Transport for London – Using Tools, Analytics and Data to Inform Passengers

Simon REED

Abstract

Transport for London (TfL) uses a single technology solution supplied by Trapeze ITS (CH) to track and manage its fleet of over 8,500 vehicles to provide accurate location information, service control and real-time passenger information. This information is used as the core of a complex command and control and delegated operational performance strategy. These tools provide a unique source of information for passengers on arrival times, disruption information and the open reporting of performance statistics.

This article describes the technology used to support the complex business model and discusses one key building block to many of the passenger information initiatives, the "syndication" of data. TfL data is provided free of charge by the passenger transport authority to an "open market" to enable developments to be made by people other than TfL. The article provides evidence and practical experience based on recent London experience that may inform business and policy makers on the use of Automatic Vehicle Location (AVL), Automatic Fare Collection (AFC) and data technology to manage the performance of their transport network.

Introduction

For any transport network, managing network performance is vital to provide a good service to passengers and this, together with the ability to provide real-time information to staff, operators and passengers, has been an established aim of transport authorities for many years.

What is sometimes overlooked is the need to report performance in a way that passengers and stakeholders can understand.

Accurate real-time arrival predictions are now the most basic of passenger expectations but this and published performance information has to feel relevant to the passenger. It is not enough to approximate. Information must reflect the passengers' experience of the network and therefore become 'trusted'. Information has to be expressed in ways that the customer can understand and it must enable the service to be benchmarked against other services. Get this right and you enable the passenger to gain faith in the service and to preference your service against others in the market.

For London, the requirement for accuracy is greater as the same control and information system that feeds passenger information on service provision is used as the payments engine for mileage and performance payments to its bus service operators of over £1.6 billion ($USD2.4 billion).

Therefore, London relies on its IT systems for vehicle tracking and reporting, service control, on-board information, mileage tracking and performance measurement of its bus network that services over 6.4 million passenger journeys every day.

The paper provides a background to the London bus service, a description of the performance measures and their calculations, the issues that affect those calculations and how these have been overcome. There is also a discussion on how real-time information is now syndicated to third parties to increase passenger impact and take-up.

For London, the requirement for accuracy is greater as the same control and information system that feeds passenger information on service provision is used as the payments engine for mileage and performance payments to its bus service operators of over £1.6 billion.

Background

Transport for London (TfL) was created in 2000 and is the integrated body responsible for the Capital's transport system. Its main role is to implement the Mayor's Transport Strategy for London and manage transport services across the Capital for which the Mayor has responsibility.

These services include:

- London's buses
- London Underground
- Docklands Light Railway (DLR)
- London Overground
- Tramlink
- London River Services
- Victoria Coach Station

London Bus Services Limited (LBSL or London Bus), part of TfL, operates the extensive bus network using over 8,500 vehicles with 4.5 million bus-stop-visits and over 6.4 million passenger journeys per day. The annual bus revenue is approximately £1.4 billion and the bus service costs approximately £1.7 billion annually to run *(Figure 1)*.

The business model is complex, with fares set by the Mayor of London, policy planning route definition and service levels defined by TfL and with a 'franchised' bus service delivery model. The franchised or "Regulated" service delivery model was first established in 1992 and is unique within the UK with the bus operators responsible for the assets (drivers, buses, garage infrastructure) and accountable to TfL for their performance against a contracted service. All ticket revenue goes direct to TfL.

Every route is tendered according to a schedule with contracts being awarded for five years' duration with an optional two-year extension.

Bus operators are paid on the basis of the mileage operated against their service contract together with a variable amount based on the aggregate performance of the contracted route.

Figure 1. New buses in London

In 2005 TfL contracted with Siemens VDO for a real-time service control, vehicle tracking and information service that has been branded "iBus" by LBSL. The contract was subsequently purchased by Trapeze ITS (CH) and hence all references within the paper will be to Trapeze.

Calculating Bus Performance Information – Operated Mileage

The total in-service "mileage" operated by the London bus network is a key performance indicator of the service delivered and is used to calculate the payments due to London bus operators *(Figure 2)*.

Ideally, the scheduled mileage that should be operated by the vehicle is recorded by the iBus on-bus equipment and a complete set of data showing when the bus visited (stopped at or passed) every stop along its journey to completion. With this data, the distance (mileage) and adherence to schedule (performance) can be calculated and the bus operator recompensed accordingly.

In reality a number of factors affect the provision of this data *(Table 1)*:

- Latency: some vehicles do not return to garages for a few days owing to operational requirements.
- Operational conditions: the bus operator may have had to curtail (turn) buses so that some stops were not visited owing to traffic conditions, diversions or other in-service problems.
- Bus operational failures: mechanical, driver or other reasons may have prevented the vehicle from completing its service.
- Equipment failure: the system may not have recorded the data required.

Therefore the bus operator must review all missing trip data and categorise it to explain why the data is not available for processing. This process and the system by which this is performed is known as Missing Trip Validation (MTV) and consists of a Trapeze application accessing the data that has been off-loaded from the vehicles.

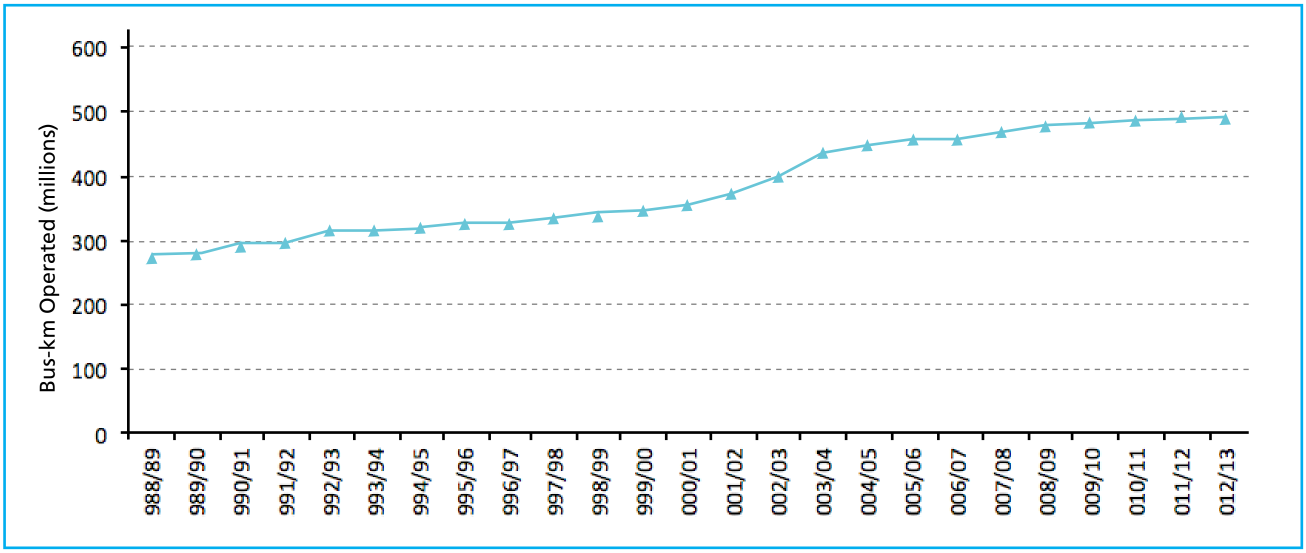
The business model is complex, with fares set by the Mayor of London, policy planning route definition and service levels defined by TfL and with a 'franchised' bus service delivery model... with the bus operators responsible for the assets (drivers, buses, garage infrastructure) and accountable to TfL for their performance against a contracted service.

Table 1. Example of published table of operated mileage (see www.tfl.gov.uk)

Network Performance | *Fourth Quarter 2012/13 (5 January to 31 March 2013)*

	4th Quarter 2012/13	4th Quarter 2011/12
Vehicle km scheduled (million)	118.8	118.1
% Vehicle km operated	98.1%	97.8%
% Km lost for staff reasons	0.0%	0.1%
% Km lost for mechanical reasons	0.4%	0.4%
% Km lost for traffic reasons	1.4%	1.8%
Vehicle km operated (million)	116.6	115.5

Figure 2. Operated bus mileage over time



The operator is presented with all 'missing' data and is invited to codify the reasons why the data is not present. The reasons why some data is not available are then reviewed by LBSL and the figures for operated and 'lost' mileage will be reported for the period under analysis.

With the completion of this process, the data set will have either the recorded times of all bus stop visits or an agreed reason as to why the data is not available.

Calculating Bus Performance Information – Performance to Schedule

While mileage data is a product of the operation of the bus service and produced as described in the previous section, 'performance' calculations use the same data set, but the data is used to measure the performance of the route against the scheduled (contracted) service. There are numerous methods of calculation that could be used and indeed there are several different methods in use globally, but London uses two specific methods.

Table 2. Example of published average EWT figures by Bus Operator (see www.tfl.gov.uk)

Operator	Minimum Performance Standard (%)	5 January 2013 to 31 March 2013 (%)	Variance from Standard
Abellio London	1.14	0.74	0.40
Quality Line	0.70	0.35	0.35
Selkent	1.14	0.79	0.35
Metrobus	1.10	0.76	0.34
Arriva The Shires	0.98	0.67	0.31
London Central	1.19	0.88	0.31
Arriva London South	1.13	0.85	0.28
London General	1.16	0.89	0.27
Blue Triangle	0.79	0.56	0.23
Network Average	1.14	0.91	0.23

1. Low frequency services

Where a service operates with four scheduled buses an hour or fewer, London treats this as low frequency and the contract with the bus operator and the report of performance is based on a percentage "on time" calculation where on-time is within -2.5 minutes (early) or +5 minutes (late) of the scheduled arrival time.

2. High frequency services

Where a service operates more than four buses an hour, London treats this as high frequency and the calculation method is based on the calculation of excess wait time (EWT), this being the difference between the scheduled wait time (SWT) and the actual wait time (AWT) as recorded at points along the route.

The EWT methodology, used by LBSL, is a measure of perceived regularity, measuring the average additional waiting time passengers experience as compared with the waiting time they expect. The lower the EWT, the more likely it is that passengers will not wait more than the scheduled time and will perceive the service as regular.

EWT is defined as the difference between the actual wait time (AWT) and scheduled wait time (SWT), that is, EWT = AWT – SWT, with

$$AWT = \frac{\sum_1^N Away^2}{2\,x\sum_1^N Away}$$

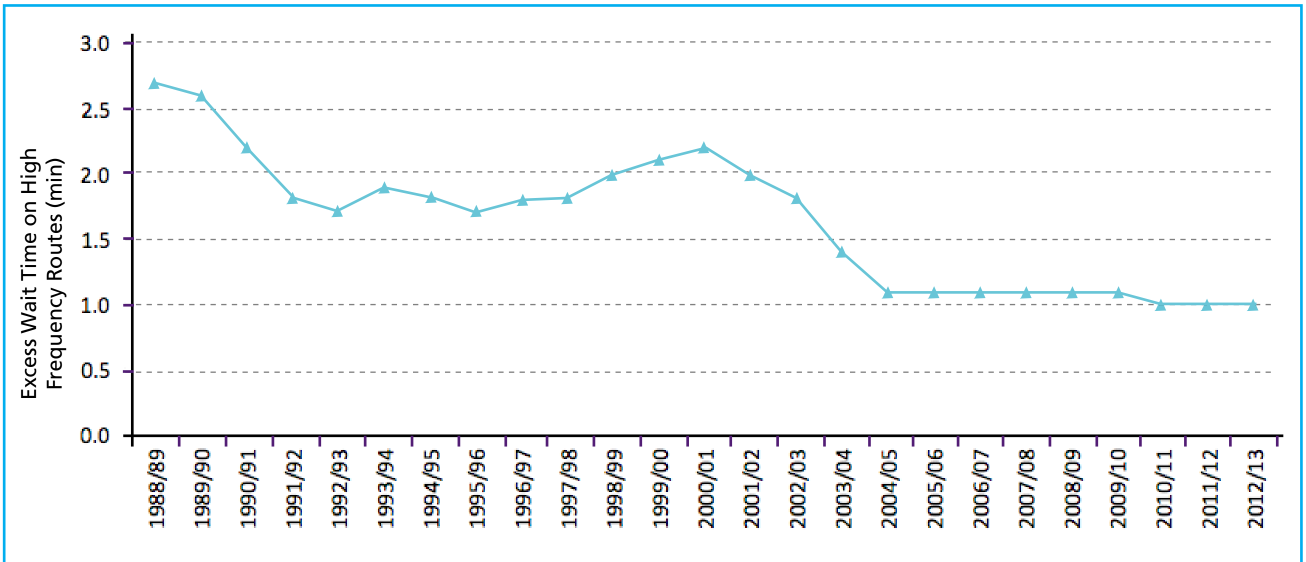
$$SWT = \frac{\sum_1^N Sway^2}{2\,x\sum_1^N Sway}$$

Where: Away is the actual headway;
Sway is the scheduled headway.

For example, a 10-minute scheduled headway route has an SWT of 5 minutes (e.g., half the headway). The EWT methodology assumes uniform arrival of passengers. An EWT of 1 minute means that customers are likely to wait 6 minutes instead of the expected 5 minutes.

EWT is not the only measure of bus performance used by Metropolitan authorities but has been proven to be a reliable measure of calculating bus performance and is a recognised method of benchmarking services between operators.

Figure 3. Excess Wait Time on high frequency routes



The use of EWT as a "benchmarking" tool was researched by Trompet et al. (2011) who explored the use of various performance indicators as benchmarking measures and they highlighted the issue with (any) bus performance calculations, which was that ". . . challenge lies with both data collection and data cleaning" the effects of which were investigated and explored within their paper.

For the performance measures to be accurate, reliability of vehicle location data is the essential start point. Trials conducted by London Bus and Trapeze proved that GPS signals alone were insufficiently accurate to track vehicles in the London environment. London has a number of narrow streets, tunnels, and other characteristics that trigger "canyon" and other "reflective" behaviours on GPS signals that confuse GPS decoders and these would report the vehicle away from its true location.

London Bus now uses an 'enhanced' algorithm designed and built with Trapeze that uses a number of inputs including a Turn Rate Sensor (analogue gyroscope), odometer, definitions of the route, a map and dead-reckoning and the inputs of all of these sources are processed by a Kalman filter to establish the true position of the vehicle.

Accurate, predicted bus arrival times for every stop on London's bus network was a primary objective for LBSL and these are made available to bus operators, operational staff and bus passengers through ... an algorithm and methodology was derived that uses the evidence of previous bus journeys as the default prediction base and not the bus schedule.

Prior to calculating the performance figures, the data is cleansed to remove issues that could be caused by incomplete data sets. This process is complex and has been developed over several years by LBSL working together with Trapeze to develop computational processes to remove the problems experienced in a complex operational environment.

Calculating Bus Performance Information – Predicted Bus Arrival Times

Accurate, predicted bus arrival times for every stop on London's bus network was a primary objective for LBSL and these are made available to bus operators, operational staff and available to bus passengers through a number of digital and analogue information services.

The core to the system is the method of calculation. Following extensive research it was clear that many of the methods of calculation in place globally would be unsuitable for the London environment. Working with Trapeze, an algorithm and methodology was derived that uses the evidence of previous bus journeys as the default prediction base and not the bus schedule.

The iBus system continuously calculates arrival predictions for bus stops 'down-stream' of the current bus position. To do this the back-office systems have the schedule for the route within the base version (definition of the service being operated), and the real-time location information for the bus detailing its position and login details (which route trip etc) received every 30 seconds from the vehicle.

For each segment of the journey the system checks the time of all buses that have used that route segment in the previous half hour and then all vehicles that have covered the same leg over the same time period in the previous 10 similar day-types. For example if a prediction is being calculated on the Friday of a Mon-Fri service, the figures for the previous 10 working days would be compared. Outliers are removed from the dataset and the forecast arrival time is derived and added to the current bus time to provide a time for the segment. This is potentially recalculated by iBus with each bus location update (every 30 seconds) for each bus-stop.

The reality is that although the calculations may take place, the output is only forwarded to an output stream if the predicted arrival time has changed.

Presenting Information to the Passenger

Performance Reports

TfL releases all of its performance information described in earlier sections via its website www.tfl.gov.uk/corporate Here one can find performance data under the "London Buses" tab detailing mileage and performance data for the network, London Borough or ranked by the bus operator (service provider).

Passenger Information

The iBus system is integrated with various distribution and display technologies to provide:

- On-board next stop and destination information in audio and visual forms on vehicle.
- Real-time live bus arrival prediction information through a number of supported channels including the Web (fixed/mobile), SMS and over 2,500 roadside signs.

Figure 4. Bus arrival information

TfL has also released its bus arrival prediction information as 'raw data' through the process of data syndication.

Data Syndication has formed part of a wider TfL Digital Strategy since 2010 and is in line with UK Government Open Data policies and the objectives of the Greater London Authority and the Office of the London Mayor, Boris Johnson. The immediate consequences of this strategy are that TfL (the Authority) would not generally produce Smartphone or other "apps", but would rely on the emerging "digital economy" or third party developers to do this using freely provided data from the Transport Authority.

The primary reasons are that:

a) the cost of keeping up to date with the number of emerging platforms was prohibitive;
b) Smartphone app development is not a core business,
c) by following a Data Syndication route, innovation was far more likely. Following this policy, real-time bus arrival data for over 19,000 bus stops and 8,500 vehicles has been made accessible through an official, supported and open interface. The data is used in third party information services like mobile apps, websites, corporate signs etc.

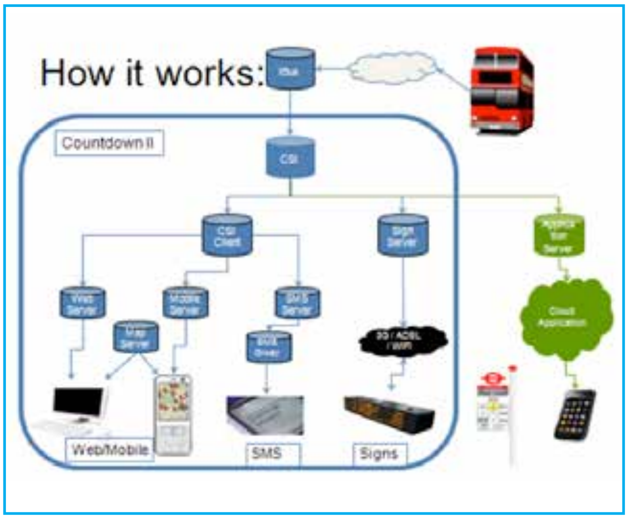
Within the few weeks following the release of the service, the usage of the new service grew rapidly, serving many 100 of megabytes of information, whereas the server load of TfL's conventional Countdown website dropped by almost 50% as App developers stopped using screen-scrape technology to harvest data.

This proves that there is high demand for this service and also that prior to the service release, people were "abusing" the conventional website to retrieve the real-time data – regardless of whether it was officially provided.

The interface service is an integrated part of TfL's information system, which provides real-time information on London bus services consistently through all information channels: 2,500 classic signs at bus stops, mobile and fixed web services via www.tfl.gov.uk/buses, SMS (text message) service and the new real-time API *(Figure 5)*.

To date, some 70-80 apps are now using the syndicated data to deliver information to our customers via Smartphone solutions.

Figure 5. System Overview – live bus arrival information



The API design allows for direct access by clients with limited resources (CPU, RAM, storage) e.g., smaller mobile phones, without the need for an intermediate server process. The same API can also used to feed the data in stream format into other industrial data users like the BBC, Google and other data aggregators.

The server infrastructure that provides the interface is operated in "the cloud", which gives high performance and flexibility to adjust to growing/varying usage.

TfL is one of the first large public transport executives (PTEs) to provide free and open access to real-time data and has provided over 100 different data sources via its portal on www.tfl.gov.uk/developers.

The specification of the API interface is open and the intention is that TfL and our contractor, IVU will submit it as a future standard for transport operators. The standard is not tailored to TfL specific terminology, so it can immediately be adopted at other passenger transport executives. The interface is already starting to be used for some projects in Germany.

Partnership with Developers

As explained, TfL is not producing any Smartphone 'Apps'. So, who is the client in this new model? TfL engaged some of the developers of existing Smartphone apps and worked with them on the design of the API and allowed them access to early versions and worked in a constructive partnership with them to exploit the possibilities that the data provided.

As a result the release of the data played a part, albeit a small one, in the continuing development of the UK digital economy.

TfL engaged some of the developers of existing Smartphone apps and worked with them on the design of the API and allowed them access to early versions and worked in a constructive partnership with them to exploit the possibilities that the data provided.

Additional information

- TfL now publishes over 100 independent data sources for syndication.
- TfL has over 1,100 subscribers for live bus arrivals.
- There are between 70-80 Smartphone 'apps' for London Bus Arrival data.
- The number of Smartphone 'apps' using TfL data can be reviewed by using either the 'iStore' or 'Google Play' websites and search for London bus.
- Details of TfL's Open Data sources are available at www.TfL.gov.uk/developers.
- In addition to the API, TfL has provided live bus arrival information at 2,500 roadside signs, on a fixed and mobile web interface, via SMS (text message) service as well as via the Open Data Syndication service.
- More details of the Countdown project can be obtained at www.TfL.gov.uk/countdown.

Conclusion

London operates one of the largest 'regulated' bus networks in the world and calculates passenger information and performance indicators for all services using a single IT architecture and employing a number of technical and business processes to ensure that accuracy is maintained in testing operational conditions.

The performance of the network continues to improve despite a backdrop of increasing congestion and a growing City population. By concentrating on accurate performance and passenger information, customer confidence has also increased with passenger numbers continuing to climb year-on-year despite a difficult economic landscape.

References

Trompet, Mark; Xiang Liu; Daniel J. Graham. (2011). *Development of Key Performance Indicator to Compare Regularity of Service Between Urban Bus Operators*. London: Transportation Research Board.

Simon Reed is the Head of Technical Service Group, London Buses within Transport for London (TfL), United Kingdom. He joined TfL as iBus Project Director in 2006 and was responsible for the rollout of the world's largest Bus Fleet implementation of an Automatic Vehicle Location (AVL), Real Time Passenger Information (RTPI) and integrated radio system, a £112million contract. iBus provides London's bus passengers with on-board RTPI, bus operators with improved service control and provides London Buses with performance reporting and monitoring. Each of these assists London Buses deliver a more accessible, reliable and consistent bus service and enables the payment of approx £1.6billion to contracted bus operators. He is now managing teams responsible for the operation of London Buses ITS systems. This included the delivery of "Countdown II" project (£32million) providing real-time bus arrival predictions for buses at 19,000 bus stops across London on the internet, SMS and via other digital channels including a data interface supporting over 30 independent Smartphone 'Apps'. In addition, Mr Reed is responsible for coordinating Data Syndication of TfL to the developer community for third party 'App' development and chairs the Wireless Steering Group for TfL's Surface Transport Division.