The process of Measuring and Validating Bus Stop Location Data for the iBus project at London Buses

Dr Steve Robinson, London Buses, Transport for London, UK, steve.robinson@tfl-buses.co.uk
Felipe Radillo, London Buses, Transport for London, UK, felipe.radillo@tfl-buses.co.uk

Abstract

London Buses (part of Transport for London) is in the process of installing a new GBP120m Automatic Vehicle Location (AVL) solution called iBus across all 8,400 buses and 700 bus routes in London. This system requires accurate bus-stop location data. This technical note discusses various approaches to collect this data, and presents research undertaken at London Buses to measure the performance of each one. The optimal approach to collecting bus-stop location data is to plot the bus-stops on an electronic map, on-site, using a lap-top. In addition a novel approach for validating this stop-location data is presented. This validation approach has proved to be very useful in identifying erroneous network data.

1 Introduction

London Buses (part of Transport for London, TfL) is in the process of installing a new GBP120m Automatic Vehicle Location (AVL) solution called iBus across all 8,400 buses and 700 bus routes[1] in London. iBus will turn into reality many of the advantages that Intelligent Transportation Systems (ITS) can provide the bus industry. For example it will allow passengers to use mobile phones to find out real-time service information, provide passengers with on-bus location information, and provide London Buses management with automated reports to allow them to measure the performance of the bus network and the private bus operators which run the routes.

One important aspect to the success of iBus is to provide high quality data to the system. Many ITS projects are guilty of overlooking the importance of data quality, resulting in severe degradation of the resulting system (Turochy & Smith, 2000). London Buses have thus implemented new processes to ensure that iBus receives data of sufficiently high quality. In particular, it is necessary for London Buses to provide the following data to iBus:

- **Item #1:** The longitude and latitude of all 19,000 bus-stops in London
- **Item #2:** The heading that the bus is travelling in when it approaches each bus-stop (expressed in degrees, i.e. 0 - 359° where 0° = North, 90° = East, etc…).
- **Item #3:** The distance that the bus travels between adjacent bus-stops on a bus route. There are over 40,000 unique stop-to-stop links in London's network.

The performance of the iBus systems will be significantly degraded if any of the data in the above categories is wrong. Due to the difficulties in measuring items #2 and item #3 these are

[1] NB. A London Bus 'Route' is equivalent to a Transmodel 'Line'

calculated automatically from knowledge of the stop location (item #1) and knowledge of the routes the buses traverse. Therefore it is vital for the success of the iBus project for London Buses to accurately know where all 19,000 bus-stops in London are located. This technical note thus discusses the method London Buses have implemented to measure and validate this. This will be of use to many other major network authorities. Section 2 briefly describes the problem of measuring the location of a bus-stop in an urban environment. Section 3 discusses an experiment that was carried out by London Buses to identify the best approach to measuring all 18,000 bus-stops in London. Section 4 describes the method London Buses are using to validate the stop location data. Section 5 presents the conclusions and describes future work.

2 Problems with measuring bus-stop locations

The task of measuring the locations of bus-stops may at first appear trivial since GPS does not now have selective availability (SA) degradation. However, GPS devices have limitations, particularly in urban environments. This has been proved by previous research in London by Patchett et al (2005). In addition the location of bus-stops frequently changes – there are in the region of 500 new or relocated bus-stops in London each year. It was thus necessary to identify the best approach to measure the location of bus-stops in London. There are many ways of obtaining the location of a bus-stop. These are outlined below in Table 1.

Method	Advantages	Disadvantages
Standalone GPS device (including DGPS)	• Cheap • Relatively easy to learn and can be used by normal staff.	• Will not work in urban canyons (i.e. near tall buildings or trees) • Prone to errors due to multi-pathing
Automatic Vehicle Location sensor (i.e. combined GPS, dead reckoning and map matching)	• Easy to collect data	• Expensive to procure • The bus-stop is defined as the flag displaying all the services. This is mounted on the sidewalk where vehicles cannot access
Plotting location on an electronic map (i.e. a person takes a laptop / tablet pc to the bus-stop and plots its location on an electronic map)	• Relatively cheap • Relatively easy to learn and can be used by normal staff.	• Requires up-to-date and accurate electronic maps • Requires licenses for GIS software • Requires training in the use of the software and map reading
Plotting of textual location data. (i.e. given a textual description of the bus-stop, plot this on a map)	• Data is already available • Easy to learn and can be done by normal staff.	• Prone to outliers since the textual description of the location of the bus-stop may be inaccurate or vague
Aerial Photography		• Bus-stops are likely to be occluded by trees or tall buildings • Photos may be out-of-date • Since person is not on site it may not be possible to identify the asset

		number of the feature.
Advanced Surveying Techniques (e.g. theodolite, total station)	• Very accurate	• Very expensive • Time consuming and unable to cope with frequent changes in location of bus-stops • Requires specialists to perform survey

Table 1 – Advantages and disadvantages of existing methods of measuring stop-locations

3 London Buses approach to measuring bus-stop locations

In order to assess which methodology London Buses should undertake to measure the location of all 18,000 bus-stops, a pilot exercise was undertaken. It should be noted that London Buses describes the location of a bus-stop by using the *Ordnance Survey*[2] (OS) Eastings and Northings coordinate reference system. The pilot evaluated the following approaches:

- **Textual Description**: This approach consisted in using the description of the physical location of a bus stop (for example: 5.3 metres north from the entrance of 172 Buckingham Palace Road; 28.6 metres south from the central line of Elizabeth Street) to pin point the coordinates using a GIS application.
- **GPS**: This approach used a standard GPS receiver (the GPS used on this pilot was a Garmin eTrex GPS receiver) to gather the Eastings and Northing coordinates of the Bus Stops.
- **Tablet PC with Electronic Map software**: Tablet PCs (Pentium 4, 1.1 GHz, 520 MB in RAM) were loaded with a mapping application.
- **Laptop with GUI**: Laptops (with Pentium 4, 1.6 GHz, 1 GB in RAM) were loaded with a mapping application.

The last two approaches require the normal staff to be computer literate and trained. To assess the affect of this two separate people were asked to do this approach. Person A was not very experienced with using computers whereas person B was.

In order to test the accuracy of each method, two routes (route 2 from Baker Street to West Norwood, and 68 from West Norwood to Euston Station) which go through central London were professionally surveyed to provide the 'true' location of the stops on these routes. The stop locations were then measured using each of the 4 methods described above (the last two being repeated with different people). Person A collected the textual description data, and person B collected the GPS data. Data was collected by driving along the route in a London Buses van.

The accuracy of each approach could then be determined by calculating the Euclidean distance (straight line distance) between the measured and the 'true' location of the bus-stop. A summary of the accuracy of each methodology in measuring the 39 bus-stops on route 2 is given in Table 2.

[2] ***Ordnance Survey*** is Great Britain's national mapping agency, providing the most accurate and up-to-date geographic data, relied on by government, business and the public for more than 200 years.

Method	Mean Error, m	Median Error, m	95% Error, m
GPS	16.45	9.61	31.53
Textual Description	6.24	1.22	11.48
Laptop with GUI: Person A	6.31	1.13	28.22
Laptop with GUI: Person B	2.68	1.19	6.45
Tablet PC with GUI: Person A	3.97	1.53	11.44
Tablet PC with GUI: Person B	3.75	1.72	7.70

Table 2 – Comparison between location measurement techniques, of the errors in measuring the location of bus-stops on the route 2.

From Table 2 the following conclusions could be drawn:

- The least accurate method was the GPS device, highlighting the limitations of this technology in heavily built up urban environments (although it is noted that newer more accurate devices are being introduced and that from 2008 Galileo satellites may improve this method's performance)
- Textual descriptions performed very well when the building numbers (which can be seen on the electronic map) were used to describe the stop location. However, in their absence the performance of this method was poor.
- There was little difference in performance when using a laptop or tablet pc. However, both person A and B preferred using the laptop.
- The computer literate person (person B) was less prone to outliers than the less experienced computer user. This suggests that training of all staff is necessary when using a laptop plotting approach. Anecdotal evidence suggests that the performance of person A has improved since training and experience of using the system.

Results from the 42 bus-stops on the route 68 showed similar results as above and are therefore not reproduced here.

Following the results from the pilot exercise, London Buses chose to measure all 19,000 bus-stops in London by plotting the location on electronic maps using Laptops. An example of the GUI uses to collect the bus-stop data is shown in Figure 1. This clearly show the high quality maps (which include building numbers) available from the Ordnance Survey, and on whose accuracy this technique depends on.

All 19,000 bus-stops have now been surveyed by this approach and London Buses are very pleased with the results. Nevertheless, it is always important to check the quality of the data. The following section will describe the validation process that has been introduced to ensure the quality of data.

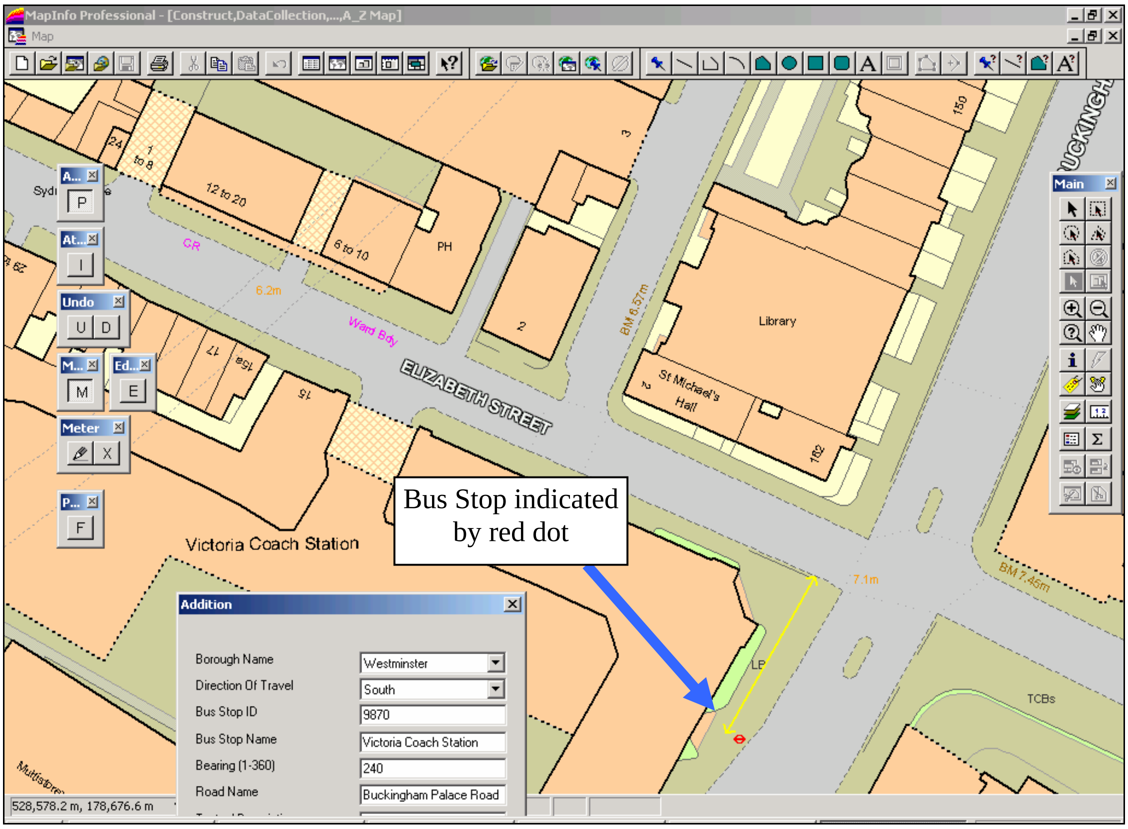

Figure 1 - Screenshot illustrating capture of bus-stop around Victoria Coach Station in central London using GIS loaded on the laptop. The high quality and accurate Electronic maps used are Ordnance Survey MasterMaps.

4 Validating bus-stop location data

4.1 'Stop-snapping'

In addition to measuring stop-locations, London Buses also have to store the route geometry and the stop-sequence for each direction of the route (called the 'route-run'). The route geometry describes the sequence of roads that the bus traverses 'in service' in each direction. It is stored as Ordnance Survey ITN links – i.e. the geometry of the road centre line. The stop-sequence of a route-run describes the sequence of bus-stops that a bus visits 'in service'.

In order to validate the bus-stop location data, London Buses uses the principle that each bus-stop in the stop-sequence of a route-run must be located very close (say within 10m) to the road centre line describing the route geometry data for the route-run.

The process of doing this is called 'stop-snapping'. Essentially stop-snapping places the bus-stop on a line describing the route of the bus. An example of this can be seen in Figure 2. In this diagram the red dot represents the measured location of the bus-stop by the side of the road. It has been 'snapped' to the location represented by the orange dot. This 'snapped location' lies on the road centre-line used to describe the route geometry. The 'snap distance' is simply the distance between the red and orange dot (i.e. the measured and 'snapped' location).

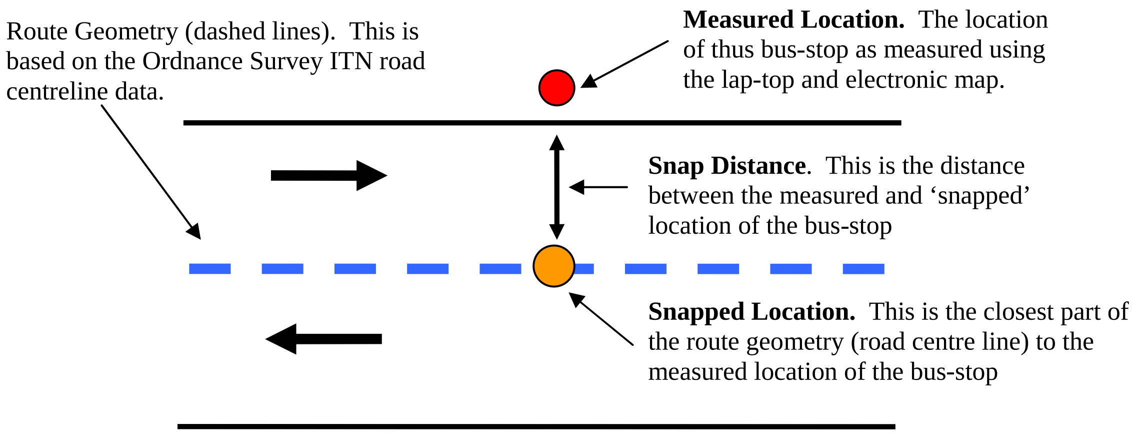

Figure 2 – Snapping a bus-stop

4.2 Process of 'stop-snapping'

The process of 'stop-snapping' is fully automated in a London Buses software called BusNet (Ravikanth et al, 2005). It uses the functionality of Oracle Spatial to achieve this.

- **Step 1:** BusNet identifies ALL route-runs which use the bus-stop
- **Step 2:** BusNet looks at all the route-geometries of all the route-runs which use the bus-stop and identifies all the common ITN links (i.e. the route geometries shared by all routes which use the bus-stop).
- **Step 3:** BusNet finds the closest point from the bus-stop to any common ITN link identified in the previous step.
- **Step 4:** The bus-stop is snapped to this point and related to the ITN link.
- **Step 5:** The geo-status is determined (see Table 3). In addition the distance between the measured stop-location and the point on an ITN link where the link has been snapped to is measured. This metric is termed the 'Snap Distance'.

Geo-status	**Description**
Snapped	The bus-stop has been satisfactorily snapped to an ITN link, based on routes that use the bus-stop.
Unsnapped	No ITN link has been found to which the bus-stop can be snapped.
Remote	The bus-stop has been 'snapped' to an ITN link (based on routes that use the bus-stop) but has a 'snap distance' greater than 100m
Misplaced	The bus-stop has been snapped to an ITN link (based on routes that use the bus-stop) but there is a closer ITN link.
Routeing	The bus-stop has been snapped to an ITN link (based on routes that use the bus-stop). However, not all the routes using the bus-stop use the link.

Table 3 – Geo-status of stop snapping

Bus-Stops with a geo-status that is not 'snapped' or have a 'snap-distance' greater than 15m indicate that there is a problem with either the bus-stop location, stop-sequence, or the route geometry. It is therefore a very powerful technique which London Buses have used to identify and cleanse errors in the data that will be sent to iBus. Without this technique it would have been very difficult to provide the quality of data required for the iBus ITS system.

4.3 Example of the use of 'stop-snapping'

As an example of the usefulness of 'stop-snapping', consider bus-stop 673 used by around 10 bus-routes near Euston Square. It was found to have a status of 'misplaced' with a snap distance of 97m. This means that it was snapped to a centre-line which all routes use but there was a closer ITN link. Investigation quickly found the cause of the problem. The route 18 had been accidentally routed down a tunnel (see Figure 3). Bus-stop 673 had therefore been snapped to the location where the road using the tunnel and road above ground converged. This was around 97 metres away from the location it should have been correctly snapped to.

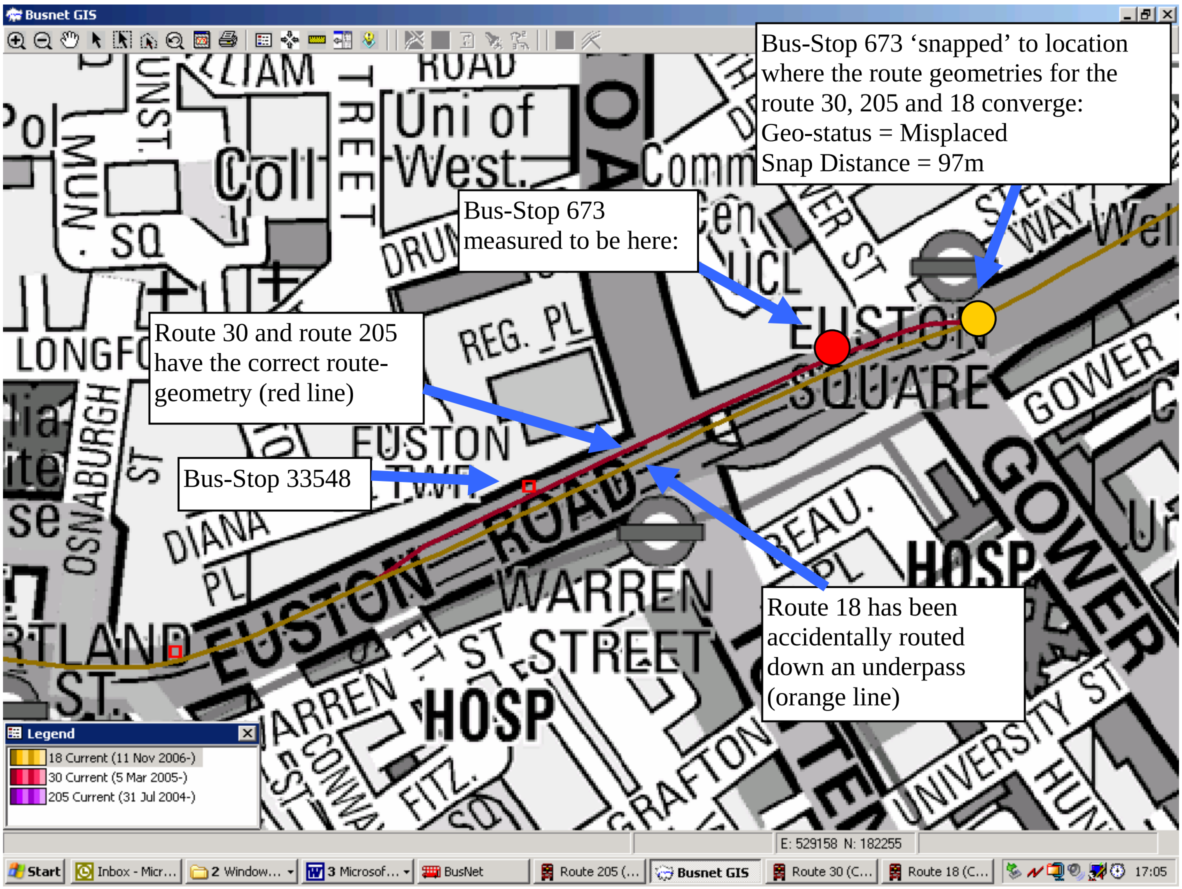

Figure 3 – Example of 'stop-snapping' process identifying problem in the route geometry of the route 18.

If this error was not detected and corrected then the distance between bus-stop 673, and the previous bus-stop in the sequence (bus-stop 33548) would have been wrong by nearly 97 metres.

This would have resulted in bus-stop 673 being announced to passengers on the bus after the stop had visited it! The usefulness of 'stop-snapping' has thus been demonstrated. It should be re-iterated however that the success of 'stop-snapping' is dependent on the initial measurement of the location of the bus-stop.

5 Conclusions

This paper has described the process that London Buses are using to measure and validate the bus-stop location data used in the new GBP120m iBus project. This methodology has been invaluable in ensuring that iBus obtains the high quality data that it needs in order to work effectively. This methodology used by London Buses is a good example of how to provide Intelligent Transportation Systems with the high quality data that they require to function correctly.

Further work is currently underway to explore the possibility of a feedback mechanism from the live iBus system, to improve the stop-location, stop-headings, and link-distances used by the iBus system.

References

- Patchett, N, Evans, J, Williams, N, Kowalski, A, (2005), 'Assessing the use of GPS for congestion charging in London', Traffic Engineering & Control, March 2005 - Vol 46, No3
- Ravikanth K. V., Godfrind A., Beinat E, (2005) 'Pro Oracle Spatial: From Professional to Expert'. Apress, ISBN-10: 1590593839
- Turochy, R. & Smith, B. L. (2000), 'A new procedure for detector data screening in traffic management systems', Transportation Research Record 1727, 127–131.