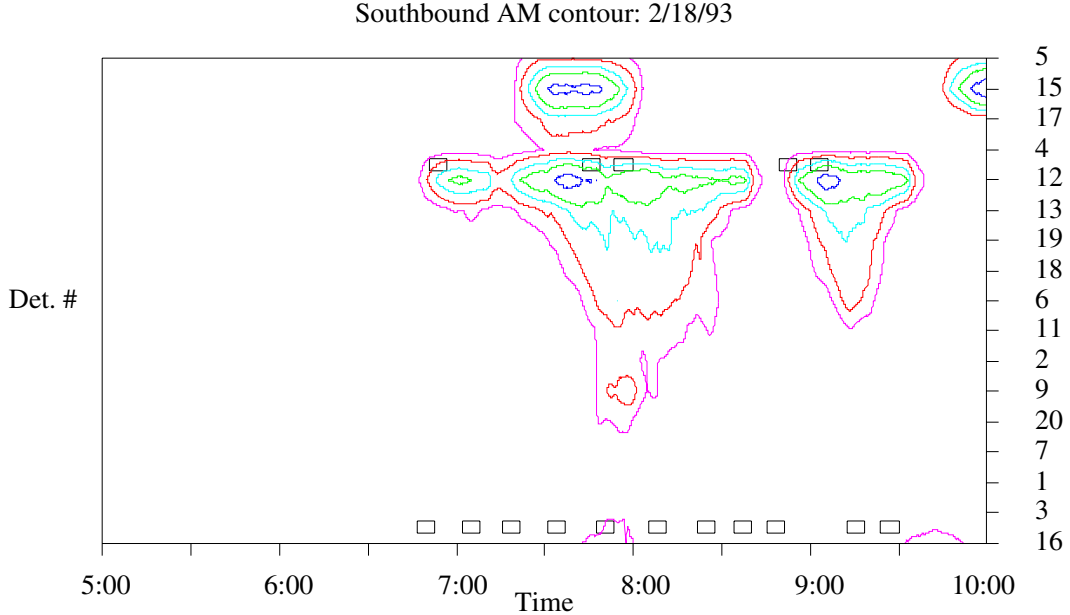


Figure 4: Delay contour plot with dispatch points (marked by boxes) based on the calls from Figure 3. While we dispatch correctly to incidents #104 and #117, the dispatch for incident #108 is too late to be of benefit. The eleven dispatches at the bottom are false alarms.

that no tow trucks are dispatched, in which case  $S_0$  and  $(S_1, \theta_{s_1})$  coincide.

### 3.7 Multiple Day Results

The performance curve given in Figure 6 is based on a single three hour period which, as previously indicated, is well-suited for detection. In order to fully evaluate the performance of an algorithm, we need to tune the algorithm with a larger and more varied data set. The results presented below use eight days from the I-880 database. For each day we take four three-hour periods: northbound and southbound for the AM and PM commutes. During these times there was recurrent congestion in the southbound direction during the AM shift and in the northbound direction during the PM shift.

These 96 hours of data contain a total of 76 incidents witnessed by at least two probe vehicles. Figure 7 displays some key characteristics of these incidents, where delay is calculated as in Section 3.1. Note that most incidents have zero delay, and consequently a cost-effective AID would ignore them. Many traditional studies would screen out by hand these zero-delay incidents as “undetectable” before attempting to estimate DR-FAR curves. In our approach no such screening is necessary. Responding to a zero-delay incident simply incurs a cost without the benefit of reduction in delay.

The performance curve from training on all 96 hours is given in Figure 8. The break-even point for the California Algorithm drops substantially to  $K_T/K_D \approx 8$ , and the reduction in cost is slight at best over all values of  $K_T/K_D$ . The presence of recurrent congestion is likely to blame for this diminished performance relative to our single-day results. Further, with eight times as much data in our training set, the AID faces a greater range of conditions and incidents over which to optimize its parameters. Also shown is the performance curve for the Minnesota Algorithm [Stephanedes and Chassiakos, 1993] generated over the same eight days. The Minnesota Algorithm appears to perform better than the California Algorithm for this data and the proposed

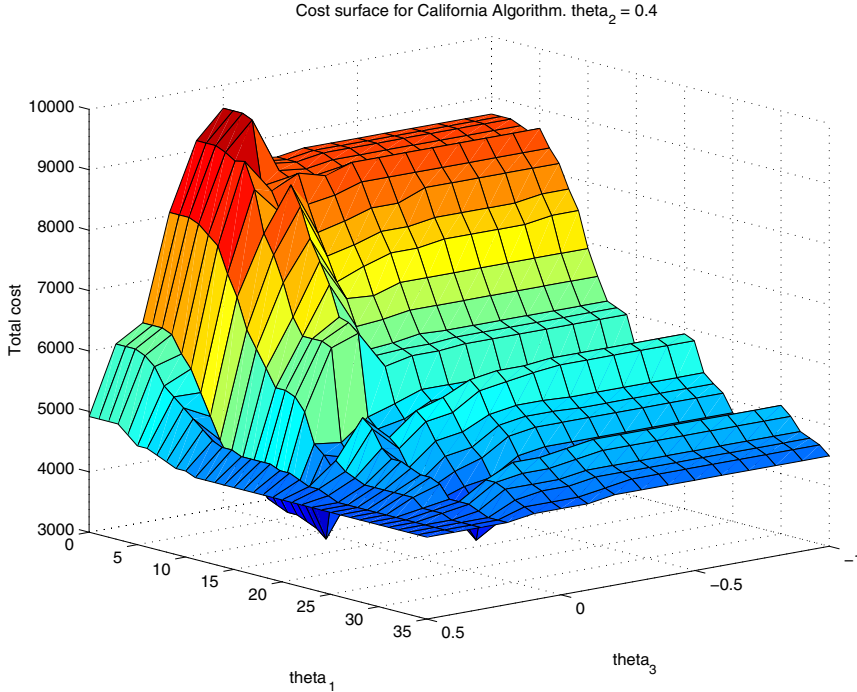


Figure 5: Cost surface for California Algorithm holding second parameter fixed at 0.4. Global minimum of \$3505 occurs at  $\theta_{s_1} = [19, 0.4, 0.3]$ . The nearest corner shows the “do-nothing” base cost of \$4355 for this day.

cost assumptions.

## 4 Discussion

The cost analysis we propose is quite flexible and allows for building cost models appropriate for particular problems. In this section we mention a few possibilities.

One major assumption is that  $T = \Delta = 10$  minutes for all incidents. That is, each incident is reached ten minutes after its dispatch and is cleared exactly ten minutes later. A natural generalization is to make  $T$  depend on location, time of day, and the speed and occupancy near the incident. Also,  $\Delta$  could depend on the type of incident (accident, breakdown, debris, mid-lane, shoulder, etc.). Allowing random  $T$ 's and  $\Delta$ 's with distributions depending on the above variables is possible as well. In some rare circumstances, data may even exist on driving and/or clearing times for tow trucks that actually arrived at the scenes of incidents. Such data could be used to estimate  $T$ 's and/or  $\Delta$ 's.

The cost assumptions (6) and (7) could be generalized in a similar fashion. The cost of delay could depend on time and space. For instance, delay during the AM commute could be made costlier than delay during the PM, since the former reduces productivity by shortening the work day. Extra components of cost could also be incurred for delay experienced during certain types of incidents (eg. injury accidents). We then generalize (6) to

$$f(S, \theta_s, D) = \sum_k d_s^k \left[ K_D(x_k, t_k) + \sum_j \beta_j I(\text{incident } k \text{ is of type } j) \right] \quad (8)$$

where  $x_k$  and  $t_k$  are respectively the location and time of the  $k$ th incident. Here  $j$  sums over a set

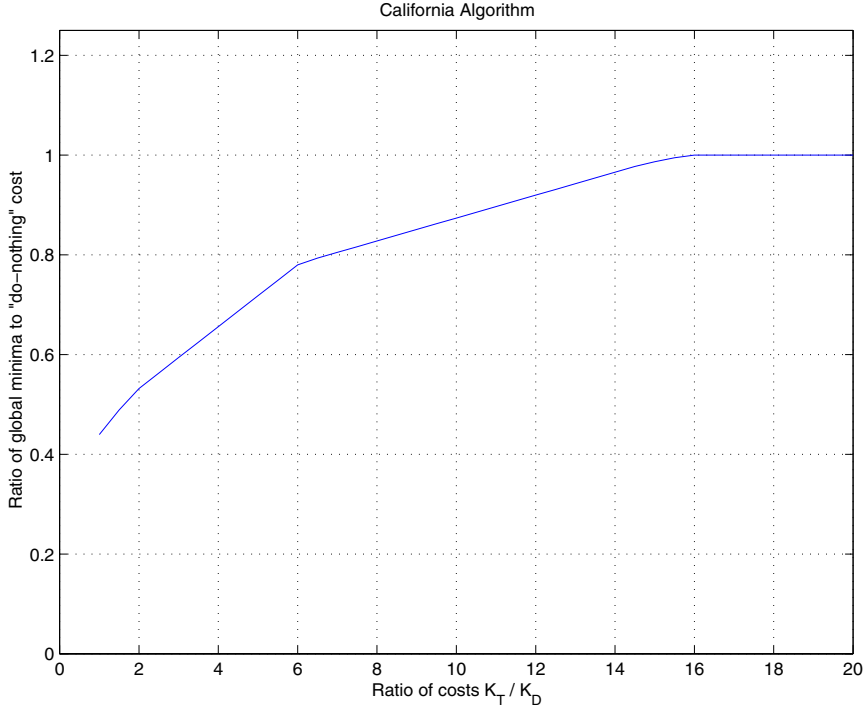


Figure 6:  $h(S_1, D)/h(S_0, D)$  as a function of  $K_T/K_D$ . As dispatching becomes more expensive relative to the cost of delay, the best-case benefits of  $S_1$  diminish relative to  $S_0$ .

of incident types, and the  $\beta_j$  are additional cost factors for delay experienced for the corresponding type of incident.  $I(\cdot)$  is an indicator function. This is by no means the most general form possible, and cost need not even be a linear function of delay.

Likewise, a potentially useful generalization of (7) is

$$g(S, \theta_s, D) = \alpha + \sum_{x,t} D(x, t) \left[ K_T(x, t) + \sum_j \gamma_j I(\text{incident } k \text{ is of type } j) \right]. \quad (9)$$

Here the  $\gamma$ 's play a similar role as the  $\beta$ 's in (8). This is quite sensible since certain types of incidents may be costlier for a tow truck to remove than others. The constant  $\alpha$  is overhead incurred even if no dispatches are made. Resources for processing the loop data, additional infrastructure, and extra labor are among the items that could be considered overhead. Note that if  $\alpha > 0$ , then the ratios in Figure 8 are no longer bounded by one.

Another assumption is that of the standard queuing model which underlies our link between duration and delay (3). However, other models are possible. For instance, one can postulate and analyze a model in which an incident leads to a velocity drop which is eliminated after clearance. With enough incident data and tow truck arrival data one might even try to establish an empirical relationship between tow truck arrival times and durations for different types of incidents. Whether such modifications are important and/or worthwhile will require additional research.

One final generalization is worth mentioning; our method allows comparisons between an AID and an incident management method that does not inherently have a performance curve. For example, consider the Freeway Service Patrol (FSP) roving tow truck system [Skabardonis et al., 1995, M. Morris, W. Lee, 1994]. We can still estimate implementation costs and delay costs using historical data and a few assumptions about how the FSP would operate. The common currency of cost

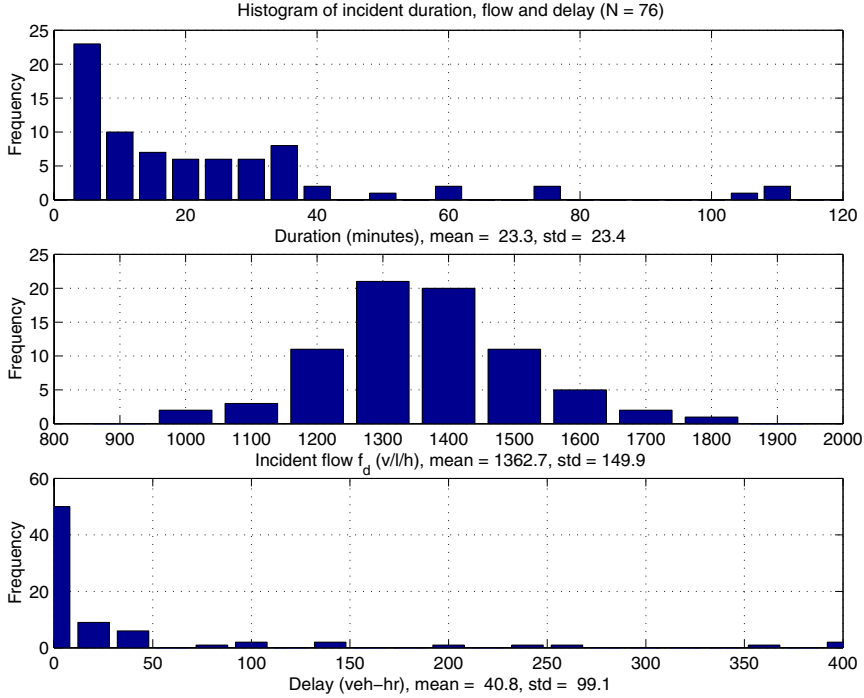


Figure 7: Characteristics of the 76 incidents from the extended training set. Histograms show incident duration,  $t$ , starting flow,  $f_d$ , and delay,  $d_0$ .

allows for direct comparisons with more standard AID's, even though the usual notions of detection rate and false alarm rate do not make sense for FSP.

## 5 Conclusions

We have presented a methodology for systematically tuning the parameters of an AID and for making truly fair comparisons between different types of AID's. This approach avoids several problems inherent in the generation of the traditional DR-FAR curve. Namely, subjective pre-screening of incidents for detectability is not necessary, and both the severity of an incident (in terms of congestion) and the time to detect are factored into the analysis automatically.

While the cost assumptions underlying the analysis can never completely capture all of the subtle aims of incident detection, our flexible framework allows one to tailor the cost functions to the problem at hand to produce a meaningful univariate cost. The advantage of this reduction is that AID parameters can be set systematically by searching over the parameter space for the lowest cost with respect to a given set of training data. Moreover, by mimicking the real costs of implementation and congestion, our results have the advantage of being practically interpretable.

## References

- [Abdulhai and Ritchie, 1997] Abdulhai, B. and Ritchie, S. G. (1997). Development of a universally transferable freeway incident detection framework. In *The 76th Annual Meeting of the Transportation Research Board*, number 971121, Washington, D. C.

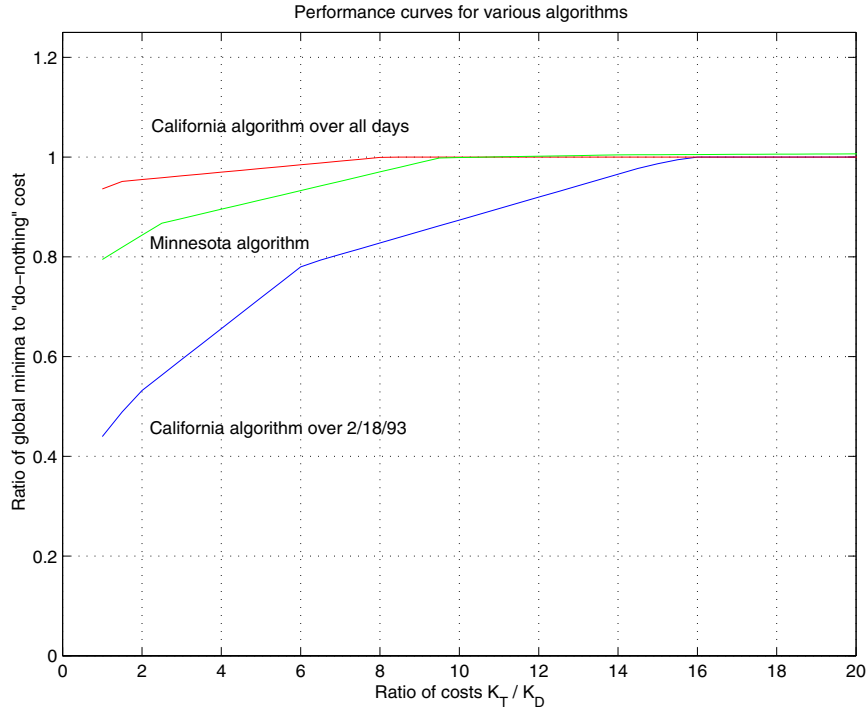


Figure 8: The top two lines are  $h(S_1, D)/h(S_0, D)$  and  $h(S_2, D)/h(S_0, D)$  versus  $K_T/K_D$ , respectively, for the eight-day set of training data. For reference, the third line is that of Figure 6.

- [Balke, 1993] Balke, K. N. (1993). An evaluation of existing incident detection algorithms. Technical Report 1232-20, Texas Transportation Institute, Texas A&M University System, College Station, Tex.
- [Epps et al., 1994] Epps, A., Cheng, J. C., and May, A. D. (1994). Developing methodologies for quantifying freeway congestion delay. Technical Report UCB-ITS-RR-94-2, Institute of Transportation Studies, University of California, Berkeley.
- [Hsiao et al., 1994] Hsiao, C., Lin, C., and Cassidy, M. (1994). Application of fuzzy logic and neural networks to automatically detect freeway traffic incidents. *Journal of Transportation Engineering*, 120(5):753–772.
- [J. A. Lindley, 1986] J. A. Lindley (1986). Qualification of Urban Freeway Congestion and Analysis of Remedial Measures. Technical Report RD/87-052, FHWA, Washington, D.C.
- [M. Morris, W. Lee, 1994] M. Morris, W. Lee (1994). Survey of Efforts to Evaluate Freeway Service Patrols. *Transportation Research Record*, (1446):77–85.
- [May, 1990] May, A. D. (1990). *Traffic flow fundamentals*. Prentice Hall, Englewood Cliffs, N.J.
- [Payne et al., 1976] Payne, H. J., Helfenbein, E. D., and Knobel, H. C. (1976). Development and testing of incident-detection algorithms: Vol. 2, research methodology and detailed results. Technical Report FHWA-RD-76-12, Federal Highway Administration, Washington, D.C.
- [Payne and Tignor, 1978] Payne, H. J. and Tignor, S. C. (1978). Freeway incident-detection algorithms based on decision trees with states. *Transportation Research Record*, (682):30–37.



- [Persaud and Hall, 1989] Persaud, B. N. and Hall, F. L. (1989). Catastrophe theory and patterns in 30-second freeway traffic data - implications for incident detection. *Transportation Research - Part A*, 23A(2):103–113.
- [Petty, 1997] Petty, K. F. (1997). *Incidents on the Freeway: Detection and Management*. PhD thesis, Department of Electrical Engineering and Computer Sciences, University of California, Berkeley. As of 1/1/98 this can be found on-line under the web page <http://www.path.berkeley.edu/~pettyk/>.
- [Petty et al., 1996] Petty, K. F., Noeimi, H., Sanwal, K., Rydzewski, D., Skabardonis, A., Varaiya, P., and Al-Deek, H. (1996). The Freeway Service Patrol Evaluation Project: Database, Support Programs, and Accessibility. *Transportation Research, Part C (Emerging Technologies)*, 4(2):71–85.
- [Press et al., 1992] Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T. (1992). *Numerical Recipes: The Art of Scientific Computing*. Cambridge University Press, 2nd edition.
- [Ritchie and Cheu, 1993] Ritchie, S. G. and Cheu, R. L. (1993). Simulation of freeway incident detection using artificial neural networks. *Transportation Research - Part C*, 1C(3):203–217.
- [Skabardonis et al., 1995] Skabardonis, A., Noeimi, H., Petty, K., Rydzewski, D., Varaiya, P. P., and Al-Deek, H. (1995). Freeway Service Patrols Evaluation. Technical Report UCB-ITS-PRR-95-5, Institute of Transportation Studies, University of California, Berkeley.
- [Skabardonis et al., 1996] Skabardonis, A., Petty, K., Noeimi, H., Rydzewski, D., and Varaiya, P. P. (1996). The I-880 field experiment: database development and incident delay estimation procedures. *Transportation Research Record*, (1554):128–135.
- [Skabardonis et al., 1997] Skabardonis, A., Petty, K. F., Bertini, R. L., and Varaiya, P. P. (1997). The I-880 field experiment: analysis of the incident data. *Transportation Research Record*, (1603).
- [Stephanedes and Chassiakos, 1993] Stephanedes, Y. J. and Chassiakos, A. P. (1993). Freeway incident detection through filtering. *Transportation Research - Part C*, 1C(3):219–233.
- [Tsai and Case, 1979] Tsai, J. and Case, E. R. (1979). Development of freeway incident-detection algorithms by using pattern recognition techniques. *Transportation Research Record*, (722):113–116.

## Appendix: Description of Incident Detection Algorithms

### The California Algorithm

The Basic California Algorithm [Payne et al., 1976, Payne and Tignor, 1978] compares three functions of the occupancy at two adjacent stations to three parameters  $(\theta_1, \theta_2, \theta_3)$  which must be set by the practitioner. If we denote the occupancy at station  $i$  over time period  $t$  as  $o_i(t)$ , with  $j > i$  being a station downstream, then the three functions are:

$$docc_i(t) = o_i(t) - o_{i+1}(t), \quad (10)$$

$$drocc_i(t) = \frac{o_i(t) - o_{i+1}(t)}{o_i(t)}, \quad (11)$$

$$drtocc_i(t) = \frac{o_{i+1}(t-d) - o_{i+1}(t)}{o_{i+1}(t-d)}, \quad (12)$$

An incident is declared to exist during time period  $t$  between stations  $i$  and  $i + 1$  if all three variables exceed their respective parameters.

### The Minnesota Algorithm

The Minnesota Algorithm [Stephanedes and Chassiakos, 1993] also looks at functions of occupancy at adjacent stations. However it first takes temporal moving averages in order to filter out spurious differences that would produce false alarms. Four parameters must be set  $(m, n, \psi_1, \psi_2)$ . The integers  $m$  and  $n$  control the amount of averaging and are usually set such that  $m\Delta t = 5$  minutes and  $n\Delta t = 3$  minutes (where  $\Delta t$  is the length of each time period). We follow these conventional choices. The free parameters  $\psi_1$  and  $\psi_2$  are thresholds against which two functions of occupancy are compared.

Using the notation from above, define

$$K_i(t) = \max \left( \frac{1}{m} \sum_{j=1}^m o_i(t-j), \frac{1}{m} \sum_{j=1}^m o_{i+1}(t-j) \right) \quad (13)$$

Then the two functions of occupancy are

$$u_i(t) = \frac{\frac{1}{n} \sum_{j=0}^{n-1} docc_i(t+j)}{K_i(t)} \quad (14)$$

$$w_i(t) = \frac{\frac{1}{n} \sum_{j=0}^{n-1} docc_i(t+j) - \frac{1}{m} \sum_{j=1}^m docc_i(t-j)}{K_i(t)} \quad (15)$$

An incident is declared to have occurred during time period  $t$  between stations  $i$  and  $i + 1$  if both variables exceed their respective parameters. Notice that (14) and (15) can not be evaluated until time period  $t + n$ .