Detection of incidents and events in urban networks

T. Thomas  E.C. van Berkum

Centre of Transport Studies, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands
E-mail: t.thomas@utwente.nl

Abstract: Events and incidents are relatively rare, but they often have a negative impact on traffic. Reliable travel demand predictions during events and incident detection algorithms are thus essential. The authors study link flows that were collected throughout the Dutch city of Almelo. We show that reliable, event-related demand forecasting is possible, but predictions can be improved if exact start and end times of events are known, and demand variations are monitored conscientiously. For incident detection, we adopt a method that is based on the detection of outliers. Our algorithm detects most outliers, while the fraction of detections due to noisy data is only a few percent. Although our method is less suitable for automatic incident detection, it can be used in an urban warning system that alerts managers in case of a possible incident. It also enables us to study incidents off-line. In doing so, we find that a significant fraction of traffic changes route during an incident.

1 Introduction

Traffic flows consist of recurrent and non-recurrent patterns. A recurrent pattern repeats itself with a known period and is, therefore, predictable. An example is a recurrent event, like a match of a professional football club. Non-recurrent patterns are caused by unexpected events or incidents. Despite the fact that they are relatively rare, incidents often have a very negative impact on the traffic situation (e.g. [1]). An important issue for policy makers has been to reduce these negative effects.

It is, therefore, not surprising that much effort has gone into the development of automatic incident detection (AID) algorithms (e.g. [2, 3]). The detection of incidents, however, is not an easy task, because incidents are rare and not the same. In the past AID has depended on the collection of flow data and velocities through the use of induction loops. Unfortunately, these data trace the results of incidents, i.e. congestion, so that incidents are relatively difficult to detect. Nowadays more measurements are collected along complete trajectories, achieved through the use of closed-circuit television (CCTV) and radar.

There are two types of algorithms that use these measurements for incident detection. The first group of algorithms is based on ‘positive identification’ or recognition, i.e. an incident is identified when the data is similar to those during an incident from the past. This group contains learning algorithms, such as neural networks (e.g. [4, 5]), and the McMaster algorithm, which distinguishes different traffic situations (e.g. incidents) based on the location of measurements in the occupancy-flow diagram (e.g. [6]). Algorithms based on recognition may be sophisticated, but their performance is only superior when all of the incidents provide rather similar data. This may be the case for a highway link or tunnel, but is less so for an urban network. However, even for highway links, each incident is unique, and, therefore, difficult to detect.

The second group contains algorithms that identify outliers. These algorithms are based on ‘negative identification’, i.e. an incident is identified when the measurements deviate significantly from the ‘normal’ situation. In time-series algorithms, for example, an incident is detected when the measurements differ significantly from the predictions (e.g. [7, 8]), and in decision structure algorithms incidents are detected when measurements exceed thresholds in a decision tree (e.g. [9–11]). Algorithms based on ‘negative identification’ are less sophisticated, but they may be more suitable for an urban network.
Incidents, however, are not the only source of concern for traffic managers. Events, for example, also have a negative impact on the traffic situation. Fortunately, large events are known in advance, and the amount of traffic they generate can, in principle, be predicted. Unfortunately, relatively little research on this subject has been done so far. One contentious issue, for example, is how to define an event. In the USA the definition of an event (e.g. [12, 13]) is specifically related to mobility, but in The Netherlands other characteristics (e.g. noise pollution) are also included. Authors may also adopt different classifications of event types (e.g. [12, 14]). However, it is widely accepted that special monitoring and management measures are necessary during 'events'.

In this paper we develop a prediction scheme for event-related demand forecasting. In addition, we use an algorithm to identify outliers, which enables us to study link flow patterns during incidents. In Section 2 we describe the data, and in Section 3 we describe the prediction method for recurrent events. In Section 4 we apply our detection algorithm to identify outliers. We separate the outliers that are most likely caused by incidents, and we analyse these in more detail in Section 5. We discuss our results in Section 6.

2 Data

The study area for this research consists of the urban network of the Dutch city of Almelo. Data were collected at about 20 signalised urban intersections from September 2004 to September 2005. Vehicles were detected by means of inductive loop detectors. The data were processed into link flow measurements per link per time interval. We define a flow profile \( Q_{dl} = (q_{d1l}, \ldots, q_{dnl}) \) as a time-series of \( n \) intervals for day \( d \) and link \( l \). In most cases, measurements were provided in 5-min intervals, so that \( n = 288 \). However, for about 30% of the links only measurements in 30-min intervals were read out by the software of the detection instruments. For these links, \( n = 48 \). In Fig. 1 we show the study area and the links for which data were collected. Note that according to Fig. 1, few data were collected in the south east of Almelo. The municipality probably decided to collect only data along the main roads, which are absent in that part of Almelo.

The measurements were inspected and invalid data were rejected by Weijermars [15]. Invalid data are the result of errors in the measurements (e.g. by failures in the electronics or by miscounts in the number of passing cars). These errors were detected by using certain criteria (e.g. flows should be equal or larger than 0, flows should not exceed a certain maximum volume and 24 h flows should be larger than 0). Due to the malfunctioning of detectors during several days or even weeks, a significant fraction of the profiles was rejected. It is worth stressing that the remaining profiles only contain raw data. Thus, we excluded bad data, but did not include ‘artificial’ data or in any other way manipulated the measurements. To create a
homogeneous sample with good statistics, we only included links with at least ten volume profiles per weekday (Mondays, Tuesdays etc.). This leaves us a sample of 48 intersection links which have 5 min time-series. Note that we excluded time-series with 30 min intervals, because the resolution of these time-series is not suitable for predictions of event-related traffic flows, which vary on shorter time-scales.

It is worth mentioning that we only used loop detector data in this study. We were not able to use incident data, which are difficult to obtain. Unfortunately, it was, therefore, not possible to validate some results from our incident detection scheme.

3 Demand forecasting of a recurrent event

When large events take place, traffic flows are influenced by the large number of visitors that attend these events. At certain locations this will result in a significant increase of traffic before and after the event. In Almelo most of these demand peaks occur when the local professional football team, Heracles, plays their home matches. Their football stadium is located in the southern part of the city. In this section we analyse the flow patterns, related to these football matches, in more detail. This analysis forms the basis for event-related demand forecasting. In fact, we will predict link flows, but because there is no (real) congestion in Almelo, link flow and demand mean the same thing here.

The concept behind our forecasts is straightforward. The demand during an event can be separated into two parts, i.e. demand generated by visitors of the event and demand generated by background traffic that has no relation to the event. We assume that both demands are uncorrelated. In that case, the background demand is similar to the demand on a day without an event. The event-related demand is the difference between the total demand and the predicted background demand. Accurate predictions, $q_{\text{pred}}$, of this background demand are thus essential.

The average profile for a day without a football match can serve as an estimate for the background demand. This average profile is obtained from historical data. Because weekdays show different flow patterns, the average profiles were obtained for different weekdays (i.e. Mondays, Tuesdays, Wednesdays, Thursday, Fridays, Saturdays, Sundays and Holidays). However, link flows also show seasonal and weather-related variations. To factor in these variations into the predictions, the relations between external factors (such as weather) and travel demand should be studied. Such a study, however, is complicated and requires many reliable data sources, which are often unavailable. A more practical approach was, therefore, applied, for which the following assumption was made. Seasonal and weather related variations change relatively slowly (e.g. seasonal variations have got time-scales longer than one day) with time. This implies that if there is more traffic than average on a particular day, there is a high probability that there will be more traffic the next day as well. Similarly, if there is more traffic than average at a specific hour, there will most likely be more traffic in the next hour. This assumption could be translated into quite reliable predictions, which are presented in a forthcoming paper.

We selected 10 days on which Heracles played their home matches, starting at 20.00 h. At 26 links we could detect extra traffic due to the football match. For these links, we predicted the background demand, as described above. We then subtracted the background from the observed profile. The average (over the selected days) of the residual (difference) profile is shown in Fig. 2, in which we show some examples from different links. Note that only a few links have profiles for all the 10 selected days, because a significant fraction of profiles was rejected due to the malfunction of detectors (see Section 2).

In the left panel of Fig. 2 we show links that serve traffic towards the event, and in the right panel links that serve traffic from the event. The time plotted on the $x$-ordinate is relative to the start and end times, respectively. The link flows show the well-defined peaks, although the data is quite noisy in some cases. The observed patterns in Fig. 2 can be approximated by normal distributions (solid lines). According to these distributions, the peak intensity occurs, on an average, slightly more than 35 min before the start time of the event, and slightly less than 25 min after the end time. The typical width (standard deviation) of the arrival peak is about 15 min. The width of the departure peak is even smaller, i.e. less than 10 min, because most visitors leave at the same time. In general, the residual profiles are approximately 0 before the visitors start to travel to the stadium (before the arrival peak) and during the

![Figure 2 Event-related link flow patterns before the start of the event (left) and after the end of the event (right) at different intersection links](image-url)
event itself (in between the arrival and departure peak). We, therefore, conclude that we have estimated the background demand quite well. However, the departure peak often shows a tail. This surplus of traffic is shown in the upper and centre right plot of Fig. 2. This tail is probably caused by the fact that some visitors do not leave immediately after the match, but rather hang around.

The normal distributions were estimated for all arrival and departure peaks, and served as an estimate for the event-related demand. Thus, by adding them to the prediction of the background traffic, we obtained the total demand forecast for a day with an event. In Fig. 3 we show three examples of observed flow profiles (dotted lines), $q^{obs}$, for days with a football match. The links in the example serve traffic from the stadium, and they show a clear peak after the match (around 22.00 h). The predictions of the total demand are shown by the solid lines.

From the upper panel in Fig. 3, we conclude that some peaks in the sample are predicted quite well, except from the tail in the departure peak. However, a significant fraction of peaks cannot be predicted accurately. In some cases, the peak is shifted with respect to the prediction, which is illustrated in the centre panel. This situation, for example, occurs when the match does not finish at the planned time. Variable start and end times cannot be predicted in advance. Predictions may be improved, although, when organisers communicate changes in start or end times to the traffic centre. In addition to uncertain start and end times, significant differences between predicted and observed demand are also observed. An example is shown in the bottom panel, in which the demand is clearly under estimated by the prediction. The predictions are based on the average profile from historical data, which thus implies that the demand must vary significantly from event-to-event. Because visitor numbers are rather constant, we suggest that demand variations are related to changes in distribution, i.e. the origins from which the visitors arrive, or modal split, which probably depends on the weather.

It may prove to be difficult to predict variations in event-related demand. However, the same visitors that arrive at an event must also leave afterwards. Thus, from the observed demand before the event, it should be possible to make predictions about the demand after the event. For two intersections, we estimated the total demand to (‘in’) and from (‘out’) the stadium. In Fig. 4 we show the results for several events. The event-related in- and out-flows ($q_{in}$ and $q_{out}$) are shown for intersections 30 (filled symbols) and 32 (open symbols). Intersection 30 is the intersection between the ring road and the main provincial road towards the west (most left in Fig. 1). Intersection 32 is an intersection on the radial road between the southern part of the city and the city centre. The in-flow of event-related traffic is from the north and west (intersection 30) or north and east (intersection 32), whereas the out-flow is from the south, and in case of intersection 32 also from the west. Note that only for six events, data were available for all the intersection links used in this analysis.

From Fig. 4 we conclude that the demand is indeed variable. Fig. 4 also shows the expected correlation between in- and out-flows, although there is some scatter in this correlation. The scatter may be caused by the fact that people take other routes on the return journeys (e.g. because the trips are part of a chain, i.e. from work to the event, and afterwards back home). However, it is more likely that the scatter is caused by inaccuracies in the demand estimates. The out-flow, for example, has a tail which has not been taken into account. This may as well explain the systematic difference, shown in Fig. 4, in which the total out-flow is lower than the total in-flow. It is also
telling that the two 'outlying' data-points (for which the out-
flows are much smaller than the in-flows) of both inter-
sections were measured on the same day.

Although event-related demand estimates may show some
inaccuracies, the correlation between in- and out-flow can be
used to improve predictions of out-flows. If spatial
correlations are taken into account, i.e. if the flows of all
incoming roads are measured and compared with each
other, then predictions of in-flows could also be improved.
For this, we need measurements from all roads that enter
the city, and from all intersections near the stadium. In our
case, these measurements are not available. Many
important links are missing in the sample. We, therefore,
 refrained from an extended analysis of spatial correlations.

4 Detection of outliers

When traffic accidents, road works or other unique events
occur, link flows can differ significantly from the normal
situation. These link flows are called outliers. In this
section we describe how outliers are detected. The concept
behind our detection method is quite straightforward. First,
we need to define the 'normal situation'. In the previous
section we argued that we can predict the regular demand,
\( \hat{q} \), quite well. These predictions, therefore, represent the
normal situation. Then, we need to define what is a
'significant' difference from the normal. This can only be
done, if the amount of noise in the measurements is known.

From Fig. 3 we conclude that the observations show
variation that is missing in the prediction. Most of this
variation looks quite random, and is, therefore, called noise.
Contrary to seasonal or weather-related variations, the noise
in different measurements is uncorrelated. The noise is,
therefore, unpredictable. Noise can have different causes,
such as the random arrival process of cars, which is an
important source of variation on highways. In urban areas,
variable and unknown cycle times of traffic signals can also
contribute to the noise. In practice, all variations which
have short time-scales, and which do not follow a recurrent
pattern can be considered as noise. Note that random
miscounts in the measurements also contribute to the noise.

If the amount of noise increases, it will become more
difficult to separate unexpected variations due to incidents
from noisy measurements. A deviation from the 'normal' is,
therefore, referred to as significant when the absolute
difference between measurement \( q_{obs} \) and prediction \( q_{pred} \) is
more than \( n \sqrt{\hat{q}_{pred}} \). The factor \( n \) can be chosen rather
arbitrarily. For large \( n \), the algorithm will detect outliers,
for which there is a small probability that they are caused
by noisy measurements. In this case, however, real outliers
might not be detected. For small \( n \), this is not a problem,
but in that case there is a larger probability that noisy
measurements are misidentified as outliers.

We tested our detection algorithm on time-series with
10 min time-lags. With these time-lags more robust results
are obtained than with 5 min time-lags, because aggregation
reduces the noise levels caused by the variability in green
times of traffic signals (e.g. [17]). For our time-series with
10 min time-lags, we get reasonable results when we choose
\( n = 4 \). The fraction of detected outliers for this limit is
about 0.2%. This means that 2 out of 1000 measurements
were identified as an outlier. Although this is a small
number, it is still significant. In most 10 min intervals, flows
are significantly larger than 10 vehicles. For these
measurements, the Poisson noise can be approximated by a
normal distribution with variance equal to the expected
flow. Given a Normal distribution, we can calculate the
probability that a noisy measurement is misidentified as an
outlier. This probability is \( 6.3 \times 10^{-5} \) for \( n = 4 \).
The fraction of misidentifications as a result of noisy data is,
therefore, approximately 3% (\( 6.3 \times 10^{-5} / 0.002 \)). For
\( n = 3 \), this fraction would be approximately 50%, which is
probably too high to be useful in a detection algorithm.
However, the \( n = 3 \) detection limit can be used when we
require that two successive measurements are outliers. The
number of multiple detections (two or more detections in a
row) is about 0.2%, but the number of misidentifications
\( (0.0027 \times 0.0027 \approx 7.3 \times 10^{-6} \) per measurement) is much
smaller. The probability that a multiple detection is caused
by noisy measurements is, therefore, very small (\(<1\%\)). We
use the \( n = 3 \) detection limit to detect multiple outliers,
which are not very extreme (some of them would not be
detected by the \( n = 4 \) detection limit), but which have long
time-scales (20 min or more).

In the literature (e.g. [7]) the following measures are used
for evaluating an incident algorithm: detection rate (DR),
false alarm rate (FAR) and mean time to detect (MTTD).
We showed that the false alarm rate is only a few percent
for our algorithm. From visual inspection, we concluded
that we probably detected most outliers. The mean time to
detect lies between 10 and 20 min. With smaller time-lags
in our time-series (i.e. 5 min) we may reduce the detection
time to 5 min, but due to higher noise levels, only more
severe incidents will be detectable in that case.

However, our algorithm only detects outliers. Besides
incidents, other negative outliers are detected as well.
Furthermore, our MTTD is relatively large. This method is thus less suitable for AID. On the other hand, only about 15% of all daily profiles contain one or more outliers between 6 and 24 h. All the other profiles have regular shapes. Because outliers are rare, traffic managers will only need a limited amount of time to check their causes. This algorithm might, therefore, be included in a warning system that alerts managers in case of a possible incident or unexpected event. Moreover, outliers that are possible incidents can be selected and studied off-line. This is done in the next section.

5 Incident analysis

Apart from events and road works, unexpected deviations in flows are almost always the result of incidents. A negative outlier may indicate that an incident is happening. In our sample we selected more than 200 negative outliers (single or multiple detections). If these are all caused by incidents, this would correspond to about two incidents per day on the whole network. Note that we would have liked to validate this assumption. However, we were not able to obtain independent data on incidents. Some of the outliers are relatively small and may be caused by a decrease in capacity due to, for example, a blockade of a truck. Other outliers are extreme and those may be caused by a significant decrease in the capacity after a large accident has happened. In Fig. 5 we show examples of relatively small outliers (left panel) and extreme outliers (right panel). The outliers are detected with both the \( n = 3 \) (open symbols) and \( n = 4 \) (solid symbols) detection limits.

According to the figure, the outlying events show a typical pattern. A dip in the measured time-series (dotted line) is followed by an excess of traffic (compared with the prediction; solid line). This typical pattern we find for most of the negative outliers. The excess of traffic suggests that queued traffic is dispersed. We, therefore, conclude that most negative outliers must be caused by an incident during which traffic is queuing. We call the duration of the dip the incident period, \( T_{\text{incident}} \), and that of the excess in traffic the recovery period, \( T_{\text{recover}} \). We find that the incident period is 20 min or less for small incidents. The recovery period typically lies between 20 and 40 min (twice the incident period). Note that we cannot detect periods less than 20 min, because we used time-series with 10 min time-lags. Incident periods of large accidents are typically between 60 and 90 min, and recovery periods are of the same length. The incident and recovery periods are thus rather variable. These variations depend on several things, such as the nature of the incident and the time it takes for the emergency services to attend.

The depth of the dip is also variable. In the upper right panel of Fig. 5 the minimum flow during the accident is close to 0, which could be due to a temporary closure of the road. The accident in the bottom right panel is less extreme. In that case, the minimum flow is about 50% of the expected.

In the examples of Fig. 5 we discriminate between small and large incidents. We would like to quantify this rather qualitative difference by a certain strength parameter. We choose to define the strength by the equivalent width (EW), which is defined for 10 min time-series as

\[
EW = 10 \sum \left( \frac{q_{\text{obs}}}{C_0 q_{\text{pred}}} - \frac{q_{\text{pred}}}{q_{\text{pred}}} \right)
\]

The mean expected link flow, \( \langle q_{\text{pred}} \rangle \), is an average over the whole period \( T = T_{\text{incident}} + T_{\text{recover}} \) (incident and recovery period). We define the start time of an incident to be 10 min before the first detected outlier. We then estimate the equivalent width, \( EW_{\text{incident}} \), during the incident

![Figure 5](image-url)
period, in which we choose $T_{\text{incident}}$ such that $E W_{\text{incident}}$ is minimised. After the incident period the recovery period starts. We choose the recovery period such that $E W_{\text{recover}}$ is maximised, but with the constraint $T_{\text{incident}} \leq T_{\text{recover}} \leq 2T_{\text{incident}}$. The equivalent width is a relative strength, i.e. relative to the mean expected link flow. The equivalent width is given in minutes, and its length multiplied by the mean expected flow equals the amount of traffic that is missing (negative width) during the incident, and the amount of extra traffic (positive width) during the recovery period. We thus demand that $E W_{\text{incident}} < 0$ and $E W_{\text{recover}} > 0$. With this constraint about 160 incidents are selected.

In Fig. 6 we show the equivalent widths for the incident and recovery period. For small incidents, there is no correlation between both equivalent widths. In fact, in some cases, the amount of extra traffic during the recovery period is even larger than the amount of missing traffic during the incident ($|E W_{\text{recover}}| > |E W_{\text{incident}}|$). These cases are probably not caused by a blockade or accident, and they are shown by the open symbols in Fig. 6. For larger incidents ($|E W_{\text{incident}}| > 10$), however, there seems to be a correlation between the two equivalent widths. If we exclude the open symbols, the average ratio between both equivalent widths is about $-0.4$. In other words, 40% (on an average) of the vehicles that have been blocked will follow their way, whereas 60% will take another route.

6 Discussion

In this paper we presented a prediction scheme for recurrent events and a detection scheme for incidents, based on link flow data that were collected at urban intersections in the Dutch city of Almelo.

We selected days with home matches of the professional football team to study the effects of a recurrent event on travel demand. We are able to separate background demand from event-related demand, and we concluded that reliable demand forecasts during events are possible. However, in a significant fraction of the cases, predictions are less reliable due to variable start and end times or due to unforeseen variations in demand. Variable start and end times may be taken into account when they are communicated by organisers of events. With modern communication technologies this should be possible. The changing demand from event-to-event is less easy to predict. However, with a cordon of measurement instruments around the city (on the important access roads) and the stadium, it should be possible to anticipate on demand variations (with respect to the forecast). Moreover, we showed that in-flow (traffic to the event) measurements can be used to update demand predictions of out-flows (traffic that leaves the event). Our results are comparable with those from a study of the area of Amsterdam [18]. In this study link flows were analysed during matches of the football club Ajax Amsterdam. These matches generate many more visitors than those in Almelo, but it was shown that reliable demand forecasts could be obtained in Amsterdam as well. In fact, the demand of event-related traffic was much more constant for these events.

With a method that is straightforward, we are able to detect outliers in flow time-series. Our detection method is based on an $n$-sigma clipping method which can detect outlying measurements in real-time. With our chosen thresholds ($n = 4$ for a single outlier and $n = 3$ for two successive outliers) we get a false alarm rate of only a few percent, although we probably identify all large incidents (e.g. accidents) and most small incidents (e.g. blockade by a vehicle). The detection time lies between the 10 and 20 min, but can be reduced when we reduce the time-lags of our time-series. We analysed the detected incidents and find that on an average 40% of the missing volume during an incident disperses along the same link after the incident. In other words, about 60% of the traffic is rerouting. It should be stressed that these figures are network dependent, and are, therefore, not necessarily similar for other locations. It may be interesting to analyse how the traffic re-routed at a specific location where an incident occurred. This can be done by comparing the flow patterns at different locations during an incident. However, due to the limited spatial resolution (at many links no data were collected) we refrained from such a spatial analysis in this study.

Some algorithms use threshold values for detecting incidents. Some authors have pointed out that the choice of threshold values is often rather arbitrary (e.g. [19]). Although we also choose thresholds, our algorithm is very successful for two reasons. Our expected link flows are robust predictions based on historical data of many days. More importantly, however, the threshold is not arbitrarily chosen but depends on the noise in link flows, which can be described in a uniform way.

In the UK, AID systems are being integrated into adaptive traffic signal systems in urban areas (e.g. [10, 20]). In The
Netherlands, authorities like to include incident detection and event predictions into the management system of Dutch motorways [21]. In these cases, multiple data sources are used. Information about occupancy and travel times in combination with spatial correlation between locations is essential. In the future, our method could be extended to other data as well. However, our algorithm detects outliers, and should, therefore, not be seen as AID, but rather as an online warning system, which indicate possible problems on the network.

7 Acknowledgments

We want to thank Vialis and the municipality of Almelo for providing us the link flow data in Almelo. This research is funded by Transumo.

8 References


