

Real-time traffic monitoring using mobile phone data

Problem presented by

Michael Alger

Vodafone Pilotentwicklung GmbH

Problem statement

Vodafone proposes to offer road traffic information services to its mobile phone customers. Such systems usually require some sort of static infrastructure to measure traffic flow (*e.g.* magnetic inductance loops buried in the surface of the road, or cameras combined with number plate recognition technology). In contrast, Vodafone intends to generate traffic flow and velocity data by using the signalling information that is already generated by the standard operation of its mobile phone network. This report investigates the feasibility of this idea by analysing a data set provided by a pilot project on the autobahn network in Southern Germany. The initial aim is to design filters that operate on the mobile phone signalling data and whose output is a low-dimensional description of road traffic conditions. The eventual aim is to develop filters which give short-term forecasts of future traffic conditions.

Study Group contributors

Eddie Wilson (University of Bristol)

Tim Gould (University of Lancaster)

Robert Whittaker (University of Cambridge)

Nenad Radulovic (University of Southern Denmark)

Report prepared by

Eddie Wilson (University of Bristol)

1 Introduction

A mobile phone network infrastructure consists of a set of radio base stations which are wired together with high-speed data links. To each base station there is associated a *cell*, which is the area where the signal from that base station is stronger than those of neighbouring base stations, see Figure 1(a). People carry mobile phones on their travels and consequently the networks track phones so that incoming calls can be routed correctly: the network keeps track of which phone is in which cell so that calls can be connected through the nearest base station. Consequently, the network has (rough) positional information for each phone, and if that information could be stored and correlated, it would be possible (in principle) to track the motion of each phone in a coarse manner. The idea in this report is that in some situations, the flow of phones between neighbouring cells is strongly related to the flow of road traffic on major roads between those cells. Consequently, mobile phone signalling data might be used to detect and even forecast traffic jams.

In practice, the monitoring of the flow of phones is much harder than the above discussion might suggest. The first problem is that of privacy: the data acquired must be anonymous. The second (worse) problem is that of data capture. The key engineering task is to modify the existing data network cheaply so that the required positional information can be captured and stored. The data flows are enormous so it is necessary to focus on particular types of positional information which can be captured easily. Consequently, Vodafone chose not to capture the positional information generated by phones which are turned on but not involved in *active calls*.

The key event which Vodafone have chosen to store is that of a *handover* (HO), when a phone involved in an active call is transferred from one cell to another. An anonymous string which uniquely identifies the call is stored, together with the time of the handover (accurate to milliseconds) and some other types of signal timing information. If the same call lasts for long enough for the mobile phone to cross the whole cell, a HO also occurs out of the cell and the identifier string may be used to establish that it is the same mobile phone involved in the two HOs without directly establishing the identity of the phone.

Electromagnetic field calculations may be used to estimate the location of cell boundaries. Consequently, if it is assumed that a pair of HOs corresponds to a mobile phone travelling on a particular road across the cell, the approximate distance travelled between the HOs is known. Thus such *double-handovers* (DHOs) may be used to estimate the average velocities of phones across the cell, see Figure 1(b), with directional information obtained from the ordering of the HOs.

In the data capture stage of this project, Vodafone analysed 4394km of Bavarian autobahn network which uses 968 cells. EM calculations were performed in order to estimate the cell boundaries and to split the roads into segments belonging to different cells. The large majority of segments were of length 2km or less although the longest was 13.7km. Attention then focused on 110km of busy autobahn South of Munich, consisting of 55 cells and 46 useable road segments. DHO data for these cells was then captured in the period June-September 2003. A separate experiment, which is not considered here, captured data for the suburbs of Munich. The analysis of city flow situations, in

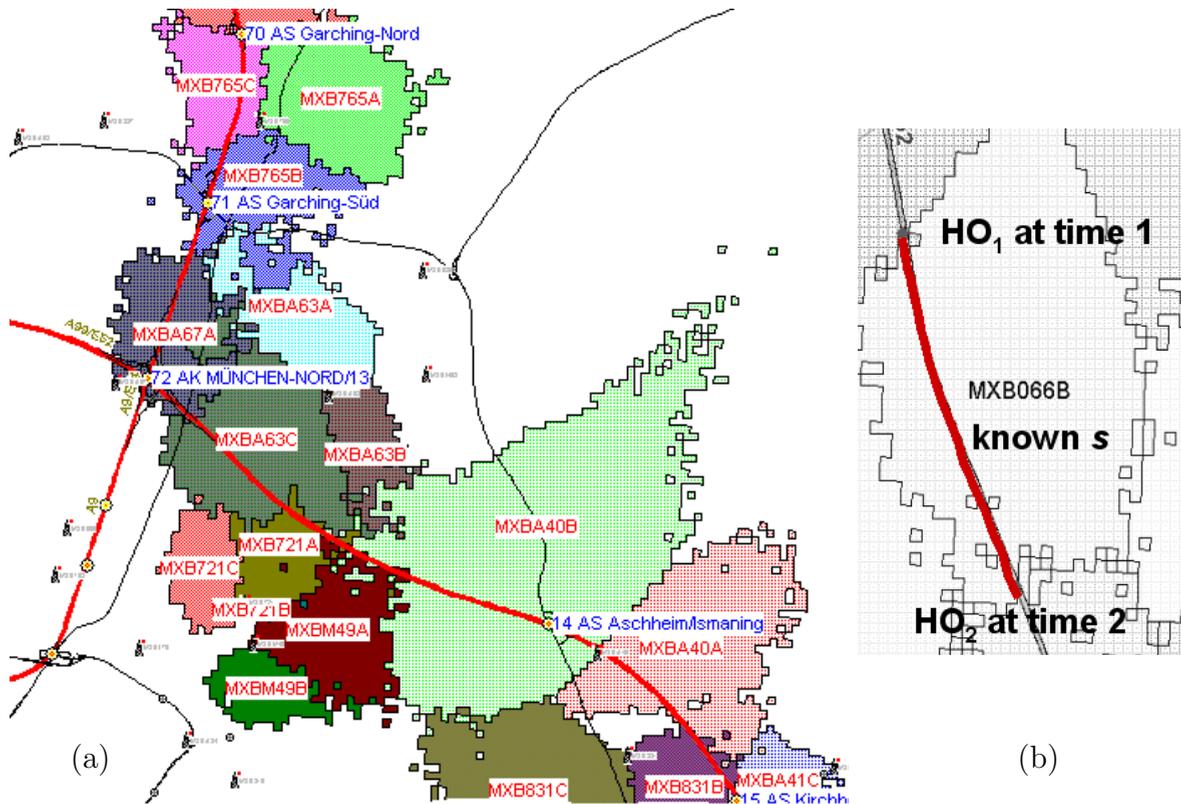


Figure 1: (a) Mobile phone network topology: each cell is shaded in a different colour. The flow of phones between the cells shown here is strongly related to the traffic flow on the two autobahns (coloured red) crossing these cells. However, the situation can be confused by traffic on smaller roads (coloured black). (b) Definition of a double-handover (DHO) event. A vehicle travelling South-bound (down the page) enters cell MXB066B at time t_1 and leaves at time t_2 . Since there is only one road that crosses the three cells shown in the given order, and since the location of cell boundaries can be estimated using EM field calculations, the approximate distance s that the vehicle has travelled, and consequently its velocity, can be estimated.

which there are many roads crossing each cell, seems much more difficult than the rural autobahn setting that we consider here, where the majority of DHOs may be attributed to traffic flow on a single road.

A sample of the data presented at the study group is given in Table 1. It consists of timing information for DHOs and estimates of average vehicle velocities based on the assumption that the DHOs do indeed correspond to travel on a particular autobahn and that the HOs occur at the points predicted by the EM field calculations. We were presented with a very large amount of this data: altogether there were 92 spreadsheets (one for each of the North- and South-bound directions of travel for each of the 46 cells) and each of these typically had in the range of several thousands to several tens of

Date	Time in	a_1	a_2	Transit (ms)	Speed (km/h)
01/07/2003	15:35	0	5	74594	144.59
01/07/2003	15:50	0	5	72078	149.64
01/07/2003	15:55	1	5	380922	28.31
01/07/2003	16:20	0	5	115563	93.33
01/07/2003	16:25	0	5	72468	148.83
01/07/2003	16:40	0	4	82469	130.78
01/07/2003	16:45	1	5	89015	121.17
01/07/2003	17:05	0	5	89984	119.86
01/07/2003	17:05	0	5	103000	104.71
01/07/2003	17:05	0	5	99688	108.19
01/07/2003	17:10	0	5	81234	132.77

Table 1: Sample of South-bound raw data from MXB117A. Here each of the eleven records corresponds to a double-handover (DHO) event that is compatible with a mobile crossing the cell on the autobahn in a South-bound direction. The time in milliseconds to cross the cell is stored, and consequently an average velocity across the cell may be estimated, based on the assumption that the mobile is indeed on the road under consideration and that the handovers (HOs) do occur at the points predicted by the EM calculation. Some extra *advanced timing* information, denoted here by a_1 and a_2 , is also given. Anomalous values of a_1 , a_2 (none here) would tend to indicate that the handovers have not occurred at the usual locations.

thousands of DHO records for the three month period. Altogether, the data set consisted of 1.3 million DHO records.

Although the above numbers may convey the impression that this is a vast data set, our chief finding is in fact that the data set is rather too sparse for the purposes proposed for it. This becomes clear from a glance at Table 1. Even at busy times of day, one may find as much as 30 minutes between DHO records at a single cell site, and gaps of 10 minutes are common. It is known that at busy times, flow rates of 40-50 vehicles per minute per lane of freeway/motorway/autobahn are usual. Thus even a conservative estimate of the flow rate would lead one to conclude from Table 1 that we have velocity measurements for only one in every several hundred vehicles driving across this cell in a South-bound direction. To clarify: we are trying to detect average traffic conditions on a busy road from estimates of the speeds of (say) 1 in every 250 vehicles.

Given that there may be a large variance in the speeds of vehicles at any one time (slow-lane trucks vs fast-lane private cars), this appears to be a hopeless situation. However, Figure 2 depicts graphically the DHO data for a single site over the course of a day, by plotting the estimated speed versus absolute time for each event. In fact, the human eye can detect trends, *e.g.* there appears to be a traffic jam (where speeds drop) shortly

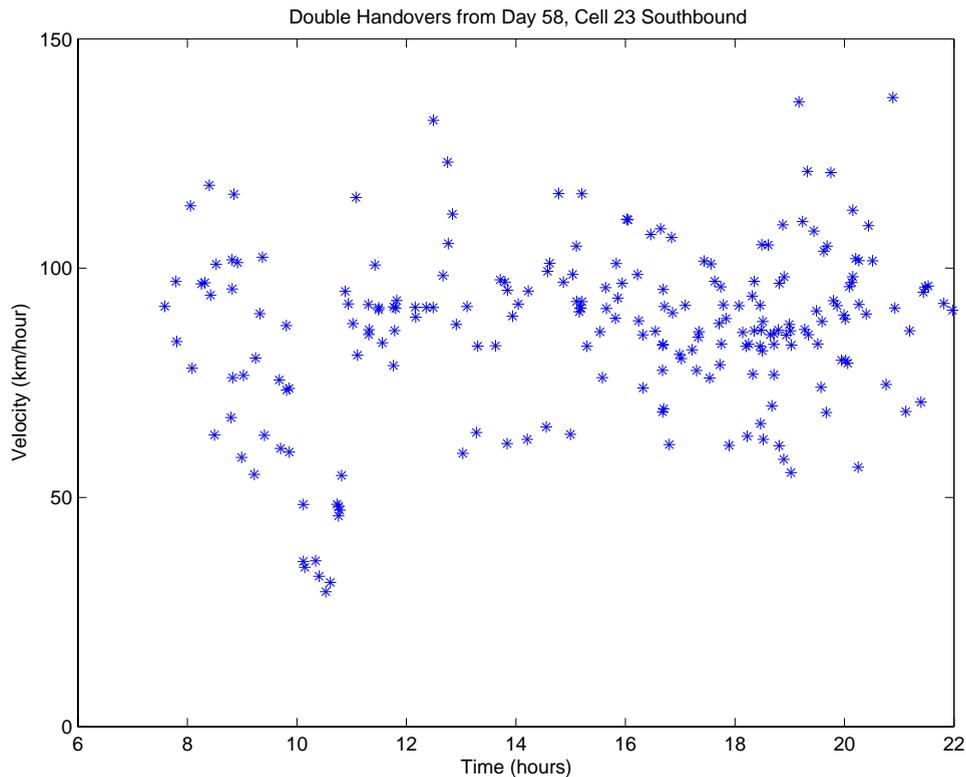


Figure 2: Sixteen hours of South-bound DHO data for a single cell. Each star denotes the absolute time and estimated velocity of a DHO record. Observe the large variance in measured velocities. Note also that there appears to be a traffic jam shortly after 10am.

after 10am in the morning. The goal is therefore to design computational filters which can identify these trends. However, the filters' capabilities are necessarily limited by the sparsity of the data, which is due simply to the fact that not many drivers make calls when travelling at high speed. A secondary factor is that to appear in our data set, those calls must be on the Vodafone network and last long enough for a whole cell to be crossed.

There is a trade-off involving the temporal width of filters: it is tempting to make filters wide so as to capture statistically significant numbers of DHOs. However, this approach leads to a loss of temporal resolution in the traffic description. Conversely narrow filters can have good temporal resolution but may give a spurious description if there are not enough DHOs.

It is clear that there are much more accurate ways of measuring traffic flow than using the DHO data analysed here. Magnetic inductance loops buried at intervals in the surface of the road are one such method, and are used in the UK Highways Agency's Controlled Motorways and Active Traffic Management projects¹. Another method is

¹<http://www.highways.gov.uk/knowledge/tcc/atm/index.htm>

that used by the Trafficmaster² traffic warning system, which uses roadside cameras and licence plate recognition technology to track individual vehicles' progress up the road and so infer typical vehicle speeds. One could also imagine using other technologies such as satellite-based GPS. However, here we have restricted attention to the problem as posed by Vodafone. Recently, a company called Applied Generics³ has announced a traffic information system based on mobile phone signalling data, but it appears from the literature that their system uses more than DHO data alone. Throughout the report we make recommendations about how the data considered here might be augmented in order to help give a more accurate traffic description.

The remainder of the report is as follows. In Section 2 we briefly describe data cleaning and pre-processing steps. Then in Section 3 we describe some of the filter techniques that we used. In Section 4 we describe an alternative strategy for identifying traffic jams which exploits the degree of sparsity of the data set. Finally, we conclude and make recommendations for future work in Section 5.

2 Data cleaning and pre-processing

We began by collapsing the DHO Excel files that Vodafone provided into a single Matlab data structure, focusing on the South-bound data only in order to save time. The resulting data structure could be indexed both in terms of the spatial index i of the cell (labelled $i = 1, 2, 3, \dots, 46$ in the downstream South-bound direction) and in terms of the record number j at a particular cell. This enabled us to write functions which searched, filtered and collapsed the data set in numerous ways including those which projected both temporally and spatially.

We then began a search through the data set to identify cells whose data appeared clean and to eliminate cells with problems. The principal tool that we used was to histogram all of the estimated velocities recorded at a particular cell, and some of the results are shown in Figure 3.

Some of the histograms, such as that shown in Figure 3(c), indicated that the corresponding cells were unusable. A common problem was that both the mean and variance of observed speeds was unreasonably high. Unfortunately we were unable to examine the tails of these distributions since Vodafone had already pre-processed the data by removing records corresponding to an estimated speed greater than 250km/h. In our view, there are two possible factors behind the problems with these cells:

- Other roads crossing the cell in the same way as the autobahn. If these roads had shorter routes through the cell, then small transit times would be possible, which would correspond to high speeds if we assumed that all vehicles travelled on the main (longer) route.

²<http://www.trafficmaster.co.uk>

³<http://www.appliedgenerics.com/index.html>

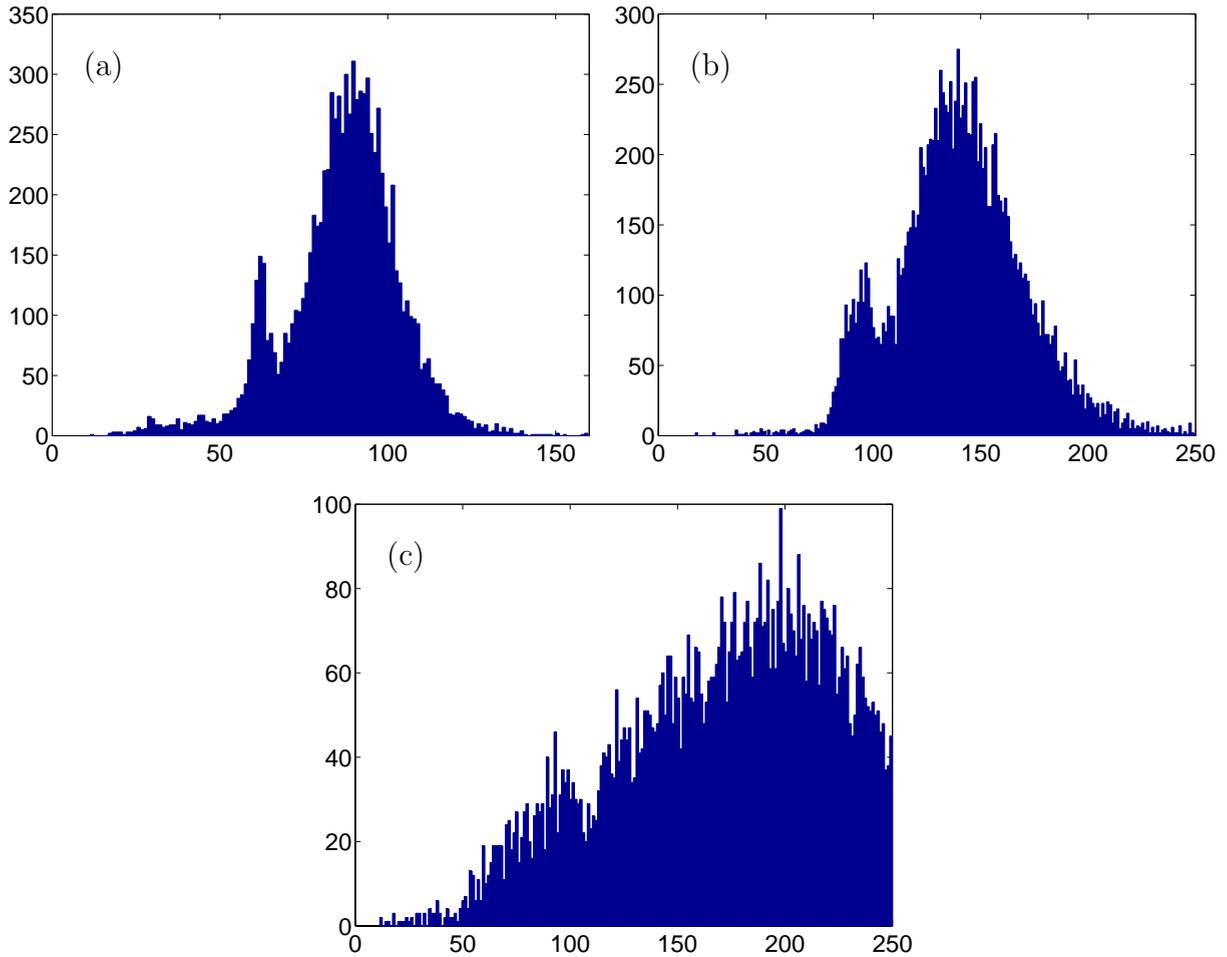


Figure 3: Histograms of vehicle's velocities (frequency vs km/h) for three different cells (South-bound traffic only). In (a) and (b) note the pronounced double-peak due to the mixture of lorries and cars. Assuming the road is level, the sharper, slower peak should occur at 80–85km/h owing to lorry drivers setting their cruise control at the speed limit. The lorry peaks in (a) and (b) are thus respectively too slow and too fast, and this corresponds to the size of the cell being respectively under- and over-estimated by the EM calculation. In (c) the distribution is wholly unreasonable: both the mean and variance are far too high. Unfortunately Vodafone had already removed outliers and thus we were unable to analyse the tail of this distribution for further clues as to what has gone wrong.

- A high degree of variability in the locations where a vehicle is handed over from one cell to the next. This variability might result from topographical features and would be worst in the smaller cells and on the shorter road segments where it would give the greatest proportional change in the distance covered between handovers. This line of reasoning suggests that there is an ideal cell size for traffic flow estimation since small cells are prone to error whereas larger cells have such long transit times that almost no DHOs are recorded.
- **Recommendation.** We suggest that some of the smaller cells should be paired up and something like triple-handover sets generated. The transit times will thus be larger and proportionally less susceptible to error.

Despite the above problems, some cells produced apparently good data, such as those whose histograms are presented in Figures 3(a,b). We expect to see a bimodal pattern with a sharp peak due to lorries at about 80–85km/h and a less well-defined peak due to cars at perhaps (about) 120km/h. Although the figures presented here have the expected pattern, the peaks' positions are not correct. We believe that this is because the EM calculation does not make accurate predictions of cell boundaries and handover points, even when the handover points are repeatable. Thus the estimates of the lengths of road segments, and consequently the estimates of speed, are out by a scale factor. For the cells on which we worked, we cleaned the data by undoing the scale factor to place the lorry peak at precisely 85km/h. However:

- **Recommendation.** The handover points should be investigated further by driving a GPS equipped vehicle along the sections of road in question and identifying the points at which an active call is swapped from one base station to the next. This experiment would need to be repeated several times to give a feel for the variability in the handover points.

One of our original goals was to use DHO data to demonstrate the well-known spatial propagation of traffic jams (upstream, against the flow of traffic). Unfortunately, we were not able to find more than three consecutive cells which produced clean data, and thus we have not been able to give a convincing demonstration of wave propagation.

3 Filter techniques

We have demonstrated that DHOs are rare and that they give velocity estimates for perhaps only one in several hundred vehicles driving through a cell. The variance of vehicles' velocities may be high and consequently we would like to combine velocity estimates of individual DHOs to give a better picture of mean traffic conditions. Further, we would like a description of traffic which is defined for all time and not just at the discrete times at which DHOs occur. We use filters to address both these issues. The two basic types that we used were:

- Moving bin averages, which estimated (continuous time) average velocity by

$$\bar{v}(t) = \frac{\sum_i I_{(t-\delta t, t]}(t_i) v_i}{\sum_i I_{(t-\delta t, t]}(t_i)}.$$

Here I denotes the indicator function and t_i and v_i are respectively the time and velocity estimate of the i th DHO at a given cell.

- Weighted averages, for example with an exponential kernel of the form

$$\bar{v}(t) = \frac{\sum_{i:t_i < t} (v_i e^{-(t-t_i)/\tau})}{\sum_{i:t_i < t} (e^{-(t-t_i)/\tau})}.$$

Note that these filters are one-sided in that they only use past-data to estimate \bar{v} : this restriction is necessary if the filters are to be used on-line. The filters have characteristic temporal width of δt and τ respectively. As we discussed earlier, a large temporal width tends to give smooth averages but a poor rate of response to rapidly changing traffic conditions, whereas a small temporal width often fails, in the moving bin case because there are not a statistically significant number of vehicles in the filter. The temporal width of the filters is thus the key design issue.

We experimented with various widths and found that those in the range 10–30 minutes seemed best. We also found that a hybrid of different filters performed well. In particular, we found that a narrow moving bin average should be used for traffic jam detection whereas a weighted average tends to give a smoother description of free-moving traffic. We also found that it might be useful for filters' widths to be state-dependent in that they should be reduced if the onset of a traffic jam was detected (perhaps defined by the moving bin average dropping below 60km/h). An example of the output from this hybrid approach is given in Figure 4.

For prediction purposes we felt that it might be useful to apply filters which estimate variance, since sharp reductions in average traffic speed seemed to be preceded by reductions in the variance: however, we did not have time to implement this last approach. Further, we believe that a solution to the sparsity problem is to bin data spatially, *i.e.* to group together data from small neighbouring cells. This technique should feature in future work.

We emphasise that the techniques presented here are necessarily ad hoc in that the question as presented is ill-posed. Since we do not know the true state of the autobahn traffic, we cannot say objectively whether our filters produce an accurate description or not.

- **Recommendation.** Provide a complete set of inductance loop data, of the kind shown in Figure 5, corresponding to the entire period of the DHO pilot study. Inductance loop data can be viewed as the ‘absolute truth’ concerning the state of the highway. A proper analysis of filter design could then be performed and optimal parameters found, provided quantitative objectives are also given, *e.g.* what should be the penalty for *false positives* (prediction of traffic jams which do not subsequently occur)?

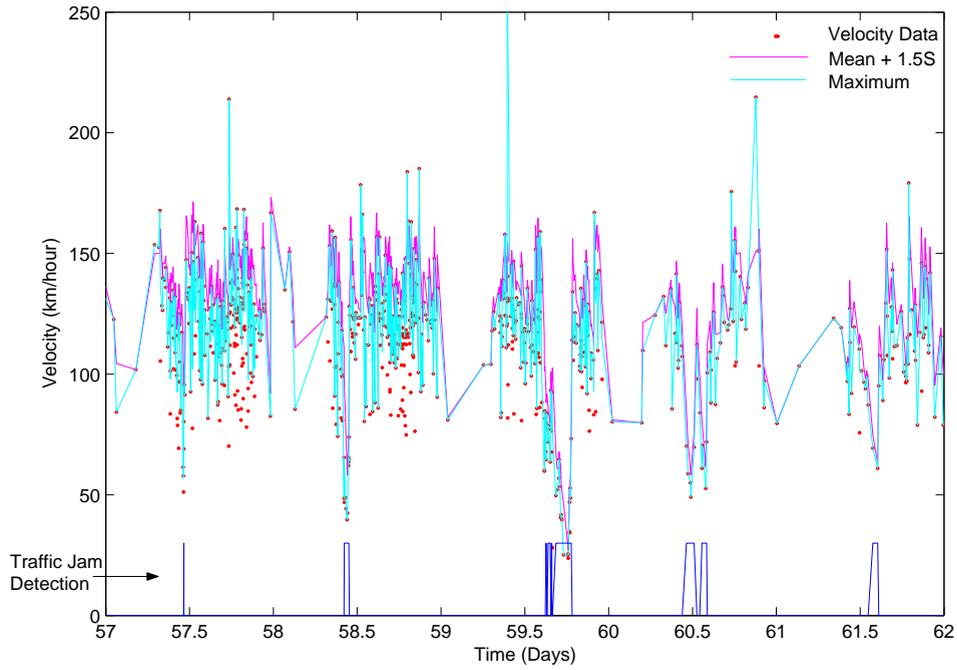


Figure 4: An example of a combination of moving bin and weighted averages that we used both to estimate average speed and to detect traffic jams. Here five working days of data are presented.

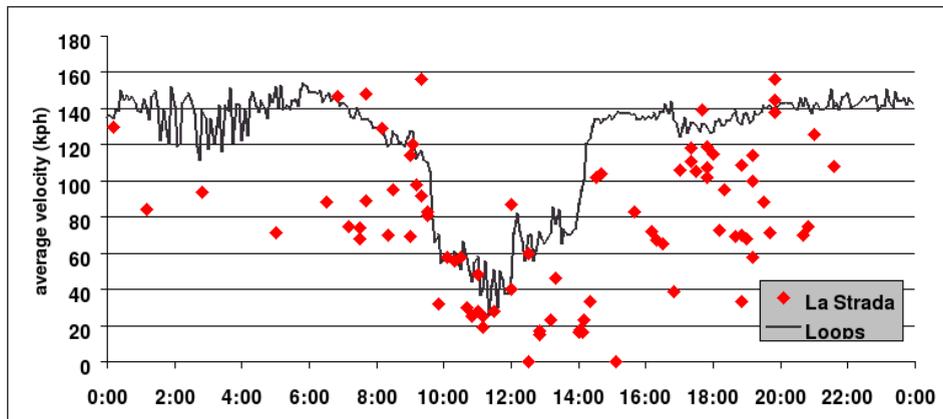


Figure 5: Comparison between inductance loop data and DHO records. We would like to have a complete set of loop data for the definitive testing of filters.

4 Frequency-based techniques

The outputs of all of the filters discussed in the previous section have units velocity, and are calculated by sums of velocity estimates v_i from DHOs. The times t_i at which DHOs

occur are incorporated only via the weights in the sums. The idea in this section is to design methods in which the DHO times t_i themselves play the main role. In particular, we want to use the rate at which DHOs are observed to identify traffic jams. If we assume that all DHOs are caused by vehicles driving on the highway, then the rate depends on the following factors:

- (1) Rate of making or receiving calls, per individual driver.
- (2) Distribution of the length of calls (since calls must be sufficiently long to last across a cell).
- (3) Traffic flow rate in vehicles per unit time. (If all other factors are equal, the rate of DHOs should be proportional to flow.)
- (4) Traffic velocity. (If all other factors are equal, more DHOs will occur in faster traffic since transit times across cells are shorter, and consequently a greater proportion of calls will cause DHOs.)

As a first approximation, one might assume that factors (1) and (2) are independent of the traffic conditions (although conceivably the call rate per driver may be much larger in traffic jams than in free flow conditions). There is then a three-way relation between the DHO rate (observable) and the traffic velocity (unknown) and the traffic flow rate (also unknown), albeit this relation is parametrised linearly by the rate of making / receiving calls and nonlinearly by the distribution of the length of calls, each of which would need to be determined.

- **Recommendation.** Gather data both on the total call volume and the distribution of lengths of calls in highway driving situations, perhaps using rural locations where the majority of calls are made by drivers on the highway.

Established traffic flow theory independently predicts a scalar relation between flow and velocity, so that if call rate and call length statistics can be gathered, one could obtain a scalar relation between either velocity and DHO rate or between flow and DHO rate.

Due to lack of data we were unable to implement this scheme, but to illustrate the principle, we built a crude traffic jam detector based on the DHO rate. For one particular cell and for one particular period during the day (mid to late afternoon) we gathered together all DHO records and analysed the distribution of inter-arrival times between consecutive DHOs, see Figure 4(a). We found (as one would expect, since DHOs should be independent) that inter-arrival times are Poisson distributed, with a rate of one per four minutes for this cell at this time of day. We then applied a hypothesis tester to one particular afternoon's data which identified when the gaps between DHO records were larger than one would expect according to the distribution that had been derived, see Figure 4(b). The hypothesis tester thus crudely indicates when the rate of the DHO process has most likely dropped, either due to a drop in flow rate or due to a drop in velocity. This method could therefore be used as part of a traffic jam detection system, possibly in combination with the techniques discussed in Section 3.

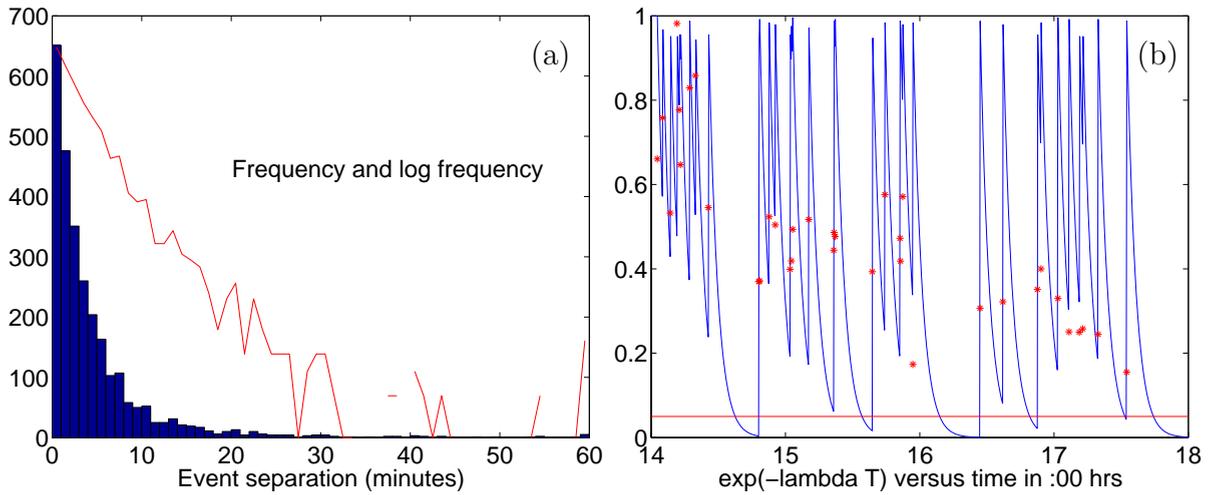


Figure 6: (a) Exponential distribution of times between DHOs, and (b) detection of traffic jams by finding drops in the DHO rate (confidence level set at 95%).

- **Recommendation.** The total call volume and single HO rates should also be collected as part of the dynamic data stream, since otherwise there would be a tendency for the methods described here to detect traffic jams when in fact the road is empty!

5 Conclusions

This report has outlined how double-handover (DHO) data might be used to monitor highway traffic conditions. We did not have sufficient time to consider how predictions of future traffic conditions might be made. Unfortunately, the study at this time is limited by a lack of data, in two different ways:

- Lack of ‘once-only’ data for calibration purposes, *e.g.* refined measurements of cell boundaries, inductance loop data for testing filters, and call rate and distribution data for the calibration of the frequency-based methods.
- Lack of additional sorts of dynamic data: in our opinion DHO data by itself is unlikely to be sufficient. As we described, single handover rates and total call volumes might also be useful.

In conclusion, a great deal more work would be needed to establish the viability of this scheme and one would need to accept that it would be inferior to an inductance loop system in any place where that is installed. However, this work may have future applications in hybrid systems which use (a reduced number of) inductance loops in combination with mobile phone signalling information.