TESTING A RELIABLE IN-VEHICLE NAVIGATION ALGORITHM IN THE FIELD

Ioannis Kaparias[1] and Michael G. H. Bell[2]

1. Centre for Transport Studies, Imperial College London, Skempton Building, London SW7 2BU, UK
   Email: ik00@imperial.ac.uk

2. Centre for Transport Studies, Imperial College London, Skempton Building, London SW7 2BU, UK
   Email: mghbell@imperial.ac.uk

Abstract

This paper presents the results of a field experiment carried out to assess the accuracy and efficiency of a new in-vehicle navigation algorithm, whose aim is to incorporate and consider travel time reliability and route the guided vehicle along uncongested roads, in the absence of real-time traffic information. Using historical travel time profiles deduced from floating vehicle data, the algorithm is implemented in a purpose-developed software tool and tested in the London Congestion Charging Zone. The experiment consists of driving a vehicle along routes computed by the algorithm and comparing the outcome with that of a conventional navigation system installed in a second vehicle. The results indicate that the new algorithm outperforms the conventional system in most cases, thus suggesting that it is a step forward towards a more intelligent navigation system.

1. Introduction

Having started off as luxury accessories installed on high-priced cars, in-vehicle information systems, and in particular satellite navigation systems are gradually becoming more and more common due to recent technological advances. Navigation systems are progressively being recognised, together with the traditional radio, as the most important source of information in the vehicle and this is why the number of users is continuously increasing. In recent years, the market has seen the introduction of portable navigation devices, besides the embedded ones, which are available at affordable prices.

Technological advances in the wireless telecommunications field have also resulted in mobile phones equipped with a GPS-receiver becoming available and thus extending the field even further. Forecasts predict that in the next few years, a large number of drivers will be making frequent use of a navigation device, which indicates that car navigation systems are rapidly becoming everyday consumer goods.

In order to keep up with the growing competition in the in-vehicle navigation field, developers and suppliers of devices, software and services endeavour to improve the quality of their products. Hence, systems containing increasingly more sophisticated functions than simple driving directions are developed. These range from more advanced algorithms aiming to calculate better routes (fastest route, shortest route, avoiding tolls, preferring motorways etc.) making use of higher quality data, to enhanced communication abilities, such that real-time traffic information aiming to inform the driver on congested areas in the road network becomes available, and even to extra functions not directly relevant to navigation, such as weather information.

Focussing on traffic information, car navigation systems can be categorised as 'autonomous' and 'centralised'. The main characteristic of autonomous systems is the fact that they have access to very limited (only traffic congestion reports from the Traffic Message Channel (TMC)) or no real-time traffic information, using only historical data, based on which they offer route guidance to the user. These systems are thus independent of any external services, apart from the GPS satellites and the TMC, and no communication of the vehicle with any other entity is required. On the other hand, centralised systems are based on two-way data exchange between the navigation device and a traffic information centre, offering thus a subscriber-based better quality route guidance service.

In recent research work [1-6] an advanced in-vehicle navigation strategy for an autonomous system has been developed, aiming at taking travel time uncertainty into consideration and supplying alternative reliable routes to the driver, under the condition of no traffic information availability. Reliability is defined as the probability of not encountering abnormal delays and is quantified as the relationship between travel time and travel time variability on a specific road or an entire route [7]. The method also considers time-dependence, i.e. the fact that travel time is not constant but varies with time. The objective of this paper is to further advance this approach and test it in the field, i.e. in a real road network using historical floating vehicle data, comparing the route guidance provided with the output of a conventional existing navigation system.

The paper is structured as follows: the next section presents the background of the study. Then, the in-vehicle navigation algorithm developed is described. After that, the data acquisition and processing procedures required prior to carrying out the field experiment are described, followed by a presentation of the experimental conduct. The results obtained are reported in the following section, while the last section concludes the paper and identifies areas of further research work.

2. Background

This section briefly reviews previous relevant research and describes the background of the work reported here. This includes the time-dependent A* route finding algorithm for road networks, travel time reliability, and travel time estimation and floating vehicle data.

2.1 The time-dependent A* route finding algorithm

Finding the fastest route in a network is one of the most frequently encountered problems in transport engineering. Although various algorithms exist, their performance tends to significantly deteriorate with increasing network size. As in an in-vehicle navigation system route finding is a subroutine that needs to be called very often and due to the fact that the size of transportation networks is usually large, it is of vital importance to have an algorithm, which is efficient enough to produce accurate results in little computation time.

While a large number of methods for finding routes between two points in a network exist and are comprehensively appraised in [8], the leading algorithm for finding the fastest route from an origin to a destination is the A* algorithm [9]. The advantage of A* compared with other route finding algorithms, such as Dijkstra's algorithm [10], is that A* is much more efficient, due to its ability to convert an uninformed search into an informed search by using a heuristic estimate of the travel time from any point of the network to the destination (usually based on the airline distance). The fastest route is found under the condition that the heuristic does not overestimate the actual travel time to the destination.

The concept of A* is summarised as follows; the algorithm holds two lists; the closed list, containing all the nodes of the network that have been expanded; and the open list, containing all the nodes that may be expanded at the next step. At each step, one node is explored and moved from the open list to the closed list, while its successors are placed into the open list. For every node n, $f(n) = g(n) + h(n)$ is calculated, where $g(n)$ is the travel time from the origin to n and $h(n)$ is the estimate of the travel time from n to the destination. The node to be expanded at each step is the node with the lowest $f(n)$ value among the ones in the open list. When $h(n) = 0$ for all nodes n, A* reduces to Dijkstra's algorithm. Moreover, the efficiency of the algorithm increases, the closer $h(n)$ is to the actual travel time.

Using a link-based approach rather than the conventional node-based method, the A* algorithm can be used to find routes for in-vehicle navigation systems, on road networks with special features, such as turn restrictions and one-way roads. Namely, an efficient technique of working with the start- and end-part of each link, and moving between link parts by calculating their f-values each time, has recently been developed [6]. Links are connected through junction movements, such that the end-part of any link is connected to the start-part of any adjacent link by a respective junction movement. The method also accounts for time-dependence of travel time by adopting a modified version of the so-called ‘flow speed model’ technique developed in [11] and implemented in [3], which enables the time-dependent calculation of link travel times based on link speed time intervals and hence not violating the first-in-first-out condition.

2.2 Travel time reliability

Travel time uncertainty has been identified as one of the most important factors affecting travellers’ decisions. Travellers are interested in how long it will take them to reach their destination, but are even more concerned with the reliability of their prediction of total travel time. A wrong travel time prediction results in either an early arrival at the destination or in a delay, none of which are appreciated.

Both empirical and analytical studies have been conducted, all of them demonstrating the importance of travel time uncertainty to travellers [12-17]. Models have been devised to measure the uncertainty encountered by a traveller during his/her trip and express it as reliability. Comprehensive reviews of this topic have also been carried out [18,19].

Much research has focussed on defining adequate measures for quantifying travel time reliability. Most measures use the characteristics of the travel time distribution, such that two types of measures can be identified: those indicating the probability that a link may be unusable due to congestion [20], and those attempting to quantify the amount of delay that may be experienced along a link [21].

Considering in-vehicle navigation, a recently defined measure of the latter category in the context of in-vehicle navigation is one consisting of two indices, namely earliness and lateness, which are derived based on the mean and the 5% and 95% travel times, assuming that travel times are log-normally distributed [7]. Thus, the earliness and lateness reliability of a link l is

$$r_E(l) = exp\left[-½\cdot T_{log}(l) - z_{0.05}\cdot\sqrt{T_{log}(l)}\right] \tag{1}$$

and

$$r_L(l) = exp\left[½\cdot T_{log}(l) - z_{0.05}\cdot\sqrt{T_{log}(l)}\right] \tag{2}$$

with

$$T_{log}(l) = ln\left(1 + \frac{var[t(l)]}{[\bar{t}\,(l)]^2}\right) \tag{3}$$

where $t(l)$ is the travel time on link l, following a log-normal distribution with mean $\bar{t}\,(l)$ and variance $var[t(l)]$, and $z_{0.05}$ = 1.65 is the tail probability of the 90%-confidence level employed. The advantage of these measures is that they are comprehensible by the driver and can be easily converted to expected time gain and delay values, such that the earliest and latest reliable arrival times at the destination can be computed.

2.3 Travel time estimation and floating vehicle data

The estimation and prediction of travel time in an urban network is of vital importance to many transport applications and has therefore been extensively researched in the past and continues to be analysed by many researchers, such as Park and Rilett [22], Park et al [23] and Awasthi et al [24]. This is also the case in in-vehicle navigation, with important contributions to the field being the works of Hoffmann and Janko [25], Boyce et al [26], Sen et al [27-29] and Kerner et al [30], introducing methods, according to which default travel time profiles for navigation systems are estimated using data obtained from probe vehicles (vehicles equipped with a navigation system).

In recent years a new accurate method of estimating travel time, called floating vehicle data (FVD) has been developed, and its concept is based on the transmission of traffic data from a number of vehicles equipped with measuring devices "floating" in traffic, to a so-called service provider, the processing of this data to convert it to traffic information and the re-broadcast of this data to the vehicles, equipped with receivers. In the UK, the first fully operational FVD system has been developed by ITIS Holdings PLC. The ITIS FVD™ system has been collecting data since February 2000, while commercial provision of the data gathered has been in place since 2002. With more than 2.5 million measurements per day, the system is now considered the largest of its kind in the world [31].

The main advantage of a FVD system, as opposed to other existing travel time data collection schemes, is the fact that data can be obtained wherever there is adequate GPS-coverage, whenever the ignition of the vehicle is on. Apart from minimising the cost that would arise from installing conventional measuring devices (such as inductive loop detectors) on many roads, it also enables the monitoring of roads that are not equipped with such devices, such that all but the minor roads can be monitored [32,33]. The fleet of the ITIS FVD™ system consists of about 50,000 private vehicles, business vehicles, trucks and coaches, equipped with 'data collection units', transmitting their locations (longitude and latitude) at pre-determined intervals. Thus, speed is worked out between two consecutive journey points and is assigned using a map-matching procedure to a specific link on a digital road network.

While the initial purpose of the ITIS FVD™ system was to obtain current traffic data and to broadcast it to vehicles, an additional function (or "by-product") that has recently gained value is the creation of a historical travel time database from past measurements, taking advantage of the fact that the system collects data 24 hours a day, 365 days a year. The resulting database of speed measurements enables the calculation of key statistical values for the travel time on each link of the network, such as the mean, the standard deviation and confidence intervals [34].

Concerning in-vehicle navigation, the ITIS FVD™ system has only been used for the provision of real-time traffic data so far. In this study, ITIS FVD™ speed measurements are being utilised for the calculation of mean speed and reliability values following equations (1)-(3) for the links of a test network in London, so as to be used by the reliable dynamic in-vehicle navigation algorithm described in the next section.

3. Methodology

The dynamic in-vehicle navigation algorithm presented in this study makes use of the time-dependent A* algorithm and the reliability measure defined in equations (1)-(3) so as to model congestion risk and suggest a set of alternative equivalently reliable routes to the driver, subject to a number of constraints imposed to ensure acceptability by him/her. The constraints applied relate to maximum route travel time, maximum route length and minimum route earliness and lateness.

The constraints introduced are enforced by the application of link travel time penalties. The main idea lies in performing an initial search in order to find the fastest route without considering reliability, applying travel time penalties to unreliable links (whether in terms of earliness or lateness or both) to exclude them from the search, and running subsequent searches, each time reducing the penalties to re-include the least unreliable links, until the computed route satisfies the constraints imposed. The travel time and length constraints are expressed in proportion to the travel time and length of the fastest route, such that for any computed route p_i, its travel time and length should be $T(p_i) < \beta T(p_0)$ and $\Lambda(p_i) < \zeta\Lambda(p_0)$, where $T(p_0)$ and $\Lambda(p_0)$ are the travel time and length values of the fastest route p_0, and β and ζ are the travel time and length permission parameters respectively. Regarding the minimum earliness and lateness of route p_i, it should be $R_E(p_i) > R_{Emin}$ and $R_L(p_i) > R_{Lmin}$, where R_{Emin} and R_{Lmin} are pre-defined thresholds.

The idea of link penalties is also introduced as an attempt to develop an efficient and acceptable method for finding alternative routes in a network, as it has been proven that strategies suggesting multiple alternative routes to the driver are superior to single-route solutions, because they decrease the risk of the congestion feedback phenomenon and they give him/her the possibility to choose a route according to his/her individual preferences. However, it has to be ensured that the routes suggested are not too similar (are maximally disjoint). This is enforced by additionally applying penalties to links that are already included in a route that was previously computed and accepted, and by applying a fourth constraint relating to maximum route overlapping, such that a newly computed route is only accepted if it does not overlap too much with any previously computed route.

Finally, to ensure termination of the algorithm, a constraint relating to a maximum number of routes to be computed, is introduced. Of course the algorithm terminates earlier if only a smaller number of acceptable routes are found. The travel time penalty, for links having an earliness or lateness value lower than a preset threshold r_{Emin} or r_{Lmin} and for links that are already included in one of the computed alternative routes, is:

$$\Delta t(l) = \alpha^{m} \left(1 - r_E(l) \cdot r_L(l)\right)^{q} W_0 \tag{4}$$

with $0 < \alpha < 1$, m = number of iterations, q = 0 if m = 0 otherwise q = 1, and W_0 = a value large enough to bring about link exclusion.

It should be noted that while the first run of the time-dependent A* algorithm is run in the normal order, i.e. forwards, the subsequent A* runs are performed in reverse order, i.e. from the destination to the origin. The advantage of doing this is that after the first run, the computed travel time from the origin to every visited link can be used as a more exact heuristic estimate for subsequent reverse A* runs. The number of links visited is thus significantly reduced and consequently the efficiency of the algorithm is improved.

In this study, the above algorithm is field tested, so as to demonstrate its advantages regarding the inclusion of reliability in in-vehicle dynamic navigation. Prior to reporting on the results from the experiment carried out, however, the data acquisition and processing procedures involved, as well as the experimental setup and method of conduct are described, including the provision of the test network and traffic data.

4. Data acquisition and processing

The data sources and the processing required prior to the conduction of the field experiments are described here; these include the test network and travel time data. The former is obtained from Planung Transport Verkehr AG (PTV AG) in the form of a PTV VISUM™ map database, extracted from a Navigation Technologies (NAVTEQ™) digital map database, while the latter is derived from speed measurements of the ITIS FVD™ system.

4.1 Test network

The test network selected should meet a number of requirements, so as to be suitable for the purpose of field testing the algorithm. Firstly, it needs to be large enough to offer a route choice, such that a set of alternative equivalently reliable maximally disjoint routes can be computed for each origin-destination pair. Secondly, it needs to have adequate ITIS FVD™ coverage, such that enough data is available for the experiments. Since the ITIS FVD™ system only covers the UK strategic highways network, some travel time data will inevitably have to be simulated; however it is aimed to use as much real data as possible by choosing a network with a high density of main roads, such as a city centre. Thirdly, the network needs to be representative of a random network used for in-vehicle navigation, that is, it needs to be diverse, containing areas with different land uses and road types.

The network that is selected for this study is the original London Congestion Charging Zone (LCCZ), consisting of 6240 nodes and 12891 links and spreading over an area around 7 km long and 5 km wide. As it covers most areas of Central London, it exhibits great diversity of land uses and contains all road types but one (it does not contain a motorway). Also, it has good ITIS FVD™ coverage, because it has a high density of main roads and because it has been the object of many transport planning studies arising from the introduction of the LCCZ scheme. It should be noted, however, that the western extension of the LCCZ is excluded, because the ITIS FVD™ data acquired (see next sub-section) refer to a period prior to the start of its operation, while the experiments are carried out after that.

The test network is supplied in PTV VISUM™ format and is shown on Figure 1. For each network element, a number of attributes are included. For nodes these are the ID number corresponding to the NAVTEQ™ database and the location with X and Y co-ordinates with respect to the OS origin. For links the following are given: ID number in NAVTEQ™ database, ID of start node, ID of end node, co-ordinates of intermediate points indicating the curvature, road type, speed limit and direction. Finally, for junction movements the ID numbers of the three nodes specifying the turn (start node, via node, end node) and the type of the turn (left, right, straight-on or U-turn) are given.

[Figure 1 goes here]

4.2 ITIS FVD™ set collection and processing

ITIS FVD™ measurements are supplied for the conduct of the field experiment in this study. For the test network, data collected by the ITIS FVD™ system during a three month period, between 1 September 2006 and 1 December 2006, is obtained. This corresponds to about 7.6 million speed measurements, each bearing the following attributes:

- Unit ID, representing the unique vehicle ID supplying the measurement
- Latitude of the recorded position of the vehicle during the measurement
- Longitude of the recorded position of the vehicle during the measurement
- Speed of the vehicle in km/h at the time of the measurement
- Date and time of the measurement
- Vehicle type (car, bus, HGV or LGV)
- Map-matched NAVTEQ™ link ID number
- Direction of the NAVTEQ™ link

In order to use the ITIS FVD™ values, a temporal data aggregation procedure by day and time of day is carried out, such that days are categorised as 'weekdays' or 'weekends', while each day is split into 15-minute intervals. Based on the time and day of the observation of an individual measurement, the speed value is assigned to a link and to a specific day and time interval. The individual measurements for a specific link and time interval form a distribution of speeds, from which the average speed can be derived by calculating the space-mean speed (harmonic mean of the distribution), and also the earliness and lateness indices can be computed through the variance of the distribution and equations (1)-(3). The result from the data aggregation is that for each link there are six lists: weekday speed, earliness and lateness, and weekend speed, earliness and lateness, each one of which has 96 values for each one of the 15-minute intervals. When plotted against time of day, the speed, earliness and lateness profiles are obtained.

It should be noted here that a filter is applied to the data aggregation, such that a minimum number of observations are needed in order to be able to calculate unbiased statistical values. As such, whenever less than three speed measurements are present, these are ignored and simulated data (see next sub-section) is used for the link and time interval in question. A filter is also applied to the data aggregation when it comes to abnormally low values. When a value from an observation is too low, it is conjectured that the vehicle stopped while the measurement was taken; hence, observations below 7 km/h are discarded and are not included in mean and reliability calculations.

4.3 Simulation of missing data

For the links and time intervals, for which speed data is not provided by the ITIS FVD^{TM} system (either because the system does not cover them or because not enough data points are available), mean speed and reliability data are simulated. The simulation procedure involves extrapolating data from the links for which data is available to the ones for which data is not available.

Regarding speed data, the chosen value to extrapolate is the fraction of the mean speed to the free-flow speed. Assuming that when no congestion is present the highest speed at which a vehicle would travel on a road is the speed limit, it can be assumed that the free-flow speed corresponds to the speed limit, which is provided for every link in the map database. In order to extrapolate the data, the ratio of the mean speed over the speed limit (speed ratio) is calculated for every link and time interval, for which data is available. For all those links and for each time interval individually, the average speed ratio is computed; the resulting value is assigned to all the remaining links of the network, for which no data is available. Speed values can then be derived for these links and time intervals by multiplying the speed ratio assigned to them with their speed limit value. On the top two graphs of Figure 2, the average speed ratio profiles for all links of the network are shown.

Concerning reliability data, the method of simulation is similar to the one used for mean speeds. Namely, the average earliness and lateness indices are computed over all links, for which data is available, and for each time interval individually. Then, the values computed are assigned to the links, for which data does not exist. In the bottom four graphs of Figure 2, the average earliness and lateness profiles for weekdays and weekends are shown.

[Figure 2 goes here]

5. Experimental conduct

Using ARIAdNE, successor of ICNavS [5], a software tool developed for the purpose of the implementation of the new reliable in-vehicle navigation algorithm, a field experiment is conducted. Two vehicles are used, such that one vehicle is driven along routes suggested by ARIAdNE (termed 'ARIAdNE vehicle'), while the other follows routes suggested by a conventional system (termed 'conventional vehicle'); comparison is made between the two systems as concerns the overall travel time experienced.

The conduct of the experiment is as follows. The two drivers involved (one for each vehicle) meet at the same point to start their trip and agree on a destination, input it in their systems, which calculate route(s) for them, and take note of the departure time (DT). The conventional vehicle driver takes note of the expected time of arrival (ETA) provided by his/her system (CETA), while the ARIAdNE vehicle driver selects a route among the set of reliable routes given and writes down the ETA, the earliest reliable time of arrival and the latest reliable time of arrival of this route, given by ARIAdNE (AETA, AERTA and ALRTA). The two drivers then begin to follow the route directions given by their systems. Upon arrival at the destination, the drivers observe which of the two arrived earlier and write down the conventional actual time of arrival (CATA) and the ARIAdNE actual time of arrival (AATA). It should be noted here, that in order to ensure objectivity and eliminate the possible advantage of one system compared to the other arising from the driving style of the one or the other driver (i.e. if one driver is faster than the other thus always arriving earlier), the drivers swap systems for half the runs of the experiment.

When it comes to the analysis of the results, travel times are computed from the arrival times that have been recorded. The conventional expected travel time (CETT = CETA – DT), the ARIAdNE expected travel time (AETT = AETA – DT), the ARIAdNE earliest reliable travel time (AERTT = AERTA – DT), the ARIAdNE latest reliable travel time (ALRTT = ALRTA – DT), the conventional actual travel time (CATT = CATA – DT) and the ARIAdNE actual travel time (AATT = AETA – DT) are thus derived. Also, the difference in the actual travel time (DATT = CATT – AATT) between the two systems is worked out, such that it can be seen which system "won" in each run of the experiment. When DATT is positive, this means that the ARIAdNE vehicle arrived earlier than the conventional vehicle, while the opposite applies if DATT is negative; a DATT value of zero indicates that the vehicles arrived at the destination at the same time. A plot of DATT against the number of runs is created, so as to visualise the magnitude of DATT.

Since minor incidents on the route may cause short delays (crossing pedestrians, delivery HGVs, parking vehicles etc), values of DATT between -3 and 3 minutes are considered as ties between the two systems. This is applied as a filter to DATT, such that if DATT lies within those values, it is set as equal to zero. The new filtered DATT is called FDATT. The filter is also applied to CATT and AATT, such that the values of FCATT and FAATT are obtained by setting them equal to the average value of CATT and AATT, if DATT lies between -3 and 3 minutes, or by adding/subtracting 1.5 minutes. In order to make DATT and FDATT values comparable between the two vehicles, the ratios DATT/CATT and FDATT/FCATT are computed (DATT/AATT and FDATT/FAATT would be equally appropriate). By ranking these values and by dividing by the total number of runs, cumulative distributions are obtained for both the filtered and the unfiltered data cases, showing the comparison of the results output by the two systems.

6. Results

23 runs of the field experiment are carried out according to the procedure described in the previous section and conclusions are drawn as regards the performance of ARIAdNE compared to the conventional navigation system. A graph displaying the trends of the CATT and AATT values

throughout the experiment, superimposed for comparison purposes, is shown in Figure 3, with the continuous line representing the CATT and the dotted line representing the AATT.

[Figure 3 goes here]

As can be seen from the graph, the ARIAdNE vehicle experiences in general similar or lower travel times than the conventional vehicle (i.e. the dotted line lies below the continuous line), with only two exceptions in the 7th and the 12th run of the experiment, while there are many occasions, where the AATT is significantly lower than the CATT (i.e. an advantage of ARIAdNE), such as in the 2nd, 3rd, 10th, 14th, 15th, 16th, 19th and 22nd run. In total, ARIAdNE outperforms the conventional system in 12 runs of the experiment, in 9 runs it is outperformed by it, while the remaining 2 runs are ties. These numbers mean that for around 61% of the runs ARIAdNE delivers similar or better results than the conventional system. This result is further demonstrated by computing the DATT from the CATT and the AATT values, and by then deriving and plotting the DATT/CATT cumulative distribution, shown in Figure 4 (the DATT/AATT cumulative distribution would also be appropriate and would be very similar). From the plot it can be seen that the resulting curve intersects the vertical axis, representing zero difference, at a cumulative probability of about 0.5, which again suggests that ARIAdNE yields better results than the conventional system. A statistical significance test of this result is described in the Appendix.

[Figure 4 goes here]

As minor incidents on a route may cause short delays, thus giving a potential marginal advantage to the one system or the other, a filter is applied to the DATT values, such that values lying between -3 and 3 minutes are set equal to zero and the result of the corresponding observation is set to ‘tie’. The new filtered DATT is called FDATT and is also applied to CATT and AATT, such that the values of

FCATT and FAATT are obtained. Similarly to the plot of Figure 3, a plot of FCATT and FAATT against the number of runs is shown in Figure 5.

[Figure 5 goes here]

The filtered result of the 23 runs of the double-vehicle experiment is hence the following: ARIAdNE outperforms the conventional system 9 times, there are 11 ties between the two systems, while the conventional vehicle arrives before the ARIAdNE vehicle only 3 times. This means that ARIAdNE arrives either earlier or at the same time with the conventional vehicle in 87% of the runs, thus suggesting again that it has an advantage. The resulting FDATT/FCATT distribution now has a higher frequency of zero values and the corresponding plot intersects the vertical axis at a higher cumulative probability value, namely 0.6. The plot is shown in Figure 6 and depicts the findings reported above. The statistical significance of this is assessed in the Appendix.

[Figure 6 goes here]

Despite the overall pattern of ARIAdNE providing better results than the conventional system, it is interesting to note that it also suffers a fairly significant “defeat” during the experiment, and to be more specific, by 19 minutes. In order to explain this, one must not forget that ARIAdNE’s underlying methodology, the reliable dynamic in-vehicle navigation algorithm introduced earlier in this paper, is a risk-minimising technique. This means that, while unpredictable delays are avoided as much as possible, it cannot be guaranteed that they will not arise at all on a route computed by ARIAdNE, nor can it be guaranteed that the conventional system will compute a route, on which all possible delays will arise. Hence, this value is in fact the “exception confirming the rule”, as is on the other side the highest DATT value obtained, which is a “win” for ARIAdNE by 20 minutes.

[Figure 7 goes here]

On a more detailed level, it can this time be observed that, not only links characterised as unreliable are avoided by ARIAdNE, but also that they are included in the route computed by the conventional system, which does not have access to reliability information, if they have a high speed limit. An example of this is shown in Figure 7, which corresponds to the 14th run of the experiment, in which the second highest DATT value (12 minutes) is recorded, and in which route guidance is sought between Bedford Square in Bloomsbury and Tyers Street in Vauxhall (South London).

While the conventional system guides the driver along the theoretically fastest route (Shaftesbury Avenue – Trafalgar Square – Whitehall – Millbank – Lambeth Bridge – Albert Embankment), shown by the thin line, ARIAdNE suggests a different, in theory longer, route (Kingsway – Aldwych – Waterloo Bridge – Kennington Road – Kennington Lane), shown by the thick line, completely avoiding the circled area containing links characterised as unreliable. The outcome is that congestion is encountered on these particular links, resulting in the conventional vehicle to experience great delay and arrive at the destination much later. As no incidents or other unpredictable events (e.g. road works) causing the congestion are neither reported nor observed in the area, it can be conjectured that the delay is due to the low lateness reliability, not considered by the conventional system but included in the route suggested by ARIAdNE.

The underlying conclusion to the field experiment is that ARIAdNE seems to be able to compute routes that avoid areas where delays are more likely to occur, whereas this is not the case of the conventional system. The result of this is that, for the same origin and destination, a vehicle equipped with ARIAdNE seems to arrive in general earlier than a vehicle equipped with a conventional system.

7. Conclusions

In this paper, a new dynamic in-vehicle navigation algorithm incorporating travel time reliability was field trialled, so as to draw conclusions on its accuracy and efficiency. Implemented in ARIAdNE, a purpose-developed software tool, the algorithm's output was compared with the output of a conventional car navigation system, using traffic data derived from the ITIS FVD™ system. The experiment took place in Central London, and more specifically in the London Congestion Charging Zone. The results pointed out that ARIAdNE outperformed the conventional system in most cases, thus suggesting that the new algorithm is a contribution to the state-of-the-art, and that it may be beneficial to use it in a more advanced navigation system.

The promising results obtained can be used as an incentive for future research in the field of reliable in-vehicle navigation, and more specifically in the further development of the new algorithm. Topics of interest are the extension and testing of the approach to real-time routing (and re-routing) responding to traffic congestion reports, the further implementation of the algorithm in a centralised navigation scheme, aiming at system-wide reliability optimisation, as well as the incorporation of further attributes to be considered, such as individual user preferences and environmental impact.

Acknowledgements

The authors would like to thank BMW AG for sponsoring the research reported here, ITIS Holdings PLC for supplying floating vehicle data, and PTV AG for providing the digital map of the test network. Also, special thanks are given to K. Park and K. Zavitsas for their valuable help in the conduct of the experiment.

References

[1] Chen, Y, Kaparias, I, Bell, MGH and Bogenberger, K, "Reliable autonomous route guidance by a constrained A* search considering intersection delays", The Reliability of Traveling and the Robustness of Transport Systems (Delft, The Netherlands, 2005).

[2] Chen, Y, Bell, MGH and Bogenberger, K, "Reliable pre-trip multi-path planning and dynamic adaptation for a centralized road navigation system", ITSC 2005 - 8th International IEEE Conference on Intelligent Transportation Systems (Vienna, Austria, 2005).

[3] Chen, Y, Bell, MGH, Wang, D and Bogenberger, K, "Risk-averse time-dependent route guidance by constrained dynamic A* search in decentralized system architecture", Transportation Research Record 1944, 2006, pp. 51-57.

[4] Chen, Y and Wang, D, "Optimum path algorithm for in-vehicle navigation system in distributed structure with un-completed dynamic information", Journal of Highway and Transportation Research and Development 2, 2007, pp. 70-75.

[5] Kaparias, I, Bell, MGH, Chen, Y and Bogenberger, K, "ICNavS: A tool of reliable dynamic route guidance", IET Intelligent Transport Systems 1, 2007, pp. 225-233.

[6] Kaparias, I, Bell, MGH, Bogenberger, K and Chen, Y, "An approach to time-dependence and reliability in dynamic route guidance", Transportation Research Record 2039, 2007, pp. 32-41.

[7] Kaparias, I, Bell, MGH and Belzner, H, "A new measure of travel time reliability for in-vehicle navigation systems", Journal of Intelligent Transportation Systems 12, 2008, pp. 202-211.

[8] Ahuja RK, Magnanti TL and Orlin JB, "Network flows: Theory, Algorithms and Applications", Prentice Hall, 1993.

[9] Hart, PE, Nilsson, NJ and Raphael, B, "A formal basis for the heuristic determination of minimum cost paths", IEEE Transactions on Systems, Science and Cybernetics 4, 1968, pp. 100-107.

[10] Dijkstra, EW, "A note on two problems in connexion with graphs", Numerische Mathematik 1, 1959, pp. 269-271.

[11] Sung, K, Bell, MGH, Seong, M and Park, S, "Shortest paths in a network with time-dependent flow speeds", European Journal of Operational Research 121, 2000, pp. 32-39.

[12] Jackson, WB and Jucker, JV, "An empirical study of travel time variability and travel choice behaviour", Transportation Science 16, 1981, pp. 460-475.

[13] Bates, J, Black, I, Fearon, J, Gilliam, C and Porter, S, "Supply models for use in modelling the variability of journey times on the highway network", AET - Proceedings of the European Transport Conference 2002 (Cambridge, UK, 2002).

[14] Liu, HX, Recker, W and Chen, A, "Uncovering the contribution of travel time reliability to dynamic route choice using real-time loop data", Transportation Research A 38, 2004, pp. 435-453.

[15] Lam, TC and Small, KA, "The value of time and reliability: measurement from a value pricing experiment", Transportation Research E 37, 2001, pp. 231-251.

[16] Abdel-Aty, A, Kitamura, R and Jovanis, PP, "Investigating the effect of travel time variability on route choice using repeated measurement stated preference data", Transportation Research Record 1493, 1995, pp. 39-45.

[17] Noland, RB and Small, KA, "Travel-time uncertainty, departure time choice, and the cost of morning commutes", Transportation Research Record 1493, 1995, pp. 150-158.

[18] Bates, J, Polak, JW, Jones, P and Cook, A, "The valuation of reliability for personal travel", Transportation Research E 37, 2001, pp. 191-229.

[19] Noland, RB and Polak, JW, "Travel time variability: a review of theoretical and empirical issues", Transport Reviews 22, 2002, pp. 39-54.

[20] Bell, MGH and Iida, Y, "Network reliability" in Transportation network analysis, Wiley and Sons, Chichester, 1997.

[21] Lomax, T., Schrank, D., Turner, S., and Margiotta, R., "Selecting travel reliability measures", Texas Transportation Institute, 474360-1, 2003.

[22] Park, D and Rilett, LR, "Forecasting multiple-period freeway link travel times using modular neural networks", Transportation Research Record 1617, 1998, pp. 163-170.

[23] Park, D, Rilett, LR and Han, G, "Spectral basis neural networks for real-time travel time forecasting", ASCE Journal of Transportation Engineering 125, 1999, pp. 515-523.

[24] Awasthi, A., Lechevallier, Y., Parent, M., and Proth, J-M., "Using data analysis to approximate fastest paths on urban networks", Institut National de Recherche en Informatique et Automatique, 4961, 2003.

[25] Hoffmann, G and Janko, J, "Travel times as a basic part of the LISB guidance strategy", 3rd IEEE International Conference on Road Traffic Control (London, UK, 1990).

[26] Boyce, D, Rouphail, N and Kirson, A, "Estimation and measurement of link travel times in the ADVANCE project", Proceedings of the IEEE-IEE Vehicle Navigation and Information Conference 1993 (Ottawa, Canada, 1993).

[27] Sen, A and Thakuriah, P, "Estimation of static travel times in a dynamic route guidance system", Mathematical Computer Modelling 22, 1995, pp. 83-101.

[28] Sen, A, Sööt, S, Thakuriah, P and Condie, H, "Estimation of static travel times in a dynamic route guidance system - II", Mathematical Computer Modelling 27, 1998, pp. 67-85.

[29] Sen, A, Thakuriah, P, Zhu, X-Q and Karr, A, "Variances of link travel time estimates: implications for optimal routes", International Transactions in Operational Research 6, 1999, pp. 75-87.

[30] Kerner, BS, Rehborn, H, Aleksic, M and Haug, A, "Traffic prediction systems in vehicles", ITSC 2005 - 8th International IEEE Conference on Intelligent Transportation Systems (2005).

[31] Cowan, K and Gates, G, "Floating vehicle data system - A smart move", Smart Moving (Birmingham, 2003).

[32] Ilgaz, RD, Gates, G and James, L, "Floating vehicle data for network management: A pilot study", 9th World Congress on Intelligent Transportation Systems (Chicago, 2002).

[33] Simmons, N, Gates, G and Burr, J, "Commercial applications arising from a floating vehicle data system in Europe", 9th World Congress on Intelligent Transportation Systems (Chicago, 2002).

[34] Storey, B and Holtom, R, "The use of historic GPS data in transport and traffic monitoring", Traffic Engineering and Control 44, 2003, pp. 376-379.

Appendix: Statistical significance tests

Two tests are carried out to assess the statistical significance of the difference in actual travel time of ARIAdNE and the conventional system, for both the unfiltered and filtered results.

To test the statistical significance of the unfiltered result a one-tailed t-test is used. Namely, the sample size is n = 23, the sample mean is calculated as $\bar{DATT}$ = 1.739 and the sample standard deviation is worked out as s_{DATT} = 8.225. The null and alternative hypotheses are thus formulated as H_0: μ_{DATT} = 0 and H_1: $\mu_{DATT} > 0$, where μ_{DATT} is the mean DATT. Based on these, the t-test value is:

$$t = \frac{\bar{DATT} - 0}{s_{DATT}\sqrt{n}} = \frac{1.739}{8.225\sqrt{23}} = 1.014$$

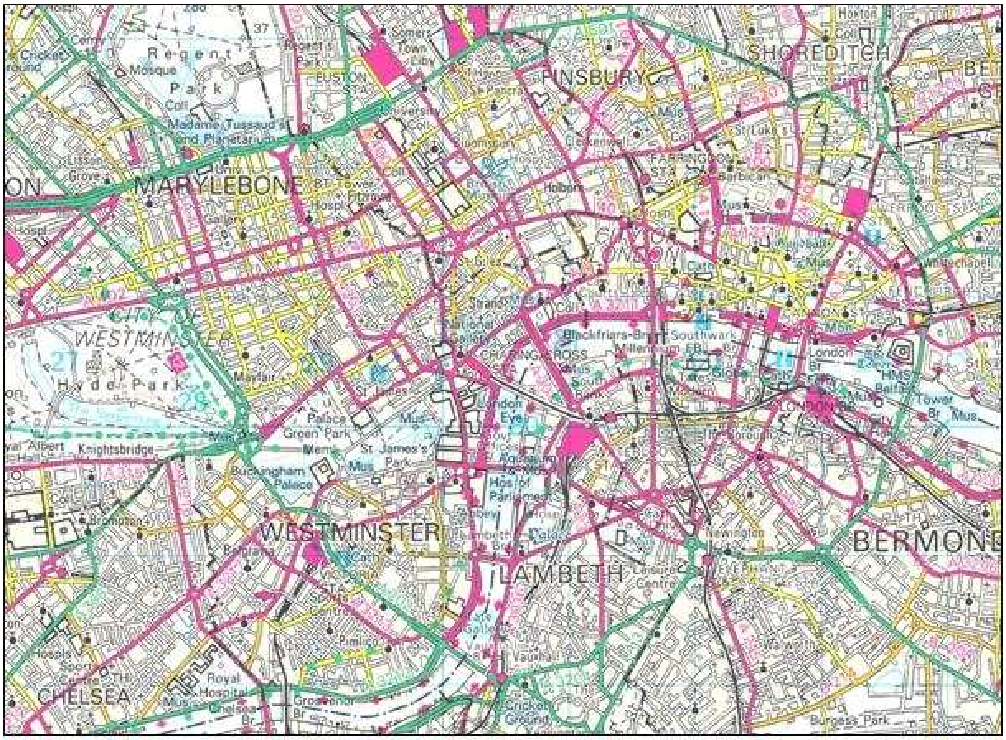
For n-1 degrees of freedom, i.e. 22, the corresponding value of Student's t-distribution at a significance level of 0.05 is $t_{22,0.05}$ = 1.717, and since $|t| < t_{22,0.05}$, the null hypothesis cannot be rejected. Thus, ARIAdNE's advantage over the conventional system is not statistically significant at the 0.05 level. However, taking a significance level of 0.2, the corresponding value of Student's t-distribution is $t_{22,0.2}$ = 0.859, and since $|t| > t_{22,0.2}$, the null hypothesis is rejected and the alternative hypothesis is accepted. As such, ARIAdNE's advantage over the conventional system is statistically significant at the 0.2 level.

To assess the statistical significance of the filtered result, a t-test is used in the same way as in the unfiltered result assessment. Namely, for the filtered result the sample size is n = 23, the sample mean is $\bar{FDATT}$ = 1.261 and the sample standard deviation is s_{FDATT} = 6.181. The null and alternative hypotheses are again formulated as H_0: μ_{FDATT} = 0 and H_1: $\mu_{FDATT} > 0$. Based on these, the t-test value is:

$$t = \frac{\bar{FDATT} - 0}{s_{FDATT}\sqrt{n}} = \frac{1.261}{6.181\sqrt{23}} = 0.978$$

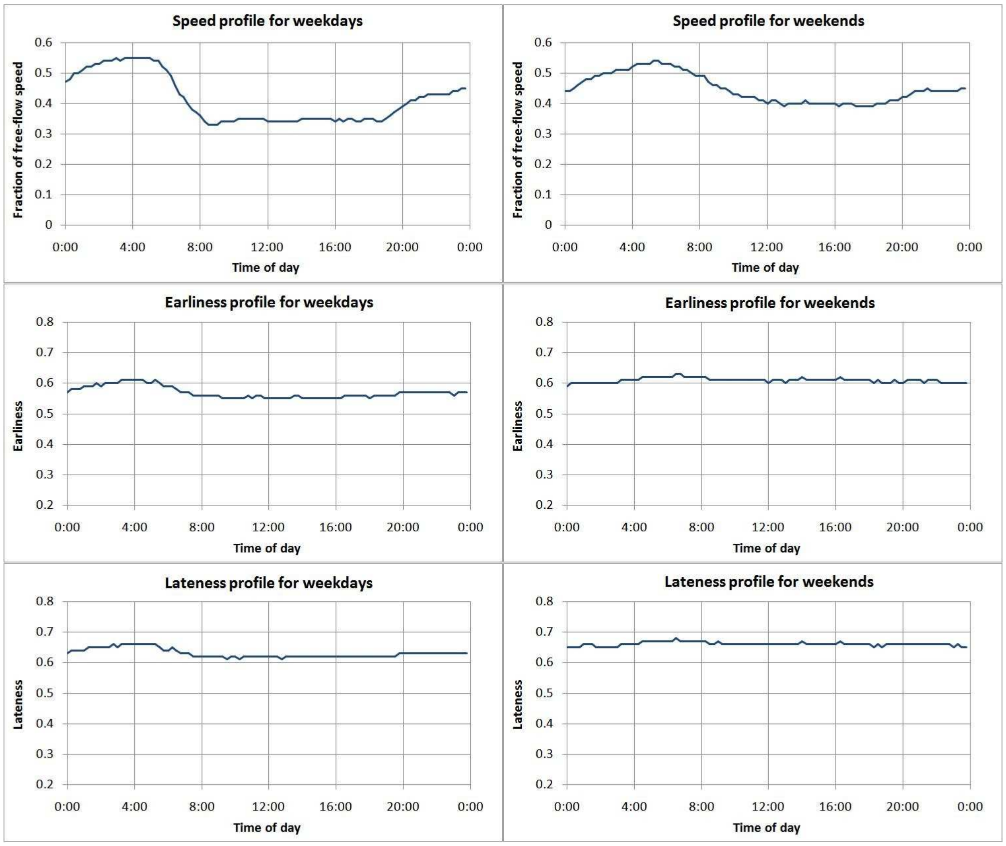
As in the case of the unfiltered result, for n-1=22 degrees of freedom, the corresponding value of Student's t-distribution at a significance level of 0.05 is $t_{22,0.05}$ = 1.717, and since $|t| < t_{22,0.05}$, the null hypothesis cannot be rejected. Thus, the outcome is that, similarly to the unfiltered result, ARIAdNE's advantage over the conventional system is not statistically significant at the 0.05 level. However, taking a significance level of 0.2, the corresponding value of Student's t-distribution is $t_{22,0.2}$ = 0.859, and since $|t| > t_{22,0.2}$, the null hypothesis is rejected and the alternative hypothesis is accepted. As such, even for the filtered result, ARIAdNE's advantage over the conventional system is statistically significant at the 0.2 level.

List of figure captions

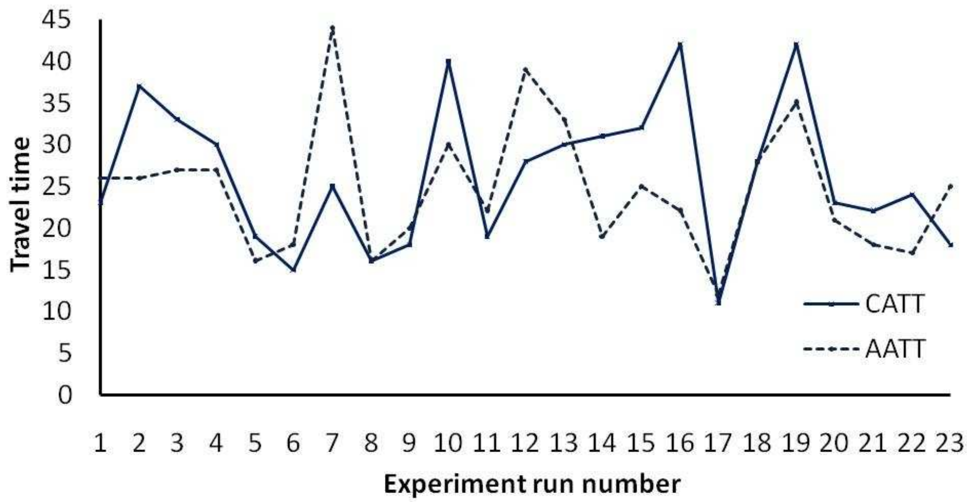


© Crown copyright/database right 2006. An Ordnance Survey/EDINA supplied service

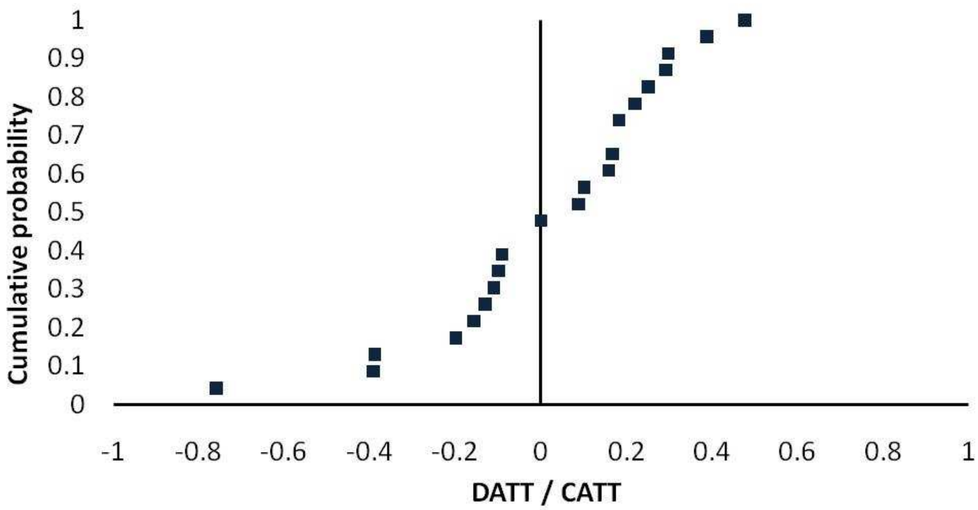
The test network
160x123mm (96 x 96 DPI)



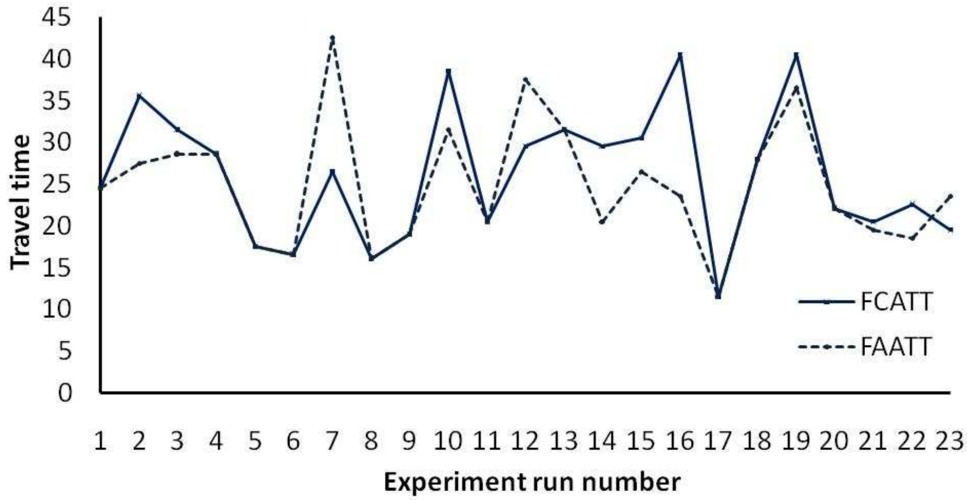
Simulated speed and reliability data: average speed, earliness and lateness profiles
400x336mm (96 x 96 DPI)



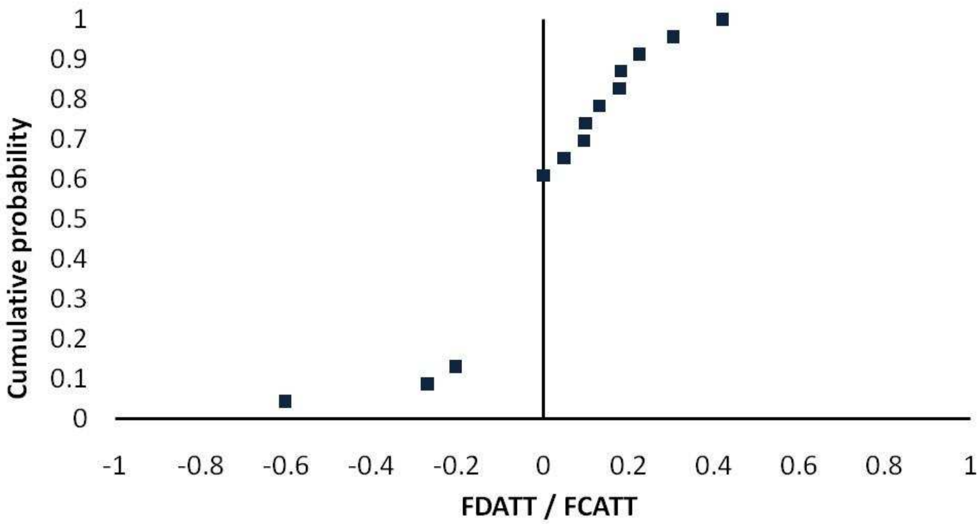
CATT and AATT values for each run of the experiment
257x135mm (96 x 96 DPI)



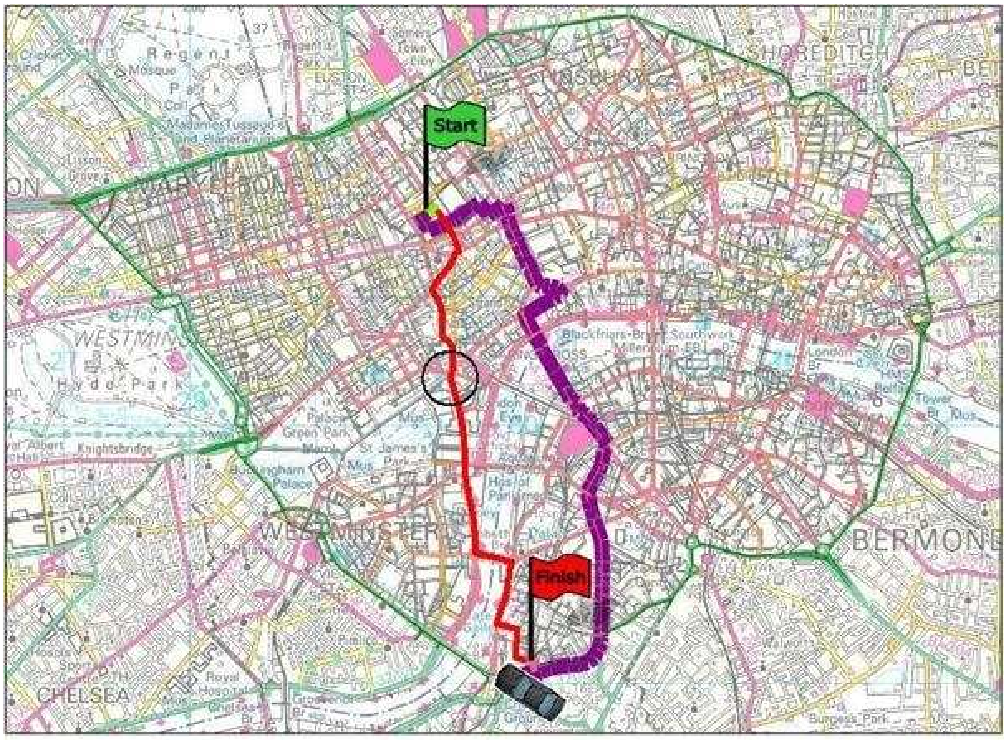
Cumulative distribution plot of DATT/CATT
257x142mm (96 x 96 DPI)



FCATT and FAATT values for each run of the experiment
257x135mm (96 x 96 DPI)



Cumulative distribution plot of FDATT/FCATT
257x142mm (96 x 96 DPI)



The advantage of the ARIAdNE route compared with the conventional route
159x122mm (96 x 96 DPI)