Detecting errors in loop-based flow data using a Long-term Integration Process (LIP)

Laurence Howe, Chief Traffic Engineer, National Traffic Control Centre (NTCC),

The National Traffic Control Centre (NTCC) collects traffic data in real time. Currently this is used for assessing the stress on the network. In the future, it is intended that flow data will be used for automatic detection of events involving capacity restriction and demand increases and to estimate travel times under delay conditions. Traditionally, the only way to check the accuracy of flow data has been to make a video recording of the traffic and manually count the vehicles passing a reporting point. However, there are over 4,000 reporting points on the Project Network so checking accuracy in this manner would take a very long time and, given that the majority of data are accurate, would be extremely inefficient. The Long-term Integration Process (LIP) is derived from Kirchoff's law for electric current. In the case of flow data, the currents are vehicle flows at upstream and downstream Reporting Points integrated over a long period of time. Providing all flows are known, including those of traffic joining and leaving the carriageway, if the upstream and downstream flows are different, one or more flows must be inaccurate. Work is currently under way to identify flow data that is reversed vis-à-vis the carriageway that it purports to represent.

The author: Dr L D Howe, Chief Traffic Engineer, National Traffic Control Centre (NTCC), can be contacted on +44 (0) 771 819 56161 or by email at Laurence.Howe@post.serco.com

INTRODUCTION

The Project Network of the Highways Agency (HA) comprises the motorways and trunk roads. Traffic flow information is used for a number of purposes. The HA operates the MIDAS (Motorway Incident Detection and Automatic Signalling) system to warn approaching motorists and help protect the back of queues on motorways. Historical traffic data is used for planning and strategic decision-making. The National Traffic Control Centre (NTCC), run by Serco on behalf of the HA, collects traffic data in real time. Currently this is used for assessing the stress on the network. Stress is defined as the actual flow at a particular point in the HA Project Network divided by the capacity (the maximum sustainable flow) at that point. This is used to identify congestion and to evaluate alternative routes for diversions. In the future, it is intended that flow data will be used for automatic detection of events involving capacity restriction and demand increases. It is also intended that it will be used to estimate travel times under delay conditions, following successful simulation results (TEC July 2006).

BACKGROUND

Traffic flows are calculated from induction loop data. The induction loops are buried in the road and count the vehicles that pass over them. In addition they can detect vehicle speed, length and carriageway occupancy (the percentage of the road covered by vehicles). The NTCC system calculates flows as vehicles per hour (vph) every five minutes. It does this by counting the vehicles that passed over a loop site in the previous ten minutes and multiplying the total by 6. The induction loops are known as monitoring sites. The points at which traffic flows are assessed are known as Reporting Points. About 30% of Reporting Points have no direct monitoring and use data derived from one or more remote loop sites. If the loop sites generate inaccurate data or if the derivation of the flows is incorrect, the flows will be inaccurate.

Traditionally, the only way to check the accuracy of flow data has been to make a video recording of the traffic and manually count the vehicles passing a reporting point. The counts are then compared with the flow data so that the accuracy of the data can be assessed. However, there are over 4,000 reporting points on the Project Network so checking accuracy in this manner would take a very long time and, given that the majority of data are accurate, would be extremely inefficient. NTCC makes random sample checks every month, but if inaccuracies do exist, they could remain undetected for many years. An additional problem associated with video assessment is the possibility of manual counting errors. These could result in an accurate site being assessed as inaccurate and vice versa. Where a site is initially

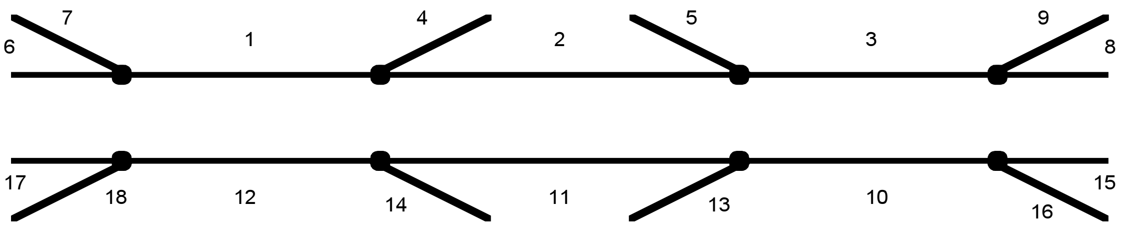

Figure 1: A typical junction set for a grade-separated junction

assessed as inaccurate, the counts can be re-checked to eliminate the errors. However, if an inaccurate site is assessed as accurate, it is unlikely to be rechecked and therefore the inaccuracy will never be detected, unless it is selected as part of a future random sample.

CAUSES OF INACCURACY

There are several possible causes of inaccurate flow data.

1. Poorly calibrated loop sensitivity results in one or more lanes either not counting all vehicles or counting vehicles in an adjacent lane;
2. Poor loop installation may result in the same effects as poor calibration, but cannot be corrected without re-installing the loops in the carriageway;
3. Loop faults may result in erratic counting, either when flows are low or when the faults are intermittent;
4. Equipment faults can lead to inconsistent counting;
5. Incorrect loop configurations can lead to traffic being miscounted;
6. Incorrect site configuration can cause data for one site to be extracted from a different site;
7. Poor calculations can lead to inaccurate flows, either because the wrong data source is used or the computation is incorrect.

Any of the above causes may result in inaccurate flow data. Once the inaccuracies have been identified, the causes can be investigated via further analysis and the faults rectified.

THE PRINCIPLES OF LIP

The Long-term Integration Process (LIP) is derived from Kirchoff's law for electric current:

> *The sum of the currents entering a node must equal the sum of the currents exiting a node.*

In the case of flow data, the currents are vehicle flows at upstream and downstream Reporting Points. The implication of this is that if the upstream and downstream flows are different, there must be a either an unknown sink into which vehicles disappear or an unknown source from which vehicles can join the road. Providing all flows are known, including those of traffic joining and leaving the carriageway, the law is inviolate.

Now, with electrical circuits, the variations in flow are transmitted at a speed close to light speed, so differences between upstream and downstream currents can never be detected. However, with traffic, the variations are limited to the maximum speed of the fastest vehicles. Thus, if we compare traffic flows between two Reporting Points, one upstream of the other, they will almost always be different, even if there is no means of joining or leaving the carriageway. But, by integrating the flows over a long period of time and taking an average, the difference between the two should be extremely small. Where significant differences exist, one or more of the flows must be inaccurate.

Upstream and downstream flow data may not always be available to the required accuracy because it is not possible to measure the entry and exit flows at unmonitored junctions. However, there is another technique that can provide additional confirmation of accuracy. The period of integration used is 168 hours, which is exactly 7 days, which means that diurnal and day of the week variations will all be averaged out. This presents the opportunity to compare opposite flows (eg northbound compared with southbound). The reasoning behind this is that if, for example, an average of 3,000 vehicles per hour (vph) are recorded as travelling southwards on the M1 towards London but only 2,000 vph are recorded travelling back, either 168,000 vehicles have stayed in London or one of the flows is inaccurate.

There is one other valid reason: one third of the vehicles travelling to London return via a different route. However, experience suggests that this is not the case. A small number of exceptions exist. These are locations on the Project Network where an imbalance does occur. Notable amongst these is the Severn Bridge, where a toll is charged to cross into Wales but not in the opposite direction. The result is that about 20% of the HGV traffic returning from Wales via the Severn Bridge took another route into Wales. The effect is less severe for cars. Once such a location is known and its effect quantified, it can be used in the same way as before, but using the required factor to compare the opposite flows.

PRACTICAL REQUIREMENTS

The Project Network was divided up into roads (eg M1, A1, etc) and each road further subdivided into junction sets. There is some variation in the configuration of the junction sets, depending on the road type. The most common type of junction set is the Motorway grade-separated junction. An example is given in Figure 1.

It can be seen that each junction set contains part of the two adjoining junction sets. Each of the assessed flows is associated with the Links (sections of carriageway) numbered from 1 to 18. The flow at Link 1 is compared with the sum of those at Links 6 and 7, with the sum of the flows at Links 2 and 4 and with the flow at Link 12. If all agree within a defined limit of accuracy, then all can be regarded as accurate. If they all disagree, the flow on Link 1 is inaccurate. If one is wrong and the others are correct, the flow on Link 1 is almost certainly accurate and the inaccuracy will be in the compared flows. All the comparisons can be verified against each other. For example, the flow at Link 2 can be verified against those at Links 3 and 5 and against that at Link 11, whilst the flow at Link 4 can be verified against that at Link 14, etc. Where inconsistencies occur, they can easily be ascribed to an individual flow. However, if there are multiple inaccuracies, the identification of the inaccurate flows may become more laborious, but these occurrences can be expected to decrease as the analysis proceeds.

The LIP is designed to run on a weekly basis. Initially, the number of undefined accuracies which were difficult to identify, because of faulty equipment that returned insufficient valid data, was large. However, as equipment was repaired the unknown accuracies fell considerably. The LIP produces average flows which are incorporated into map layers so that data can easily be checked with reference to the map of the Project Network. One of the key attributes that the LIP checks now that the process is maturing is whether or not the flow data has changed significantly. Where it remains essentially the same from one run to the next, the flows can be assumed to remain accurate. Where there is a significant change in the flows, the cause can be investigated.

The LIP is also useful for assisting the faults team in the rapid identification of faulty equipment.

SITE REVERSAL

Because the average flow in one direction is expected to be the same as that in the opposite direction, there are cases where it is impossible to determine in which direction the flow has been measured. This generally occurs on rural single carriageways where the two loop sites (one for each direction) are connected to the same equipment (although there are some other instances where this can be a problem). Initially the LIP was unable to identify whether the flow data from the sites had been reversed (ie the flow data from one site was being used for the other carriageway and vice versa). Now that the LIP is becoming mature, it has been possible to define an additional test.

This test consists of averaging the data for the morning peak and the evening peak for the 7-day period. By dividing the morning average by the evening average, a diurnal flow ratio can be derived for each flow. These can be compared with upstream and downstream ratios and checked for consistency. If there is access to a major town between the Links that create the ratio, a reversal may occur. By checking for consistency and for flow relative to conurbations, cases of reversed flows can be identified and corrected.

CONCLUSIONS

The LIP has been developed at NTCC to identify inaccurate flow data on a large-scale network. Over a period of about 18 months, it has been developed and refined to the extent where the number of undefined errors is small. Work is currently under way to identify flow data that is reversed vis-à-vis the carriageway that it purports to represent. The results of applying the LIP over this period is that the accuracy has improved dramatically and, because the LIP is also useful for assisting the faults team in the rapid identification of faulty equipment, the availability of flow data has also been improved over and above what would have been expected without it.

ACLNOWLEDGEMENT

I would like to give credit to Phil Tait, Bob Smith and Mark Nicholson who have helped to create the LIP as a practical tool.