Acknowledgements

Firstly, I would like to begin with the usual pleasantries of acknowledging the following persons and groups that contributed directly or indirectly to this research project.

Thanks must be attributed to all my close friends and family who have supported me throughout the duration of this MSc Database Systems course.

I would also like to acknowledge my work colleagues at London Buses whom have contributed their support in numerous ways, and to the company for providing the financial sponsorship.

Abstract

I'm investigating how GPS bus location data can be presented using spatial analysis and rendering technologies available as a business decision solution for London Buses – a subsidiary of Transport for London (TfL). I will research suitable technologies and approaches and select one approach to produce a simulated prototype of how the GPS bus location data can be used.

The other identified approaches I will explain in a detailed theory behind how the technologies can be applied to solve the problem domain. The strengths and weaknesses of the approaches will be compared and appraised based on their ability to solve the problem domain.

Contents

1. Introduction

Transport in London plays a major part in the economics of the city [1]; it is therefore likely that the modelling and prediction of transportation flows seeks to increase efficiency in the network – which can have the knock-on affect of increasing productivity as a whole. Service is the key London, if we take people as a commodity (resource) – transport has a major affect on peoples' wellbeing, behaviour and ultimately the local and global economies. Take the daily commute for example – even a minor road blockage can cause delays and have an affect on peoples' behaviour – increasing stress and anxiety levels as well as their journey time to work. This has the knock-on affect of less work productivity due to persons arriving late for work and not in a relaxed state of mind [2].

Transport being a geographical network derives numerous types of spatial and temporal data, the modelling of which has predominantly been used in transport planning. This passive technique of transport modelling is based on historical static data – which in turn generates static economic models of behaviour.

There is a want – however – to have dynamics in the way we model spatial data which can generate accurate predictions and forecasts [3] [4]. We can divide the dynamic behaviours of a transport network into two areas: Progressive dynamics – changes in transport infrastructure, trade and location patterns, e.g. new bus routes, stopping points, growth of industry and developments in an area, population growth in an area. And Reactive dynamics – changes in user's behaviour and / or technological dynamics e.g. increase in patronage at certain times or the impact of unforeseen events on the transport network. It is difficult to predict any of these changes, but there is no reason why we cannot attempt to model real-time transport behaviour data in order to predict behaviour and negate likely degradation in the quality of service. As (Bak and Chen, 1991) define this natural/critical state: "Many composite systems naturally evolve to a critical state in which a minor event starts a chain reaction that can affect any number of elements in the system”.

The growth of Intelligent Transport Systems (ITS) in the last decade is responsible for generating a wealth of dynamic transport performance related data [5]. Yet by its geographic nature – the data is rarely utilised in a real-time decision making environment using the very tools and applications that are ideal for this type of spatial data. The project proposed will attempt to demonstrate how the use of available and emerging technologies can be applied to real-time bus location data from the London Buses iBus system [6] (Automatic Vehicle Location) to be used as a spatial decision support system.

This paper does not however seek to delve into the mechanics of spatial-temporal modelling, but to establish at the theoretical level the reasons, as a result of the spatial-temporal dependency, why vendors have tried to introduce technologies for varying reasons to support the research of the problem domain – and to demonstrate this where possible.

2. Advances in the field of Temporal and Spatial Technology and Databases

2.1. Geo-spatial Data

Business data in certain industries such as transportation, retail, utilities etc have an element of spatiality; there are even intelligent systems that now exist that are dynamically creating spatial data.

For example, if we take a chain of supermarkets as a data set, each record can have the following attributes:

Non-spatial		Spatial	
Name	Victoria Sainsbury's	Name	Victoria Sainsbury's
No.	172 – 174	No.	172 – 174
Road	Wilton Road	Road	Wilton Road
Postcode	SW1 9TM	Postcode	SW1 9TM
Telephone no.	020 7654 3210	Telephone no.	020 7654 3210
Parking (spaces)	200	Parking (spaces)	200
Petrol Station	Yes	Petrol Station	Yes
Restaurant	Yes	Restaurant	Yes
Hot Food Counter	Yes	Hot Food Counter	Yes
Home ware	No	Home ware	No
Recycling	Yes	Recycling	Yes
		Location X	529152
		Location Y	178745

Conventional address   Attribute Data   Spatial address

If we wanted to know whether or not this store was located near to a home address – querying the Non-spatial data set in-relation to a home address would prove very cumbersome if not impossible. The conventional address is made up of three string data type fields, which (i) is in efficient for querying, (ii) requires an unenforceable uniformity (structure) within the string, and (iii) needs consistency and comprehension in the labelling of each field.

The spatial data set on the other hand defines the exact location of the store as an unambiguous[1] OS (Ordnance Survey) Grid Reference, rendering the textual address description as just another attribute of the store. Therefore using spatial operators (e.g. a buffer, contains) we can now easily define the relationship between two spatial locations. Using the supermarket example, the store location is defined as a spatial point; where as the home address is defined as a spatial area, the postal code area in Fig 2.1.1, (people are highly unlikely to know the exact spatial reference of their property).

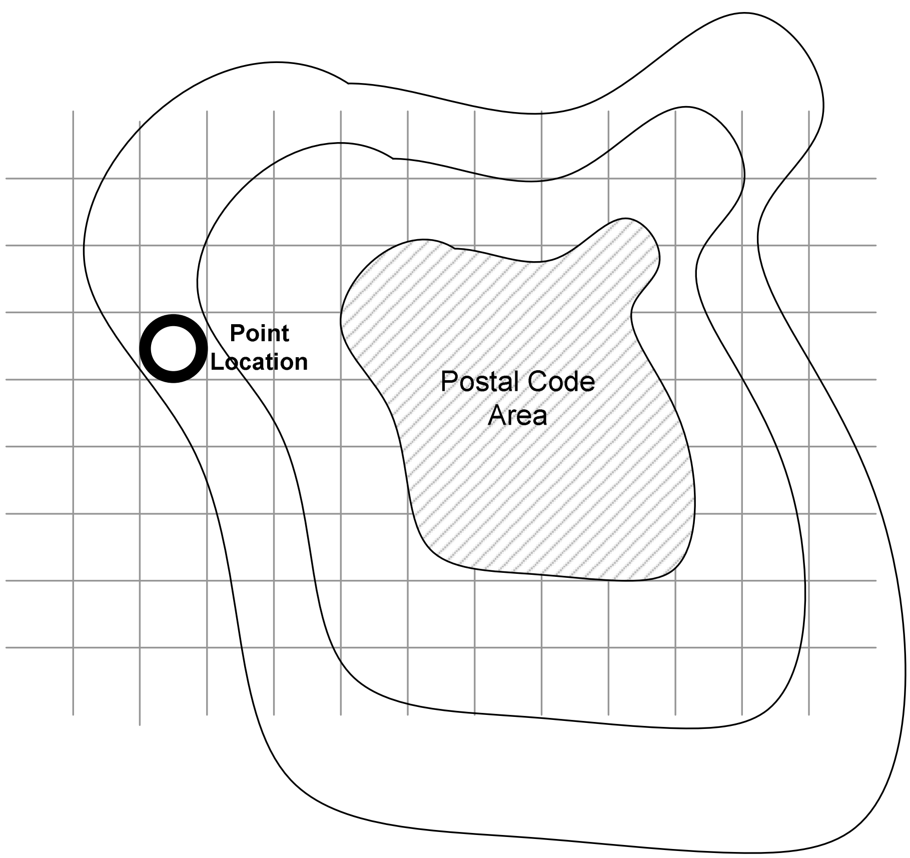


Fig 2.1.1

[1] This assumes accuracy in the recording or projection of the geographical location

So what we end up with is conventional data about a store with underlying space characteristics, which allows new querying opportunities. For example:

The travelling salesman and goods distribution – a spatial query is beneficial and cost-effective in identifying the quickest and most efficient way to serve all stores.

Population census – identification of the type of population within a catchments area of a store can help marketing strategies on what particular products to promote, for example, a store in Eastbourne may want to highlight particular products associated with the elderly and retired, as they make up a large percentage of the overall population.

One of the fundamentals of spatial data management is to provide a greater level of storage and querying intelligence than conventional data management. It would therefore figure that there is an underlying requirement for tools and applications to: manage, analyse and render the spatial data. The following sections look at associated approaches and areas of consideration for spatial data management.

2.2. GIS-Centric Applications

GIS-Centric approaches have specifically been managing spatial data for many years. Although any conventional data associated with the spatial data is redundant in the GIS centric data structures. Indeed, any relationship with conventional data is maintained outside of the application [7]. This is common in many GIS applications today – where the spatial operations are performed within the application itself and linked to relationally structured databases. The approach of GIS is to model the spatial data based on area using a map plane or planes (this varies dependent on the approach), and carry out spatial operations based on the interaction between the planes. (See Fig 2.2.1)

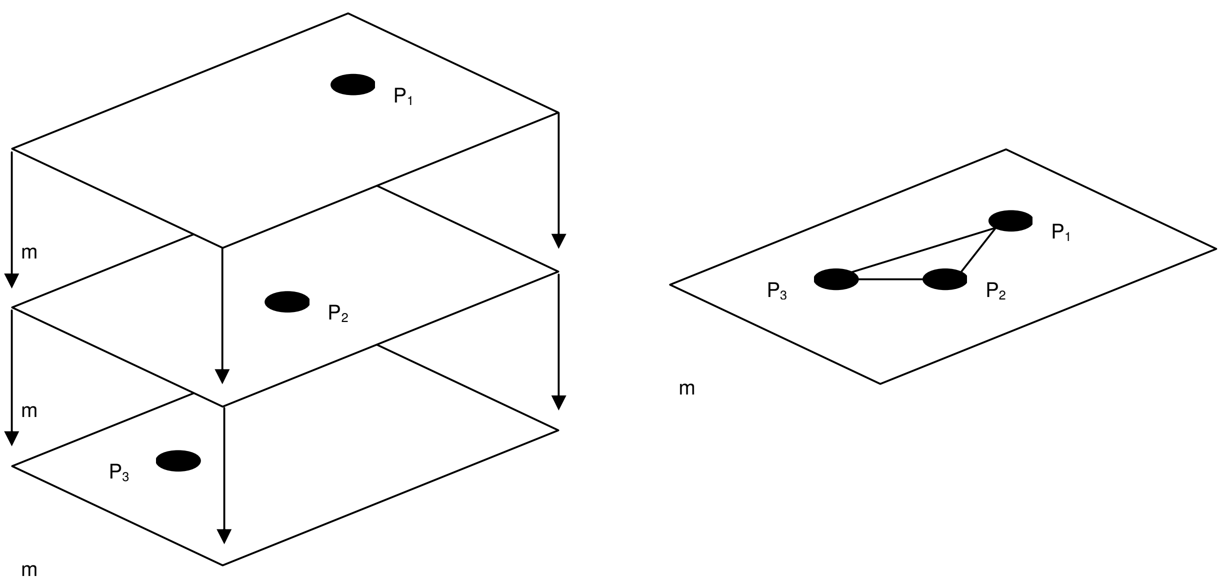


Fig 2.2.1

However, not all GIS approaches support different types of spatial modelling, i.e. point, line, area. In the same way that not all GIS approaches support the typical spatial operations, i.e. buffer, nearest neighbour etc. [7]

For years GIS were developed and used to meet the requirement of a specific problem domain, usually due to the overheads associated with the processing of large amounts of spatial data. This meant that each system developed in their own silo worlds – forming incompatible complex data structures and analytical languages [7] [8].

The different structures of spatial objects in GIS are particularly an issue when a spatial operation on two objects yields a different spatial object, which may not be supported even by the same GIS application.

The Open GIS Consortium (OGC) recognised this to be an issue where the sharing of data amongst organisations is concerned. They worked on an abstract data model for spatial data in order to enhance interoperability – and which now forms the ISO19107 Spatial Schema Specification [9].

2.3. Spatial Databases

Efficient processing of conventional data can only be achieved in DBMS [7]. Thus, recent approaches in the field of spatial data analysis have focused on managing the spatial data also in the database. This has meant adapting the way spatial data is structured and stored in-order to accommodate it, whilst maintaining the underlying fundamental characteristics, which makes a database what it is.

The Oracle DBMS approach is to store spatial data as another data type. The data type supports different data structures relating to the spatial object, i.e. line, point and polygon, as well as more complex arcs, boundaries, circles or a combination of any of the aforementioned. The following table (2.3.1) explains the data structures.

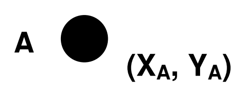
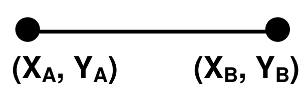
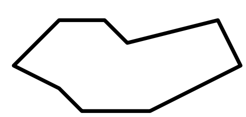
	Data Structure Syntax	**Example**
Point	SDO_GEOMETRY(2-dimensional object, coordinate_system SDO_POINT_TYPE(longitude_ordinate, latitude_ordinate, no third dimension))	
Line	SDO_GEOMETRY(2-dimensional line, coordinate_system SDO_ELEM_INFO(offset, line type, interpretation SDO_POINT_TYPE(A_longitude_ordinate, A_latitude_ordinate, B_longitude_ordinate, B_latitude_ordinate)))	
Polygon	A combination of the above.	

Table 2.3.1

2.4. Temporal Databases and Queries

The temporal dimension of databases is introduced when information requires an aspect of time in order to organise. The concept of time driven data can be found in most industries, for example, bank account transactions, reservation systems, etc. [10]

Temporal Database Research has concentrated on standardising the numerous temporal models that support specific temporal requirements and features. Early research looked into the extension of the relational model to take account of the temporal but only yielded the impossibility of incorporating the multiple facets of time. This virtual barrier is apparent in that most relational temporal models only support a linear model of time [11].

In order to introduce an element of time in a conventional RDBMS a temporal granularity is normally introduced as a data attribute in the database schema. This means the object either gets time-stamped once at a single point in time or the attribute is continually overwritten – erasing any historical temporal data charting changes through the object's life span, see Table 2.4.1 example. To maintain the historical account of time and space in a relational database the use of tuple versioning is required.

Bus	X Co-ord	Y Co-ord	Matching Link	Speed	Time Stamp
1	531001	180001	Link 1	22	01/09/2006 14:00:00
1	531011	180011	Link 1	13	01/09/2006 14:00:30
1	531021	180021	Link 2	16	01/09/2006 14:01:00
. . .					
2	531001	180001	Link 1	11	01/09/2006 14:29:30
2	531011	180011	Link 1	9	01/09/2006 14:30:00
2	531021	180021	Link 2	6	01/09/2006 14:30:30
. . .					

Table 2.4.1

Bus	X Co-ord	Y Co-ord	Matching Link	Speed	Time Stamp	Valid From	Valid To
1	531001	180001	Link 1	22	01/09/2006 14:00:00	01/09/2006 14:00:00	01/09/2006 14:30:00
1	531011	180011	Link 1	13	01/09/2006 14:00:30	01/09/2006 14:00:30	01/09/2006 14:30:30
1	531021	180021	Link 2	16	01/09/2006 14:01:00	01/09/2006 14:01:00	01/09/2006 14:31:00
. . .							
2	531001	180001	Link 1	11	01/09/2006 14:29:30	01/09/2006 14:29:30	01/09/2006 14:59:30
2	531011	180011	Link 1	9	01/09/2006 14:30:00	01/09/2006 14:30:00	01/09/2006 15:00:00
2	531021	180021	Link 2	6	01/09/2006 14:30:30	01/09/2006 14:30:30	01/09/2006 15:00:30
. . .							

Table 2.4.2

In Table 2.4.2 we extend the temporal model to include valid time into the relations for each spatial location, yet still in a linear model of time. This allows a current version of valid records to be attained for aggregation to the spatial link, as in Table 2.4.3. The temporal database will store when an event occurs (time value) and when certain facts are considered to be true (time period: start and end time).

Note the primary key constraint has changed to the bus, link and now the validation period in order to uniquely constrain the temporal dimension.

Link	Average Speed (mph)	Temporal Interval (mins)	Valid From	Valid To
. . .				
1	15.5	30	01/09/2006 14:00:00	01/09/2006 14:30:00
1	14.2	30	01/09/2006 14:00:30	01/09/2006 14:30:30
. . .				
2	11	30	01/09/2006 14:01:00	01/09/2006 14:31:00
. . .				
2	11	30	01/09/2006 14:30:30	01/09/2006 15:00:30
. . .				

Table 2.4.3

The addition of the start and end time fields resolve the issue of data expiration as soon as a new location fix is received but still creates new records of data and greater transactional processing, as well as the outstanding requirement to vary the temporal interval – application of the branch model for different days and periods during the day. This is where object orientated approaches are becoming prominent in the research field to solve the limitations of the relational approach [11].

Temporal databases explore ways in which the granularity of data or facts from one set differs to the granularity of another set can be converted or expressed to provide consistency, this is as long as the semantics behind the granularity is understood, for example, a weekend is a granularity made up of Saturday and Sunday. Fig 2.4.1 shows some common temporal reference systems.

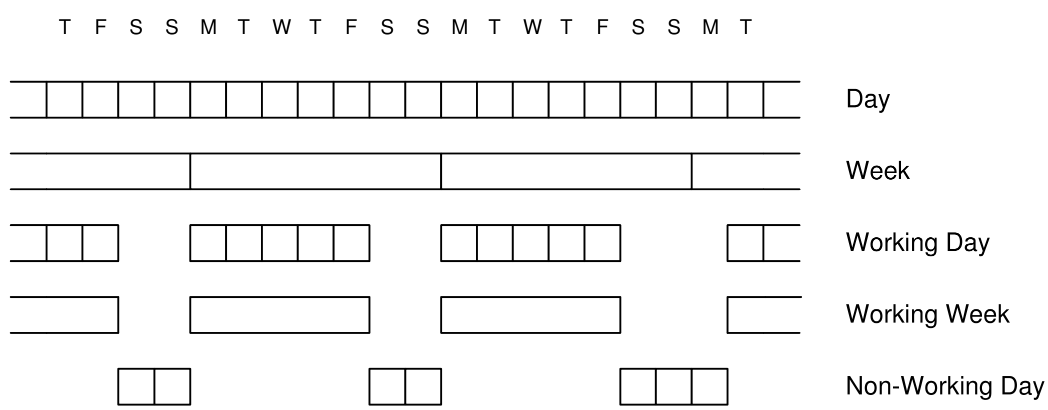


Fig 2.4.1

Geographical Information Systems (GIS), however, are unable to sufficiently handle the concept of the temporal dimension [12]. They can only produce a snapshot of time-varying objects [7] [14] [15]. Objects that move will lose any historical spatial attributes it had [14]. The creation of another object following a move (to retain the existing location) means there are now two spatial objects for what is effectively the same object – this causes repetition of data and validity issues of both objects, i.e. which object is valid at a particular moment in time? And in a dynamically changing transportation network – the objects are spatially and temporally susceptible then most geo-spatial features analysed by GIS. It is this limitation in current GIS platforms that has prompted the need to develop and incorporate tools for spatial-temporal analysis and display. The data structures are two-dimensional (see Fig 2.2.1) and are unable to incorporate another dimension or inflexible in defining relationships between objects.

With the introduction of modelling temporal data within a database comes the issue of retrieving and manipulating the data. The most common database query language SQL (Structure Query Language) did not have the capability to handle temporal querying. Hence the temporal extension to SQL was conceived in 1993 – TSQL2. The extension allowed for operations on the data that could yield temporal projections, temporal aggregations, and transformation amongst granularities of time periods or time point data based on the state or event of a relation. Following the publication of TSQL2 in 1994 further research recommended adding the valid-time and transaction time support into the next temporal SQL specification – SQL3 [10] [16].

3. Characteristics of the Transport Network

Spatial transport data is used in various elements within the transport industry from planning to operations – aggregation of data on a geographical scale has proved immensely beneficial for transport authorities [17].

Transport networks were traditionally represented by nodes and straight line "links" with no relation to geography and comprehension of dimension, i.e. distances and geometry of the network. Hence there was no rendering representation of a network. Another limitation was that nodes were just anchor points to connect links with no concept of holding attributes. Which given that in some models of networks – nodes themselves have a function to play. Take for example, a Railway Network: a Station is an Access node into the rail network. Stations have a capacity on the amount of 'people traffic' it can handle, which in the traditional modelled networks would not be included and therefore link capacity (rail lines) could falsely be represented [17]. This network is not only distributed spatially but is also dynamic from a temporal perspective which is of great interest to data modellers. For example, traffic flows along certain parts of the road network will be near to capacity at certain times of the day, i.e. morning and evening week-day peaks. This near capacity will have an impact and 'knock-on' affect to the transport network.

We can see there are now even more benefits in using spatially (and geo-referenced) related transport data than was previously being realised. Where before spatial data has historically been used for planning – there is more to be gained in the analysis of more real-time (temporal) spatial data – reduction of operating costs, improvement in the quality of services offered for instance [8].

3.1. Application of Approaches on Transport Networks

3.1.1. Geographical Information Systems (GIS)

Geographical Information Systems have been instrumental in their application to model, analyse and render transport networks from a transport planning perspective and are rapidly becoming prominent in the field of travel forecasting [12]. In the late eighties and early nineties the ideological linkage of GIS and modelling of transport networks was conceived into reality with the development of the TransCad software (1988) and software developed for the US Bureau of Transportation (1990) to geographically analyse US Census data [12].

The integration of transport modelling software and GIS highlighted major interoperability issues in the partnership between the two:

- Data updates in one package do not automatically get updated in the other. It is quite common to import and export between systems.
- Changes to GIS distinct core scripting language compromises compatibility.
- Data structures are different or incompatible.
- Users need to be familiar with different systems.

[8] [12]

These limitations have by no means been impossible barriers in the goal to achieve interoperability between different functional systems. Standards like those suggested by the Open Geospatial Consortium (OGC), World Wide Web Consortium (W3G) and the International Standards Organisation (ISO), some of which are discussed later, have emerged through research as a means of addressing some of the above issues.

3.1.2. Database Management Systems (DBMS)

The use of and advantages that DBMS offer for modelling Transport networks has come of age in recent years with the extended ability to maintain the arc-node topology, two-dimensional geo-referencing, and linear referencing, synonymous with the topography of a transport network, within a database [18]. As was discussed earlier conventional DBMS do not have the associated extensions (modules) to manage the complexity of spatial and/or temporal data. The products that are spatial compatible must be able to store this spatial data [8], for example, having a spatial data type capable of modelling the different spatial object structures found in a transport network (Table 3.1.2.1) – and even more importantly the interaction between them (Table 3.1.2.2).

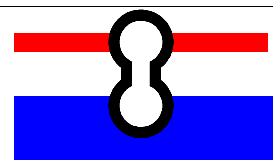
Spatial Object	**Transport Network Object**	**Example**
Points	Vehicles, Network Access Points (Bus Stops, Stations), Junctions	
Lines	Roads (Centre-lines, Lanes), Rail lines	
Polygons	Bus Garages, Bus Stations, Train Depots	

Table 3.1.2.1

And the DBMS and database must support the language to retrieve, query and update the spatial data [8], for example, creating subsets within the transport network, spatial relational queries – using spatial data or even non-spatial data.

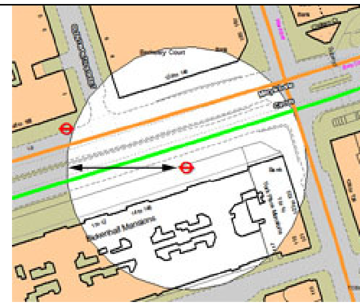
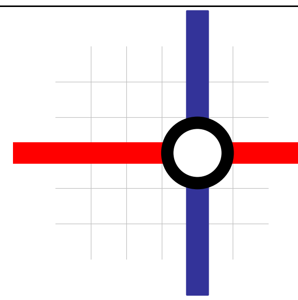
Spatial Operation	Transport Spatial Operation	Example
Within Distance (Buffer)	Query to find all bus routes within 50m of a selected bus stop.	
Nearest Neighbour	Matching of a vehicle GPS location fix to a transit link within the road network.	
Intersection	Crossing of two transport networks, the intersection of two or more different rail lines, or the junction of two or more roads.	

Table 3.1.2.2

Yet there are areas of using a DBMS approach that need to improve in order to successfully model transport networks – especially for use within an intelligent transport system (ITS) – which requires high integrity and processing of real-time information [18].

The storing and processing of spatio-temporal data on the scale of transport networks is yet to be proven. Rendering of spatial data is also very limited and by no means capable of presenting geographically using maps without interfacing to purpose third-party software.

It is conceivable that one approach would be to use DBMS to model and store the spatial and conventional data, carry out the spatial analysis and use the GIS to render the results using its successful ability to manage maps.

Though there are still the incompatibility issues between the DBMS and GIS models when it comes to transporting the spatio-temporal data required for presenting. This is a research area where bodies have come together to try and introduce some standards in geo-spatial, in particular, in order to overcome issues like these. These standards are discussed later on in this section.

3.1.3. Intelligent Transport Systems (ITS)

The growth of Intelligent Transport Systems (ITS) in the last decade is responsible for producing a wealth of dynamic transport data and creating its own issues. Using the Global Positioning System (GPS) – the real-time enabled location data, that is generated, can be used in GIS to locate spatially dynamic assets – such as vehicles in a transport network [19]. GPS can enhance modelling within a transport network by providing additional dynamic measurements – relating to network speed, for example – combined with sophisticated map matching algorithms [12].

The navigable[2] data model required for ITS – should at least support the functions, defined in table 3.1.3.1, which also compares the application support for each [18].

[2] Digital geographic data bases of a transportation system

	GIS	DBMS [3]
Translation of co-ordinate based locations	✓	✓
Support of map matching (discussed later)	?	?
Capability to represent the network in detail	✓	?
Spatial analytical capabilities	✓	✓
Temporal validity	✘	✓
Procedural Language Support	✘	✓

Table 3.1.3.1

GIS and GPS data combined have been popular for use in location Satellite Navigation (SatNav) systems – but still in its infancy in its use in the semantics of transport modelling, analysis and rendering – especially from a data performance related perspective in real-time operations – mainly due to the challenges it brings [8].

3.1.4. Temporal

The real-time temporal element introduced by Intelligent Transport Systems to spatial data [20] can be defined as 'the affect on spatial objects or areas that are likely to change their location and characteristics over time'. And it is this new dimension of time, which must be supported efficiently [7], especially in the dynamism of the transport network.

Fig 3.1.4.1 demonstrates the spatial and temporal dynamics of just *one* single bus journey – highlighting the complex model required for spatio-temporal analysis of an entire transport network.

[3] Based on a DBMS having full proven spatial and temporal functionality

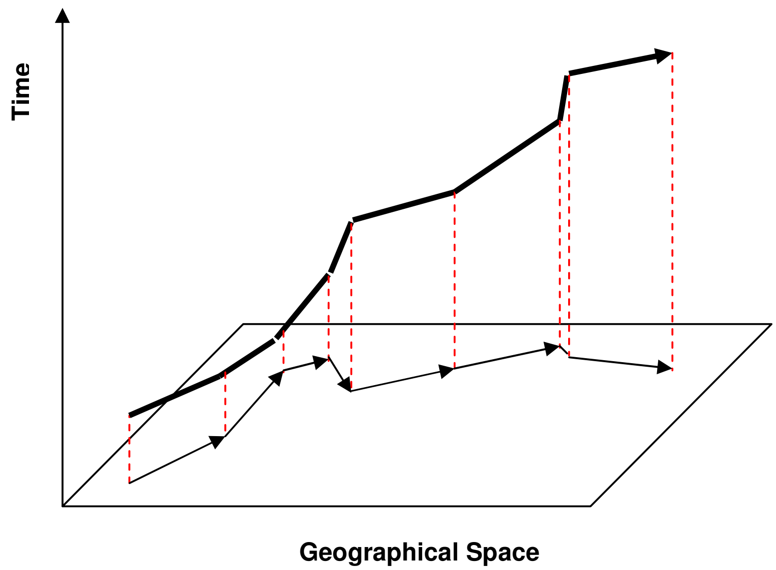


Fig 3.1.4.1 – A space-time path for a bus journey

The most recent development of the TransModeller software (Yang and Slavin, 2002) has attempted to extend the underlying GIS data model to incorporate the temporal dimension for transport modelling [12]. However, research has predominantly focused on managing the spatio-temporal data in databases [14]. DBMS' have the following advantages over GIS, a) to store and associate spatial data with conventional data, b) use of the SQL language and operators for analysis, and, c) widely adopted standards in relation management and data modelling [8]. It is the characteristics of the database that provide a stable platform to research into solving the limitations of spatial and temporal data modelling – pertinent to any transport network – where the focus has been to solve two major problems: unifying spatial and temporal data models into a single spatio-temporal data model and then integrating the existing spatial and temporal relationship operators into the select statement of SQL3 [14].

Using real-time GPS location data from a bus, as an example, the temporal functional dependency is not only used to enforce a time window of constraint on the data that we wish to process but is also used to eliminate redundant data in the temporal relations [21] where the data is no longer required due to its invalidity in its given application. In this real world example – the real-time data from over an hour ago is all but redundant due to the changing nature of traffic conditions on a local scale. In essence an event occurs that changes the state of the entity – the state of the entity is then valid for as long as the temporal granularity dictates [14].

The temporal granularity in this example could vary from Minutes to Hours. At certain times of the day the data granularity of an hourly basis would suffice for the application of the data; where as at other times the granularity could be required on a minute basis – for example during peak hours of the day traffic conditions are highly dynamic in their changeability as opposed to night where the conditions are very much static on an hour by hour basis – but we do need to be cautious. Granularity down to the minute would not be efficient in this scenario, given that it would not provide any benefit [21].

In order to successfully model this behaviour we introduce the temporal model concept of a branching model [11], an Object Orientated approach which deals with the non-uniformity multi granularities of time spans required for complex temporal data modelling of: (i) Working Days and Non-working days; (ii) Operating periods [non-standard – unique to the bus operations industry] (early morning, mid-morning, late morning, lunchtime, afternoon, evening, night) (See Fig 3.1.4.2). Consideration of using weekdays and weekends is not viable – as bank holidays, Christmas day, Good Friday are holidays that fall on weekdays but which we would wish to treat different in a transport network. The division into working day and non-working day – drilling-down to operating periods is efficient enough for the desired thresholds of data aggregation that will provide an accurate representation of its intended application. Fig 3.1.4.3 demonstrates the branching modelling of the time spans in our bus network example, which incorporates the industry time spans explained in Fig 2.4.1 and Fig 3.1.4.2.

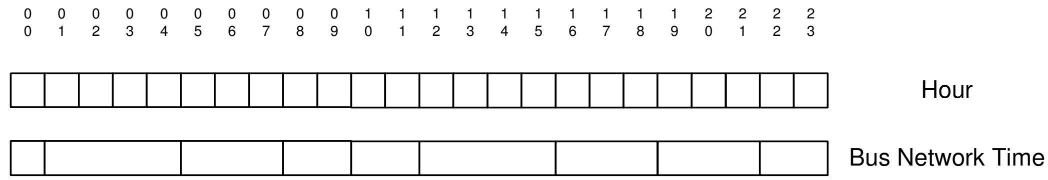


Fig 3.1.4.2

Fig 3.1.4.3 represents the bus transport network in a temporal object orientated branching model, where P_temporal Order is the object subject to an order of change in time and the P_temporalPrimitives denotes an anchored period of time, e.g. a day is from 00.00hrs to 23.59hrs.

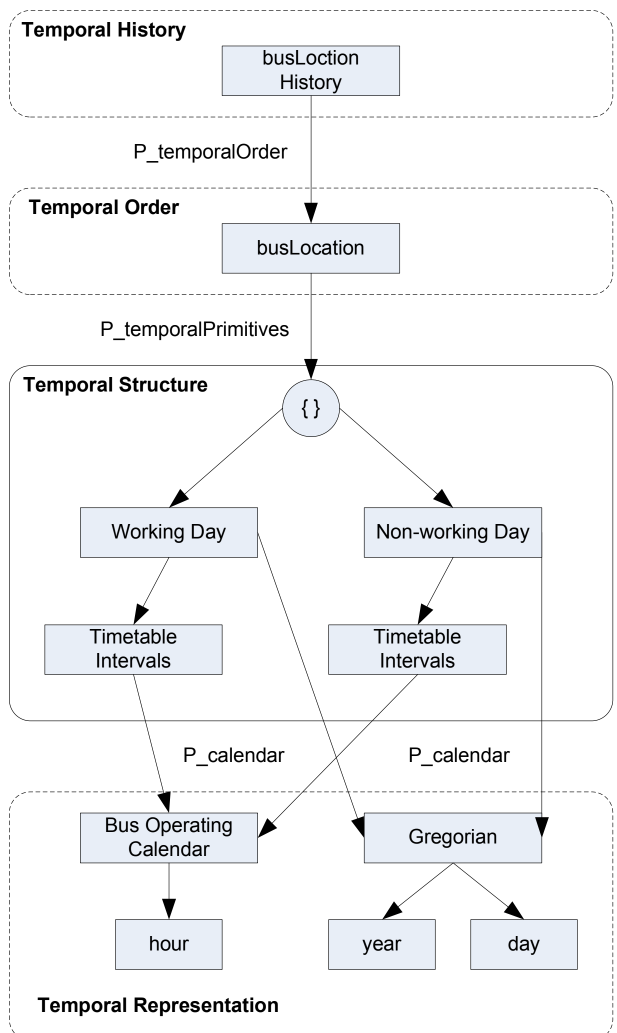


Fig 3.1.4.3

3.1.5. Geography Markup Language (GML)

The Open Geospatial Consortium (OGC) is a consortium of companies, agencies and universities setup to work towards a common goal of solving interoperability issues in the geospatial market.

In April 1999 the OGC launched a project to define standards for geospatial information. The Geography Markup Language (GML) was subsequently developed based on XML standards – and is fast becoming the standard XML encoding for geospatial information – concerned with describing geographic content as apposed to its presentation [22]. Like XML – GML exploits the same benefits that its cousin is accustomed to, providing a basis for solving some of the interoperability issues [23].

- Integration with non-spatial data
- Storage and transport of data
- Transformable
- Automated data integrity
- Easily editable

The XML encoding standard for geospatial metadata was mooted as a key element of the OGC Web Services architecture – which makes possible the search and distribution of spatial data and services across the Web and other portable wireless devices – independent of the platform. Yet one of the major breakthroughs from the GML standard was its ability to resolve many of the difficulties associated with incompatible data models [24] – which were discussed earlier. GML allows the self-documenting transfer of information between systems – which allows either system to present the information as it is programmed to do – or to transform the data into its own model by using an accompanying spatial schema [25].

The current version of the GML specification version 3.1.0 includes GML support for temporal referencing structures [26]. It identifies that values with a reference to time need to be measured relative to a temporal reference system [22] [24] – common reference systems being the Georgian Calendar and the 24-hour local or Coordinated Universal Time (UTC). Although depending on the requirement of the application – may entail use of alternative reference systems where the variations of time granularities are non-standard like the bus network example in the previous section [22]. This feature of the current GML specification should not be confused with that of time varying information (dynamic temporal dimension of data), which is yet to be determined how this dimension can be incorporated. The OGC have recently proposed recommendations for extending the current geo-spatial standards in order to incorporate [27]. Working as part of the Temporal Evaluation Assessment (TEA) programme, the OGC have been testing and demonstrating ways to manage and use time-varying geospatial information with the aim to make improvements in the areas of:

- Temporal requirements, technology gaps and potential solutions
- Capable standards and compliance
- Data schemas and databases with temporal capabilities
- Stakeholder participation

[27]

The proposals have included the addition of metadata for features with temporal information, which should meet the requirements demanded by the likes of application within transport, logistics and communications industries. The new metadata tags are summarised from the OGC Temporal Recommendations paper – reproduced below [27].

1. *Feature instance* – a single instance of features (usually the most current)
   i. Currency timestamp – the date and time at which this feature became effective
2. *Feature history* – a succession of revisions to a feature over time, to include past and future revisions.
   i. Revision – a unique sequential identifier (integer value) that defines the successive revision of features.
   ii. Valid time - real world time the change occurred (valid start time) and the change is no longer in effect (valid end time)
   iii. Transaction time - when the database was updated to reflect the change (transaction start time) and when the change is no longer in effect (transaction end time)
3. *Feature observation* – a time series of observed values from an in-situ feature sensor.
   i. Observation type – an enumerated name of the type of report (example, temperature).
   ii. Time of observation
   iii. Value of observation

4. *Moving Feature* – a feature that changes location over time. The trajectory of a single moving feature contains.
   i. Time
   ii. Location
   iii. Other attributes as defined in ISO 19141

[27]

4. Summary

There are evident advantages in spatially enabling data or using spatial data for development and use within business tools. None more so than transport industry related. Yet a distinction should be noted that Transport networks do not have to be geographically referenced [15] they can just as easily be spatially modelled as a linear topological diagram – but this does inhibit its use in the real-time environment, due to difficulties in locating problems against familiar points of reference, i.e. a map feature.

Without aggregation to the spatial link level within a transport network, GPS data modelled, especially for rendering in GIS, is meaningless as there is no fixed point in which the analysis can be referenced against – a spread of GPS point data is of no use. It also allows us to use a Linear Referencing System (LRS) [15] to overcome potential issues like indicating direction for snapping location data to the correct link. Thus the advantages of spatially enabling a transport network start to be realised.

Traditionally Geographical Information Systems (GIS) are area based analytical approaches not particularly designed for the linear and nodal topography of transport related networks and lacking in dimensional support of what is a spatio-temporal network. So, are GIS and Transport models compatible? It does really depend on the use, for example planning and asset / infrastructure management, the two are ideally compatible, but largely GIS technology as it currently stands is not suitable for the level of sophistication required for the application to transport modelling of real-time data in an operational application environment like an Intelligent Transport System (ITS) [15].

The effort of database vendors to address their short-comings and to expand their portfolio of data type support is evident by comparison. The integration of features like PL/SQL, spatial and temporal operators in one DBMS package is a powerful approach for the active database functionality necessary for use in Intelligent Transport Systems (ITS).

Both GIS and DBMS vendor solutions have performance overheads and are very much reliant working independently of each other.

Open interfaces like the OpenGIS spatial schema and Geography Markup Language (GML) help to achieve interoperability and remove barriers between different kinds of systems that process spatial data [25]. So does the answer to spatio-temporal issues lie in GML? The modelling of bus locations as GML features, that have associated temporal information, is a possible solution in solving interoperability issues between systems that in turn have their own advantageous features such as spatial and temporal storage and map rendering capabilities.

The theoretical research conducted in this paper would rank the evolution of geo-spatial and temporal applications to be as follows, where the research and development of today would be positioned at about number three.

1. Basic collection of transport infrastructure locations
   – Spatial transport related data that is substantially static which is stored and represented for reference and basic querying purposes.

2. Real-time vehicle tracking
   – The location of a vehicle is tracked from GPS location fixes and transmitted back centrally for rendering against mapping or a network diagram.

3. Near real-time geo-simulation
   – A near re-run of what has just happened – spatial data with a temporal dimension is historically recorded and the entire set is available for "play-back" from a defined time period [28].

4. Real-time condition analysis

   – Processing of numerous sources of spatial and temporal metrics to produce a meaningful aggregated outcome, e.g. network status, air quality.

5. Predictive geo-simulation

   – A predicted outcome of conditions based on behavioural and mathematical modelling in response to changes in spatial and temporal metric data.

5. Appraisal of Spatial and Temporal Informatics Approaches

This section appraises a few of the leading market products or approaches available today that we could consider to use.

5.1. Database Management Systems

Oracle 10g[4] with Oracle Locator and Spatial with Oracle Map Viewer [29] [30]

5.1.1. Background

Oracle's claim as the leading player in the geospatial database management market cannot be ignored as an approach. Spatial support has existed in some form or another within earlier releases of the core Oracle Database platform, but has really come of age in the latest 10g release.

5.1.2. Credentials

Oracle manages spatial data like any other data type, with its own native data type which introduces the flexibility and manageability available within this fully integrated Oracle platform.

5.1.3. Method

Entries in a relational table are spatially enabled by adding in a spatial column (SDO_GEOMETRY) into the database schema, an object that has a geographic meaning or representation of the corresponding data is created.

Spatial distance techniques need to be used in order to snap to the network, in order to locate to the network. Network analysis cannot be used as the location data is easily distributed either side of a network. This requires use of the spatial analysis functions: *Within-distance or Nearest-neighbour*.

5.1.4. Strengths

All the benefits of a relational database, including:
Data Mining – use of the widely standard Structured Query Language (SQL);
Data Storage – ability of associating spatial objects with conventional data;

[4] The version used for research in this paper is the Oracle Database 10g Express Edition

Data Modelling – the precision and flexibility to model basic and complex spatial objects within database exclusive data type. And when representing date-time information in Oracle it is possible, from version 9 onwards, to represent points in time and time intervals using ANSI SQL data types such as timestamp and interval.

5.1.5. Weaknesses

The use of spatial functionality is still relatively new to the Oracle community and has never yet been fully utilised. Also, although Oracle supports the temporal concept of valid time – it lacks comprehensive detail and documentation in this area.

5.1.6. Opportunities

Data in corporate databases can be spatially enabled but there has never been such desire to within industries. Database Management Systems are common place in companies – vendors like Oracle, Microsoft and IBM have a strong user base of their products within the IT industry. The product is there, the realisation of their capabilities just isn't yet.

5.1.7. Threats

It is difficult to use the temporal element in such a short time, so validity of data to use in a short setup time would not be possible to produce accurately targeted results. But is by no means as good as if we were to use more meaningful granularities where larger amounts of valid data would produce a greater accurate model. Earlier versions of Oracle represented points in time with the *date* data type, which is precise to within one second, and time intervals as numbers (where 1 = one day).

5.1.8. Comparisons

In comparison to the MapInfo platform, Oracle has far superior data type storage and modelling capabilities.

5.1.9. Summary

The Oracle 10g platform is the strongest contender by far from a familiarity, ease of use and learning curve perspective. It has all the operations and analytical capabilities that its rivals have as well as the rendering capabilities within the integrated Map Viewer.

5.2. Geographic Information Systems (GIS)

MapInfo Professional v7.8 and v8.5 [31]

5.2.1. Background

Geographical Information Systems have historically been heterogeneous stand alone applications used for spatial analysis and producing the results graphically against mapping.

5.2.2. Credentials

MapInfo Professional is a well established product in the mapping application market.

5.2.3. Method

The location data is manually projected against the Ordnance Survey British co-ordinate system to produce a spatial object point for each location on a map plane (layer). For each location point, the nearest neighbour spatial query identifies its location in relation to a reference transit link on the Ordnance Survey Integrated Transport Network (ITN). A thematic map is produced based on averaging the speed attribute of each location record matched to each individual transit link in the transport network.

5.2.4. Strengths

The advanced rendering capabilities for spatial data against mapping from a statistical perspective make it a good tool for presentation. The spatial analysis techniques are ideal for logistical operations in industry.

5.2.5. Weaknesses

For use in a real-time application environment, the automatic transaction capabilities of the data Spatial analysis of data using GIS is far from dynamic and as a standalone application not suited to a real-time operational environment. In effect MapInfo is designed for use as a planning tool and heavily relies on manual operation. Which for use in an operational critical system is not only time consuming but not practical on the size.

5.2.6. Opportunities

GIS products have database interfacing capabilities but are very much standalone applications. The latest version of MapInfo supports most widely used database applications, including Oracle 10g, Microsoft SQL Server, and Informix 9.4. The opportunity is to store more complex modelling of spatial and / or temporal objects in the DBMS and use the GIS platform to render.

5.2.7. Threats

A product like MapInfo Pro heavily relies on user interaction and is therefore not an appropriate application development tool in which to build a spatial and temporal database application. MapInfo's storage capabilities are limited; the modelling of temporal dimensions for instance is not a fluid as the database approach.

5.2.8. Comparisons

MapInfo is different type of application in comparison to the Oracle DBMS even though there are similarities in the functionality. This is where there is an overlap, but ultimately they have different uses.

5.2.9. Summary

Not the “all round” package when it comes to data storage and temporal modelling, but has comparable spatial data modelling and operations to Oracle. It is by far top of its game in statistical, logistical, rendering and map generation.

5.3. Combination of GIS and Database

Oracle Database and MapInfo Professional

5.3.1. Background

Traditionally GIS have used databases like Oracle and Microsoft SQL Server as storage of spatial data whilst the spatial analysis is carried out using the tools of a GIS (like MapInfo) and rendering of data on map planes.

5.3.2. Credentials

This is a tried and tested combination from two leaders in their respective markets. A “no frills” approach to functionality.

5.3.3. Method

i. The spatial data can be manipulated and stored in the GIS with any relational conventional data being manipulated and stored in the DBMS
ii. The spatial data can be manipulated in the GIS and stored as a spatial object along with its conventional data in the DBMS
iii. The spatial and conventional data can be manipulated together in the DBMS and rendered using the GIS

5.3.4. Strengths

The combination has all the power and functionality of the Oracle database with the rendering capabilities of MapInfo.

5.3.5. Weaknesses

If there was a requirement to web enable any application then this combination would not be suited – MapInfo is only now developing web publishing capabilities and it is difficult for any automatic processing to be initialised. MapInfo is effectively designed for use as a planning tool and heavily relies on manual operation. To create an advanced application could prove difficult due to the distributed nature of the functionality between packages.

5.3.6. Opportunities

The combination of the two platforms makes it possible for numerous applications, where each platform has its own strengths, as discussed in each of the previous relevant sections.

5.3.7. Threats

These are two separate systems which give rise to potential incompatibility issues and the link between the two is a potential single point of failure. There is also the requirement for any user or developer to be fully trained on both applications.

5.3.8. Comparisons

As a combination the two applications work, but nothing above and beyond in comparison to using them from a standalone perspective. It can be just too complex.

5.3.9. Summary

It entirely depends on the type of application for the data's intended use for whether or not this approach is suitable.

5.4. Mathematical Analysis

It is possible to ignore spatial data as objects on which analysis can be performed; an algorithmic approach can be just as accurate in achieving the aims.

5.4.1. Background

Using a combination of trigonometry and Pythagoras theorem it is possible to simulate spatial analysis.

5.4.2. Credentials

There is less reliance on vendor products and increased ability to develop an application from scratch.

5.4.3. Method

To create a procedural application that subjects location data to mathematical functions and formulae for spatial manipulation

5.4.4. Strengths

This method of approach is effectively open source – allowing for the mathematical spatial analysis to be fine tuned.

5.4.5. Weaknesses

The performance overheads to manage the processing of the functions' transactions are very high. To be realised as an application, a development platform is required. The code required to develop and manage all permutations is likely to be quite complex and detailed. There is also a lack of temporal support.

5.4.6. Opportunities

For would be programmers and mathematicians to delve into the mathematical mechanics of spatial and temporal modelling and analysis.

5.4.7. Threats

Additional elements of spatial data, such as vehicle heading, would need to be available from data sources to aid the accurate mathematical analysis. Further to this if we were to use the spatial metric of directional heading there is the 0 divided by 0 heading issues that would need to be solved.

5.4.8. Comparisons

This approach cannot seriously be compared to the other solutions. It is a concept that lacks detail in the areas of data storage, modelling and analysis.

5.4.9. Summary

Using formulae and calculation is a plausible solution but a very labour intensive approach to develop and achieve the same desired results which can be achieved from the other approaches discussed earlier in this section. It can severely be hindered by performance due to the transactions that need to take place. Overall it is not a feasible solution.

6. Experimental Research

The research undertaken in this section will demonstrate the feasibility of developing a spatial decision system based on requirements from the bus industry; and to put it into context with other similar technologies and solutions.

6.1. Requirement

The aim of service control is to regulate the buses in order to provide the bus passenger a reliable and regular bus service, like in Fig 6.1.1.

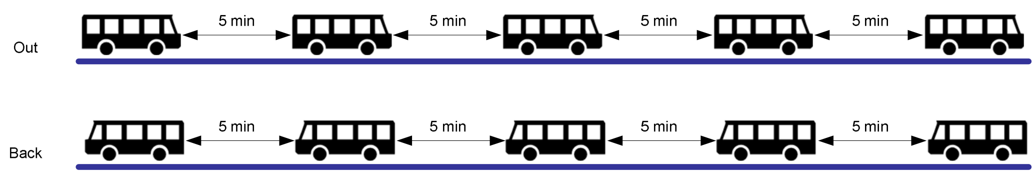


Fig 6.1.1

Reliability of buses in London is influenced by many external factors, including, but not limited to: passenger boarding times and behaviours [32], traffic congestion, on-board incidents and diversions. Yet even though some of these factors are inherently unpredictable, some are measurable in negating their affect. Of the list – the heavily influential contributing and / or resulting factor is traffic congestion – which is possible to measure.

The current AVL system (Fig 6.1.2) and indeed other similar applications offer more reactive (present status of a bus) than pro-active preventative functionalities. The lack of visualisation from a geographic and wider perspective is also evident.

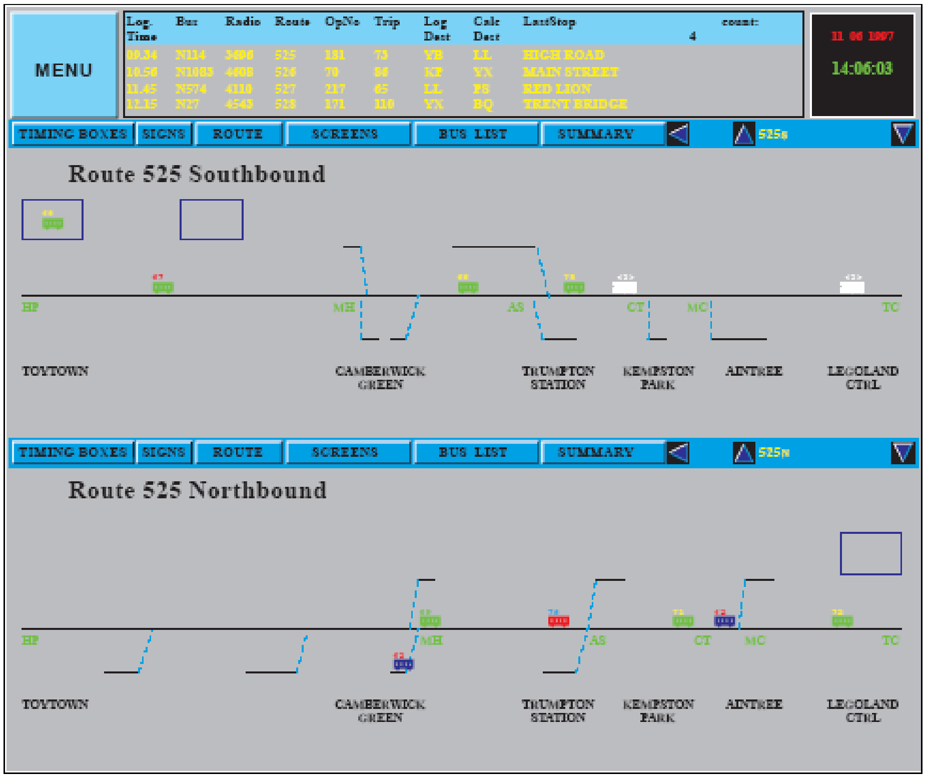
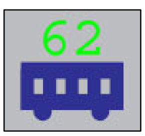


Fig 6.1.2



The colour of the number above each bus indicates its adherence to the schedule in real-time. [Green for Late / Yellow for On-time / Red for Early]

Situations that need to be avoided are buses bunching causing large gaps in service. The EWT[5] (Excess Waiting Time) [33] is the metric used by TfL to police the performance of bus routes operated by private companies.

If we take the following scenarios:

[5] EWT – Excess Waiting Time: the excess amount of time a passenger had to wait on top of the scheduled waiting time.

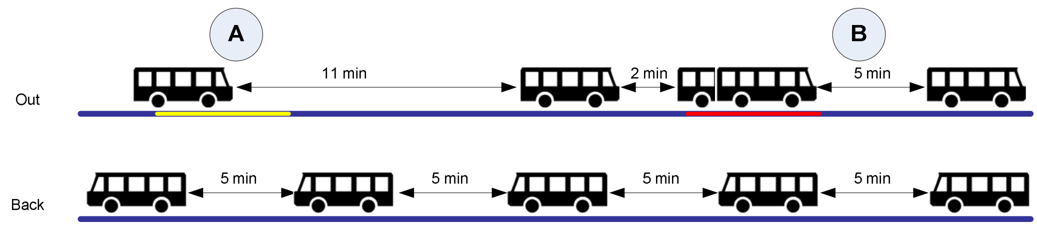


Fig 6.1.3

In Fig 6.1.3 – at point (A) there is medium congestion, which is not posing a risk to the route at present. Prior to point (B) there is heavy congestion which is causing buses to bunch – the waiting time at point (B) will be increased by an unknown amount. If the availability of accurate and precise congestion information was put to good use, in a timely manor, then preventative action can be taken, like in Fig 6.1.4.

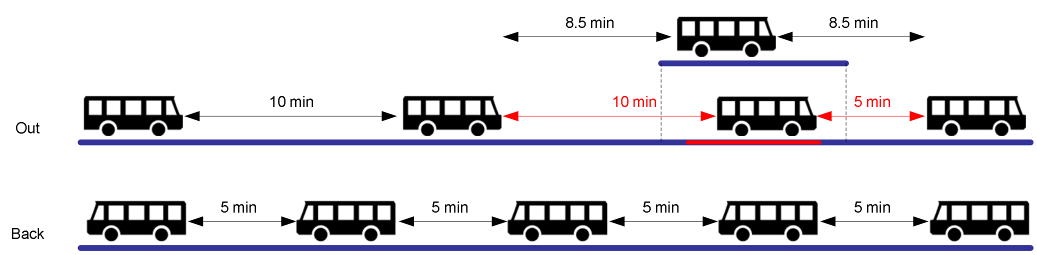


Fig 6.1.4

In Fig 6.1.3 – the second bus which was at point (B) could be diverted, like in Fig 6.1.4, well in advance of it approaching the heavily congested area – leaving the first bus essentially isolated from the route. The headway between buses as a result is increased in the 'Out' direction due to the missing bus and the extra time taken to traverse the diversion routeing. Although the headway increase between buses causes the Estimated Waiting Time (EWT) to be increased for passengers at point (B) – a frequency is maintained and the EWT is maintained within acceptable parameters.

Again, having the information available can help the following situation in Fig 6.1.5.

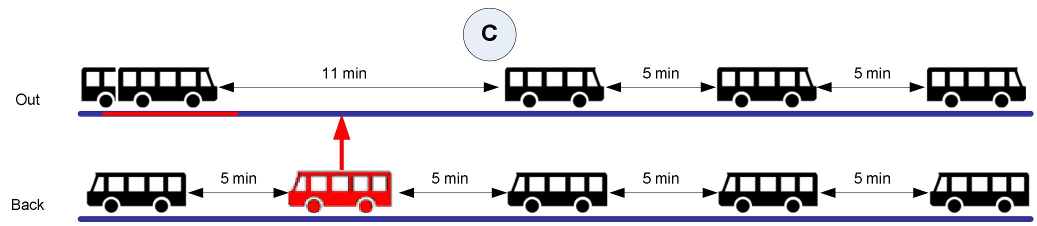


Fig 6.1.5

Heavy congestion at the start of the route is preventing buses from making progress into the route. At point (C) the EWT will have increased due to the headway of 11 minutes. The availability of accurate and precise congestion information can help the situation in the following way in Fig 6.1.6.

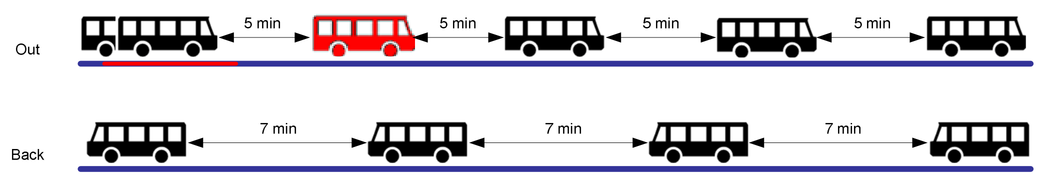


Fig 6.1.6

In this situation the end of the route can be isolated by turning buses early from the other direction – instead of allowing all the buses to eventually bunch at one end. The same headway is then maintained for 80% of the route (damage limitation).

As part of the iBus project, all buses in the 8,000 vehicle fleet are scheduled to be installed with GPS location equipment from 2007 to 2010 [6]; this raises the potential to use the data (to be received) to greater affect than its intended application. From a requirements gathering exercise point of view, any use of GPS location data would therefore have to be pro-active in modelling and highlighting potential network problems at the earliest possible time and from a geographically perspective.

Any proposed application must therefore demonstrate the ability to solve the issues discussed here and meet the requirements – which are depicted in the following ‘rich picture’, see Fig 6.1.7. A rich picture is a technique used in information systems development – as part of the Soft Systems methodology [34].

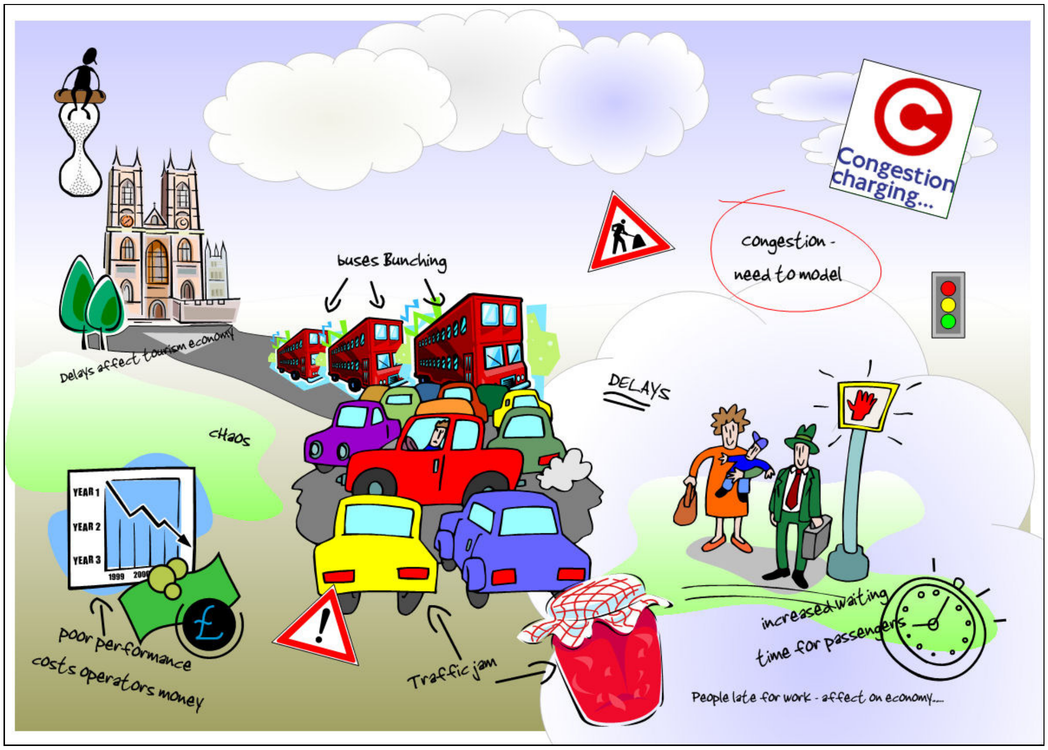


Fig 6.1.7 – A Rich Picture technique depicting system requirements

6.2. Approach

When London Buses first sought a replacement for the current AVL system – one of the main objectives was to move to a technology with greater accuracy and less infrastructure maintenance costs than the current road-side tracking Beacon based system. The popularity and development of GPS technology over the last two decades made it a strong candidate to meet the primary replacement objectives. However, GPS is still vulnerable to the affects of signal dropouts and urban canyoning – we will discuss the affects of this on the research later in this section. And in a city like London the vulnerability of the GPS signal is potentially very high. This prompted London Buses to undertake trial GPS location surveys in 2005 before explicitly stating a preferred technology in the Invitation to Tender (ITT) for a replacement system.

I am using the GPS location data from one of the trials conducted on Tuesday, 15th March 2005, 15.00hrs to 17.00hrs (Baker Street to Norwood) to use in my prototype application. (See Appendix A)

As one of the strongest technologies around, I have selected the use of Oracle spatial to develop an 'active database' prototype for the spatial matching analysis of the vehicle locations to the Ordnance Survey spatial Integrated Transport Network (ITN) links [35]. And I have selected MapInfo to render the results against geographical mapping. The advantages of these products meet the basic requirements for spatio-temporal data aggregation, interpolation and rendering.

The intended method is to:

(1) Create the database tables,
(2) Create the indexes
(3) Create the spatial query *Within Distance*
(4) Create the packages and triggers

(5) Insert the transit links

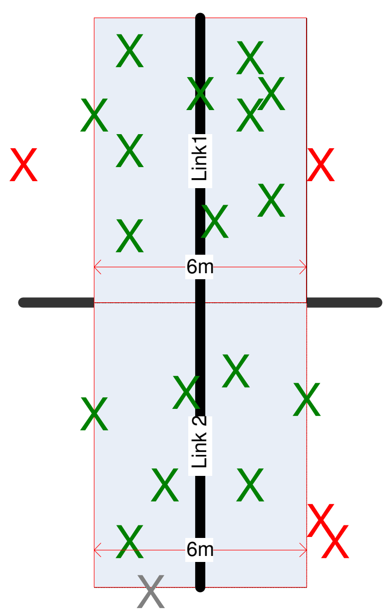
(6) Insert the GPS values

(7) *The automation of the application occurs*

(8) Export the results to MapInfo

(9) Create a Thematic Map

The data has been normalised to the 2NF (normal form) for this prototype application. I have selected to concentrate on survey data along the 'one way' roads of Baker Street, Portman Square (east side) and Orchard Street. The significance of selecting this one way system will become apparent when discussing the issues and limitations of an application.



Each ITN link and GPS fix record has an associated spatial object (SDO_GEOMETRY data structure) stored in the Oracle database.

The Oracle spatial query '*Within Distance*' works by effectively creating a buffer zone of *xx* metres around each ITN link object. And then selecting all GPS fixes objects within the buffer. This query can be created as a view.

From the example: Link 1 has 10 spatial matches and Link 2 has 7 spatial matches.

A caveat of the results published is that bus speeds (taking into account acceleration and deceleration) are approximately 20% less than that of the conventional road vehicle which was used in the trials, so the expectation in a prototype using the bus GPS data is that the speed figures will be skewed slightly lower.

6.3. Results

Following the initial setup of the active database application the GPS location values were inserted into the 'location' table. (SQL scripts and Screen shots can be found in Appendix B and C.)

SQL> start insert_values.sql

The insertion of values causes the triggers (below) to fire on the 'location' table and then subsequent triggers to fire.

```
CREATE OR REPLACE TRIGGER location_bi
BEFORE INSERT OR UPDATE ON location
BEGIN
        state_pkg.newRows := state_pkg.empty;
END;

CREATE OR REPLACE TRIGGER location_aifer
AFTER INSERT OR UPDATE OF loc_id ON location FOR EACH ROW
BEGIN
        state_pkg.newRows( state_pkg.newRows.count+1 ) := :new.rowid;
END;

CREATE OR REPLACE TRIGGER location_ai
        AFTER INSERT OR UPDATE OF loc_id ON location
DECLARE
        BEGIN
        FOR i in 1 .. state_pkg.newRows.count loop
        INSERT INTO link_loc (link_loc_toid,link_loc_id)
        SELECT  K.link_toid, L.loc_id
                FROM location L, link K
                WHERE     SDO_WITHIN_DISTANCE(K.link_geom,     L.loc_geom,
                'DISTANCE=6 UNIT=METER')='TRUE'
                AND L.rowid = state_pkg.newRows(i);
        END loop;
END;
```

The trigger automatically carries out the spatial matching of the GPS location spatial objects to a transit link spatial object in the 'link' table. The results are inserted into the 'link_loc' table – which causes the next trigger to fire. A summary of the spatial matching trigger is shown in Table 6.3.1.

Link Sequence	**Road Name**	**Link ID (OS TOID)**	**Spatial Matches**
1	Baker Street	430217380	50
2	Baker Street	430217379	53
3	Baker Street	430407949	4
4	Baker Street	430217386	13
5	Baker Street	430473662	4
6	Baker Street	430324906	68
7	Baker Street	430217372	16
8	Baker Street	430217373	13
9	Baker Street	430135703	17
10	Baker Street	430217156	65
11	Baker Street	430324655	61
12	Baker Street	430217163	22
13	Portman Square	430418631	5
14	Portman Square	430343528	6
15	Portman Square	430418630	3
16	Portman Square	430343534	4
17	Orchard Street	430343529	4
18	Orchard Street	430459115	2
19	Orchard Street	430147230	1

Table 6.3.1

The next trigger that fires on the 'link_loc' table, selects the average speed of each transit link based on the speeds of the matching locations to each transit link. The results are inserted into the 'link_speed' table, which is summarised in Table 6.3.2.

```
CREATE OR REPLACE PACKAGE state2_pkg
AS
        type ridArray is table of rowid index by binary_integer;

        newRows         ridArray;
        empty           ridArray;
END;
```

```
CREATE OR REPLACE TRIGGER link_loc_bi
BEFORE INSERT OR UPDATE ON link_loc
BEGIN
        state2_pkg.newRows := state2_pkg.empty;
END;

CREATE OR REPLACE TRIGGER link_loc_aifer
AFTER INSERT OR UPDATE OF link_loc_id ON link_loc FOR EACH ROW
BEGIN
        state2_pkg.newRows( state2_pkg.newRows.count+1 ) := :new.rowid;
END;

CREATE OR REPLACE TRIGGER link_loc_ai
        AFTER INSERT OR UPDATE OF link_loc_id ON link_loc
DECLARE
        BEGIN
        FOR i in 1 .. state2_pkg.newRows.count loop
        INSERT INTO link_speed(link_speed_toid,link_speed_avg_speed)
        SELECT k.link_toid, avg(l.loc_speed)
                FROM link_loc c, location l, link k
                WHERE c.link_loc_id = l.loc_id
                AND c.link_loc_toid = k.link_toid
                AND c.rowid = state2_pkg.newRows(i)
        GROUP BY k.link_toid;
        END loop;

END;
```

A package containing three triggers has had to be created for each table event in Oracle due to the way in which Oracle writes data to the tables. The trigger action cannot be successfully be implemented as the insert of the data has not yet been committed to the database. In order to get around this problem, a package of triggers has to be created to record the state of the modified table; 1) the first trigger resets the table state array before any insert or update of the table takes place, 2) the second trigger records the id’s of the rows in the table being modified, 3) the third trigger processes the trigger action on the modified rows.

Link Sequence	Link ID (OS TOID)	Average Speed	Spatial Matches
1	430217380	8.23	50
2	430217379	7.81	53
3	430407949	23.00	4
4	430217386	20.31	13
5	430473662	15.75	4
6	430324906	4.40	68
7	430217372	19.76	16
8	430217373	30.93	13
9	430135703	32.11	17
10	430217156	4.76	65
11	430324655	4.06	61
12	430217163	12.27	22
13	430418631	27.83	5
14	430343528	37.50	6
15	430418630	41.00	3
16	430343534	43.25	4
17	430343529	41.60	4
18	430459115	41.00	2
19	430147230	32.60	1

Table 6.3.2

Using MapInfo, the spatial matching average speed results (summarised in Table 6.3.2) (and transit link spatial object) have been exported to create a thematic map to represent the results in Fig 6.3.1. The thematic average speed intervals have been set at 5 mph at the lower end of the scale to increase the precision of the results presented, given that the average speed in London ranges between 9 mph and 12 mph [36]. This enables developing problems to be observed quicker.

This stage of the application cycle could easily be semi-automated using the MapInfo “map DBMS table” tool, or even fully automated in a Java application, that can be web enabled against mapping using the Google Maps Application Programming Interface (API).

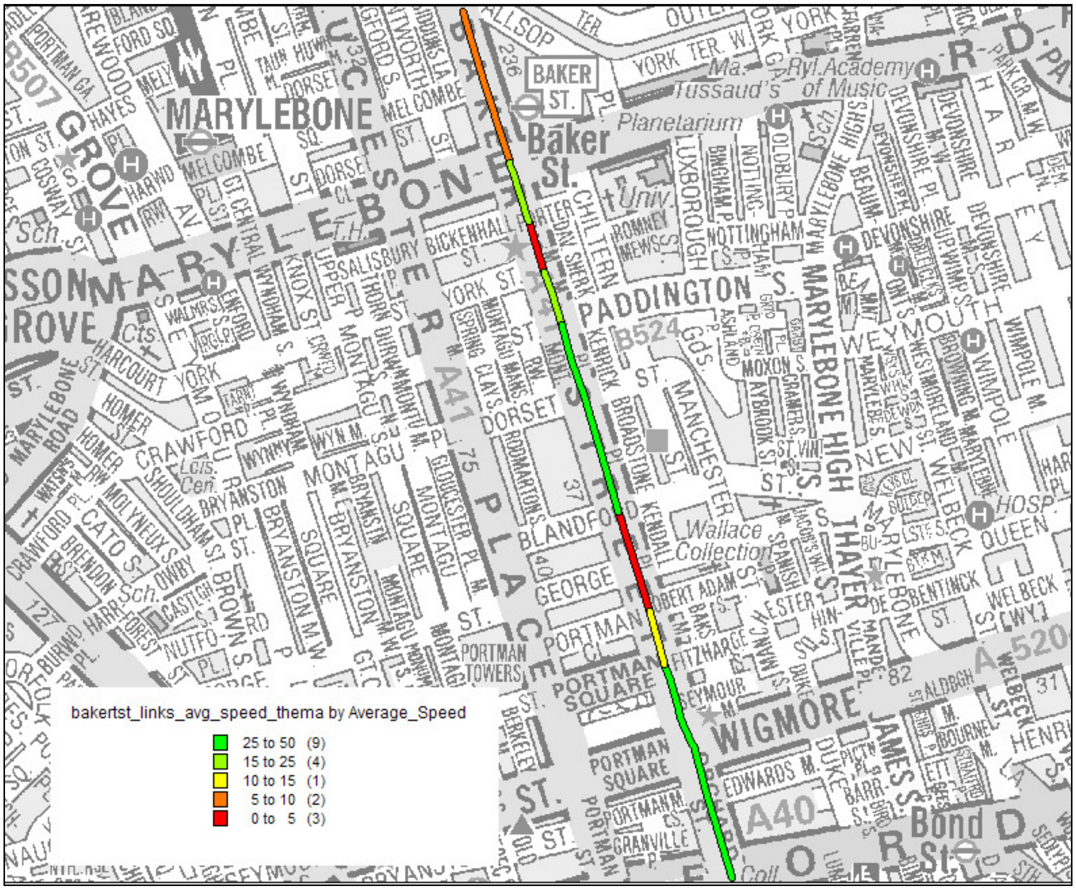


Fig 6.3.1

To demonstrate the change in the average speed of each transit link as and when more data is received – another set of location attributes, for one trip along the same stretch of road[6], are inserted and processed in the data model. The results and changes to the transit link average speeds are summarised in Table 6.3.3 and the change to the geographical representation can be seen in Fig. 6.3.2.

SQL> start insert_values_2.sql

[6] The trip data is manipulated from the original data set (Speed = + 7mph) in order to simulate a change in the average speed.

Link Sequence	Link ID (OS TOID)	Average Speed	Spatial Matches
1	430217380	11.53	102
2	430217379	11.13	102
3	430407949	26.50	6
4	430217386	23.86	22
5	430473662	20.5	4
6	430324906	7.71	134
7	430217372	22.50	30
8	430217373	35.58	26
9	430135703	35.50	34
10	430217156	7.45	130
11	430324655	30.50	16
12	430217163	14.98	42
13	430418631	32.25	8
14	430343528	40.75	8
15	430418630	44.50	2
16	430343534	46.75	8
17	430343529	43.75	8
18	430459115	44.50	4
19	430147230	36.10	2

Table 6.3.3


bakertst_links_avg_speed_thema by Average_Speed
25 to 50 (11)
15 to 25 (3)
10 to 15 (3)
5 to 10 (2)
0 to 5

Fig 6.3.2

The research application is based on the conceptual schema in Fig 6.3.3. Within the schema, the 'location' and 'link' entities are geospatially dimensional; and the 'location' and 'avg_speed' entities are temporally dimensional, these added dimensions enable the data to accurately model the behaviour of the bus (transport) network.

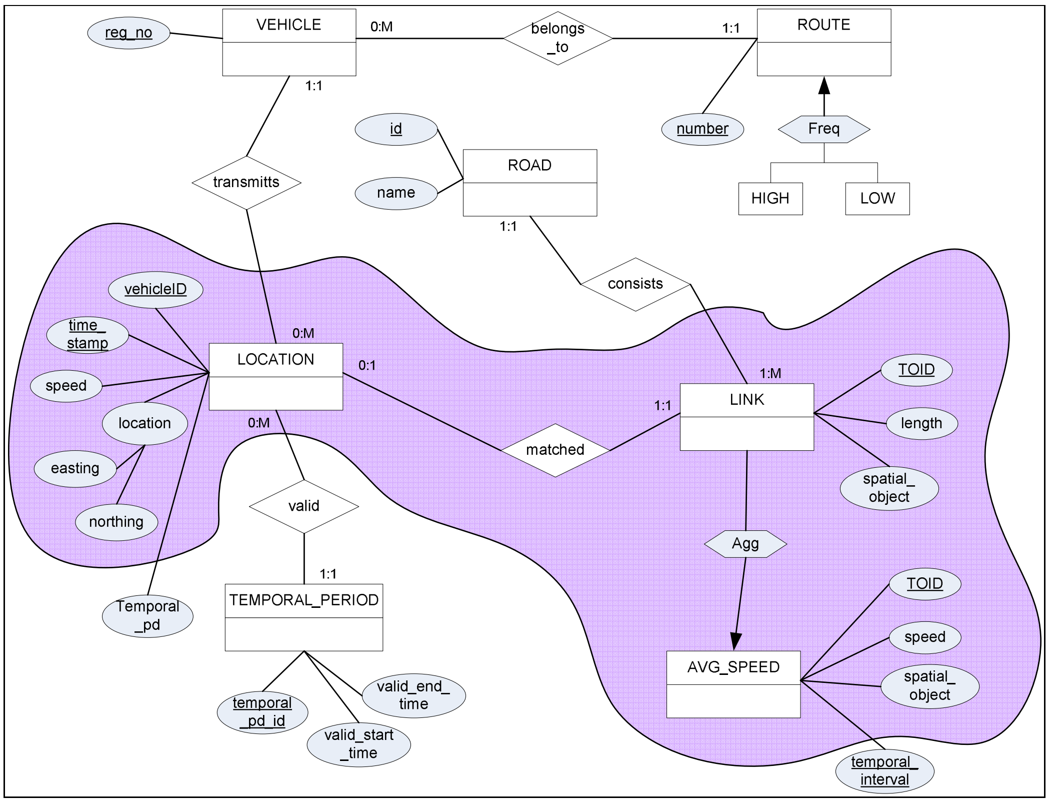
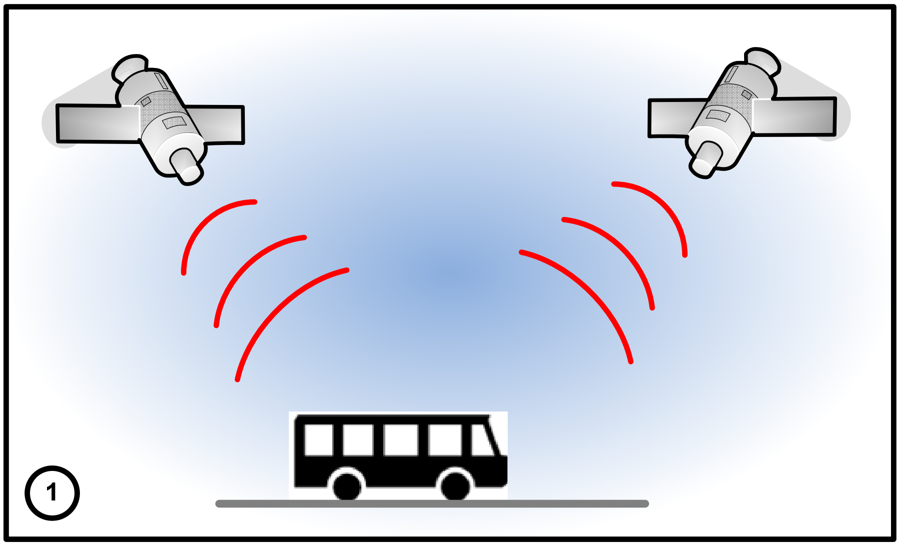


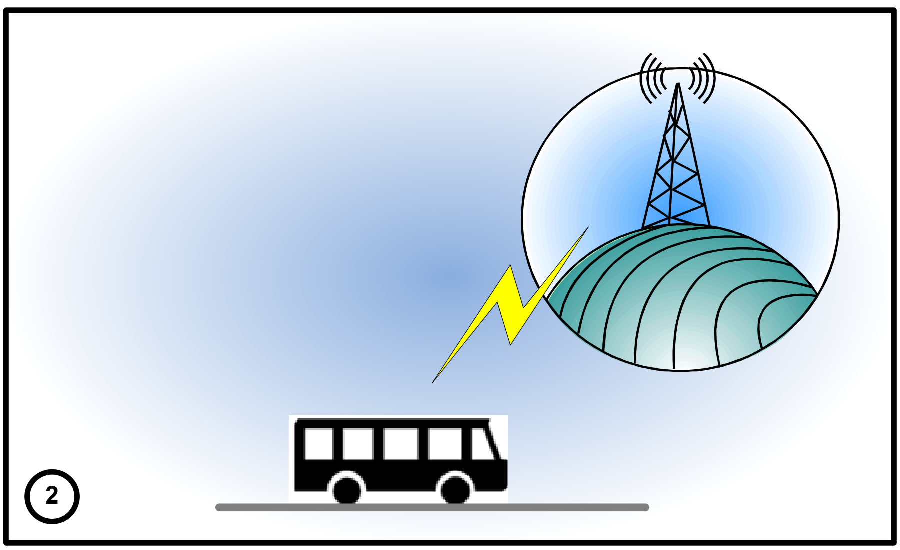
Fig 6.3.3

The shaded area of the conceptual schema highlights the extent of the implementation of this research application.

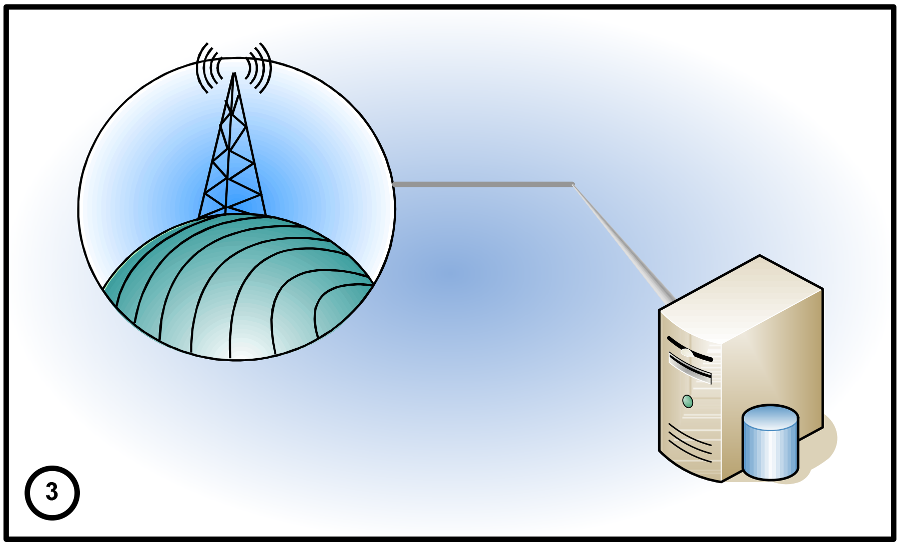
A fully working application would involve the entire extent of the proposed conceptual schema. The following diagrams visually describe how the full application would function.



The bus is constantly getting a spatial and temporal location GPS fix from the satellites in sight.



The bus communicates its GPS location, timestamp, speed and other attributes over the GPRS data carrier.

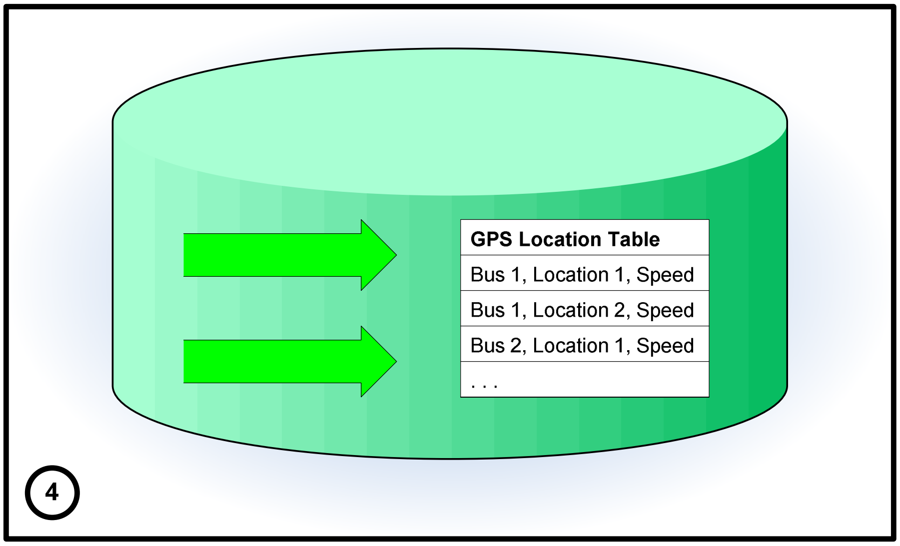


The data is received by the nearest transmitter and sent via a Wide Area Network (WAN) to the database server.

At this point we should explain the concepts of the active database and triggers. We can confidently describe our application as a real-time application – where the data inputs and outputs are dynamic and changeable. And it is the continuous processing of the data triggered by events within the database that change the data. Triggers are effectively a set of database actions that respond to active rules triggered by certain events. With this in mind we have the basis of an active database.

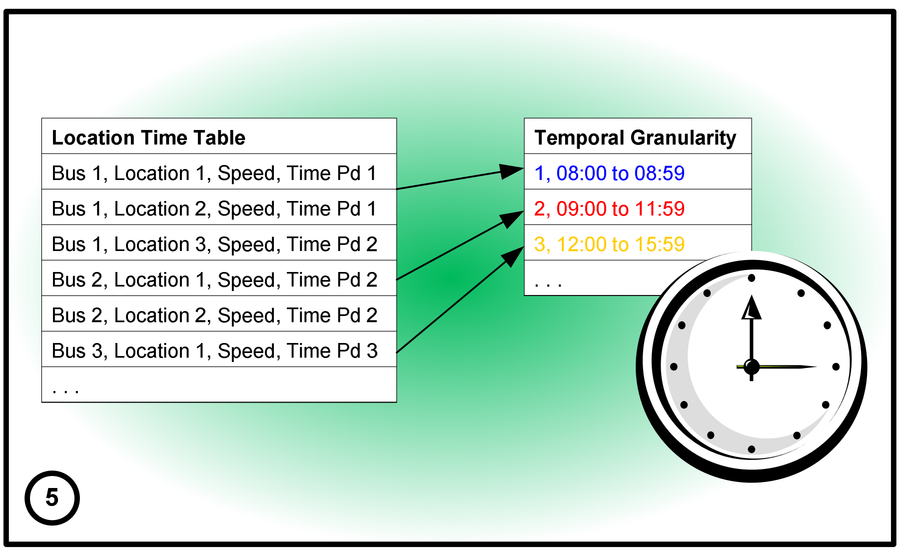
The general model proposed for active databases is referred to as the Event-Condition-Action (ECA) model, due to its three main components [10], defined below in-relation to this research application.

- Event – the cause which triggers the rule (Schematics 1 to 4)

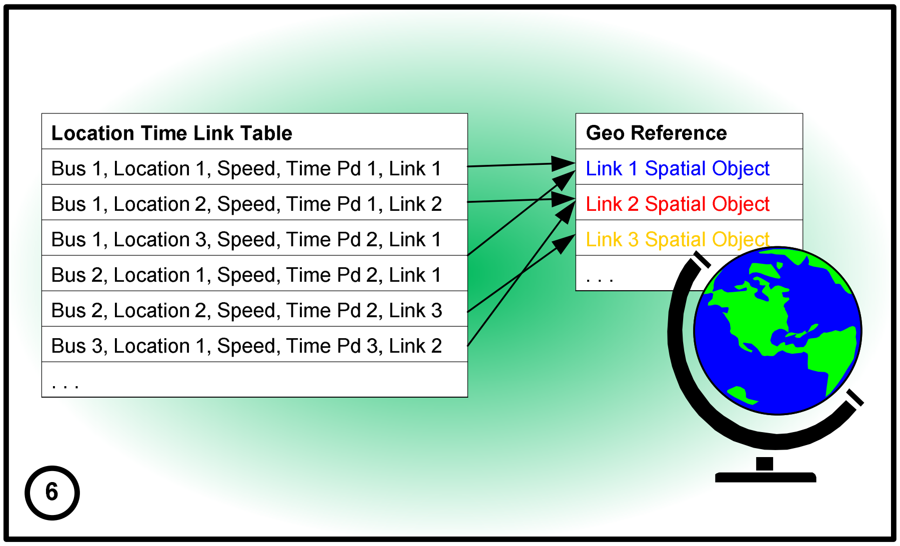


The raw location data is inserted into the database as a new record

- Condition – determines the type of action to be executed

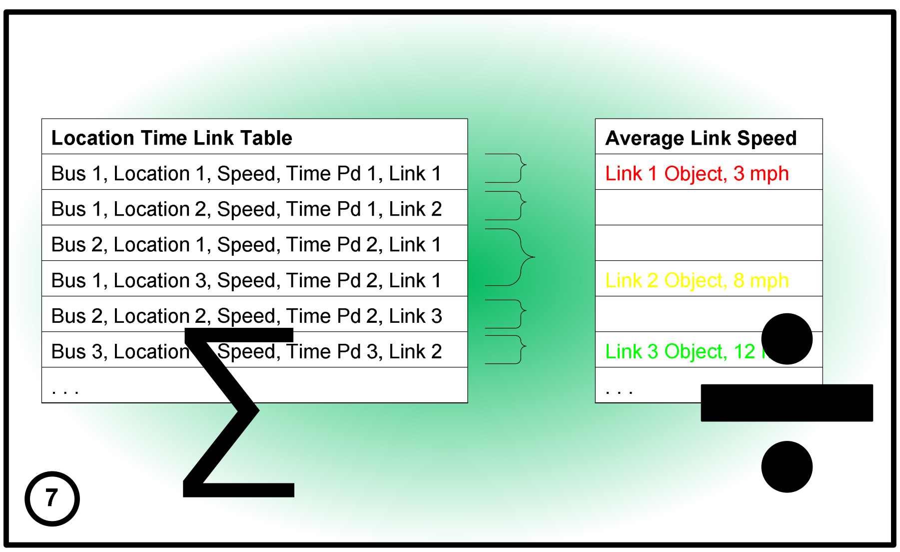


The timestamp of the location data is assessed for its validity against pre-defined temporal periods.

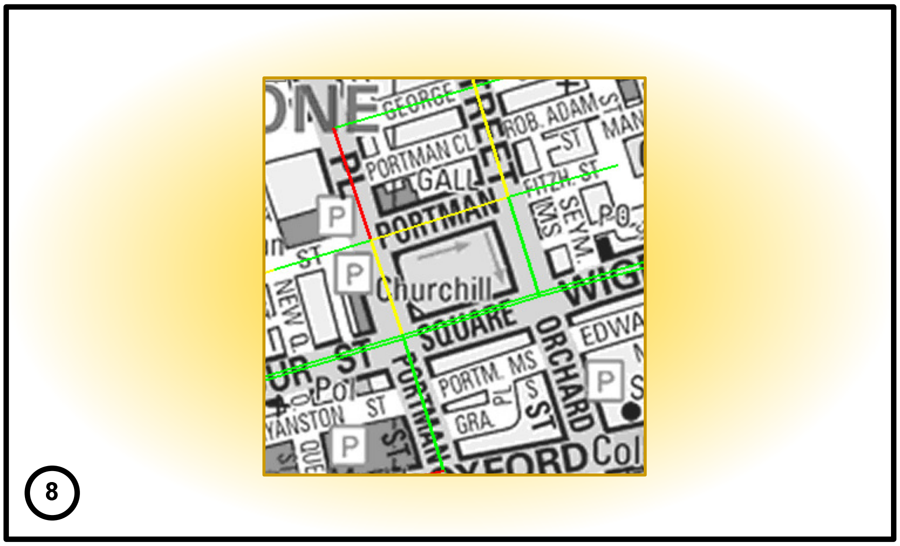


Each GPS location record is the matched to a spatial link in the geo-spatially enabled OS Integrated Transport Network (ITN).

- Action – the sequence of SQL statements to be executed



Aggregation for each link is carried out based on the average speed of the buses within a temporal time window.



Each spatial link in the transport network is colour coded based on the average speed of buses over that link and rendered against familiar geographical mapping.

6.4. Technology Appraisal

There are systems within the London Road Transportation network that already determine delay metrics: like congestion. The SCOOT [38] loop detectors and other loop detectors are fixed points in the road network; the Congestion Cameras are fixed locations on the boundary of zones.

TfL have been one of the major public transport authorities in the world to use these ITS – metrics from loop detectors and speed cameras are currently used to model performance of the London road network using GIS internet technology to render the results. This application is called COMET [39] (See Fig 6.4.1) and is a real-time map-based display providing early warning of congestion.

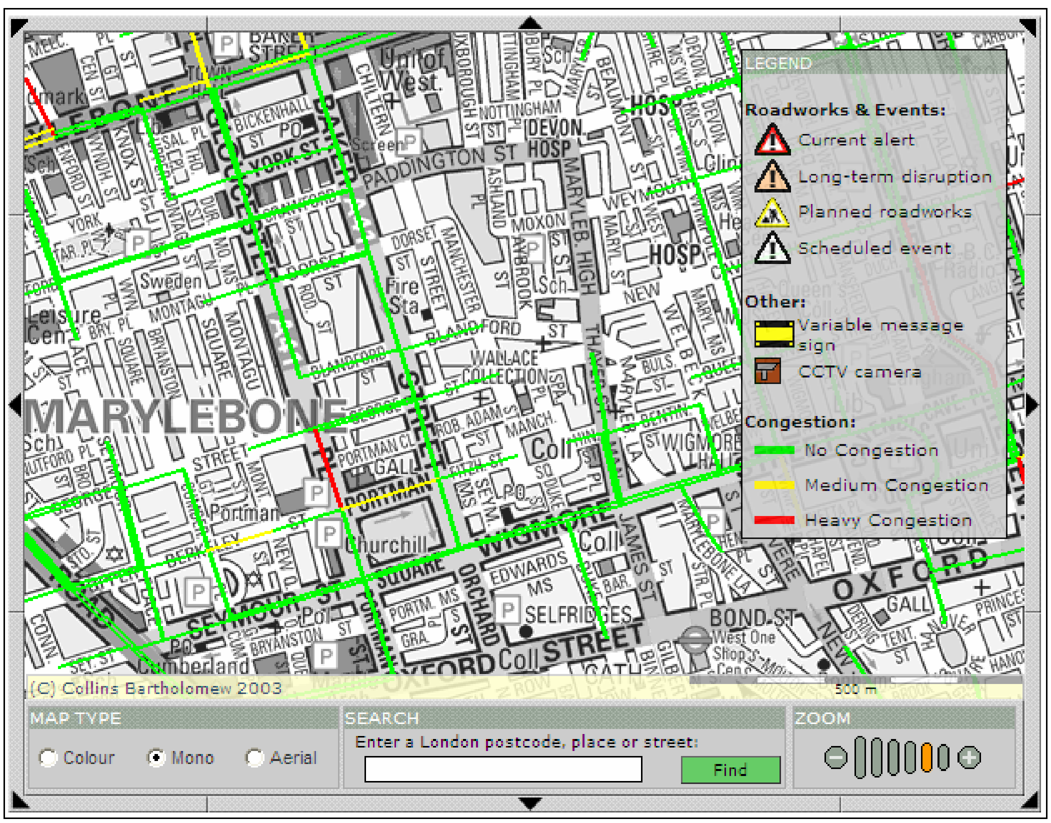


Fig 6.4.1

Spatial

The measurements from the detectors are however at fixed locations (normally traffic signalled junctions) around the Greater London road network and coverage is not extensive. Therefore Fig 6.4.1 is not strictly representative of the road conditions along the entire road length(s). To define this data as spatially enabled needs to be put into context. The London Buses iBus system (Automatic Vehicle Location) has the potential to use GPS data from each bus with a satellite fix every *nn* seconds or by defining a geo trigger point at any position – across the road network. The use of iBus spatial data to form accurate representation of road conditions at any point along the road network is therefore extremely beneficial.

Fig 6.4.2 depicts the data availability that could be *expected* from a fixed-point measurement system (loop detectors) compared to a dynamic point measurement system (Bus GPS). The collection of network performance data is of a greater spread spatially in the dynamic point system – whereas the fixed point system is limited to the spatial points that it can collect data.

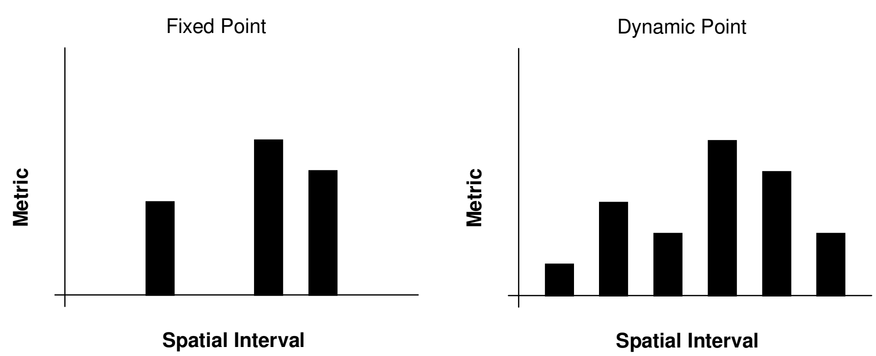


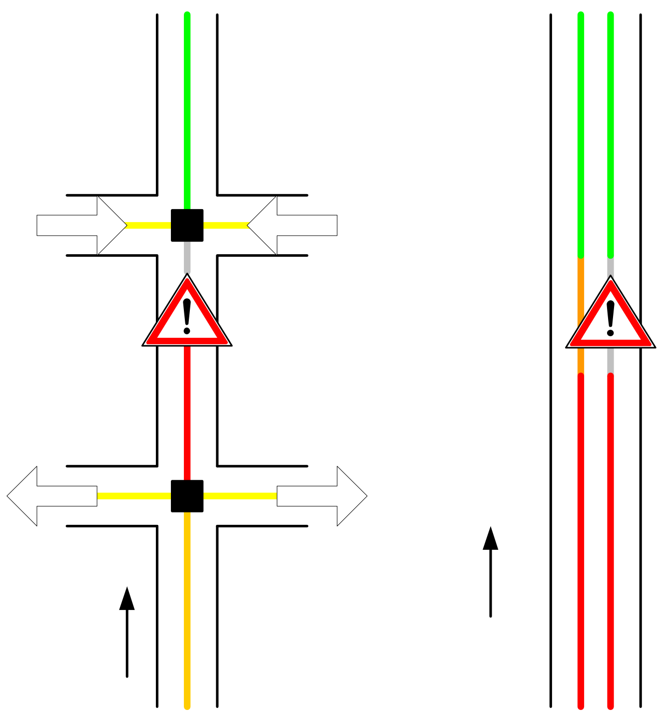
Fig 6.4.2

Fixed point system collects data from *all* vehicles but is limited to certain points (loop detectors) on the road network. The dynamic point system collects data only from buses but at any spatial interval across the road network, but limited to the bus network, and consequently there are fewer bus journeys made compared to private vehicle journeys made in London. Table 6.4.1 compares the two measuring methods.

	Loop Detectors	**GPS**
Method of Collection	Fixed	Variable
System Coverage	Low	Medium
Volume of Data	High	Low

Table 6.4.1

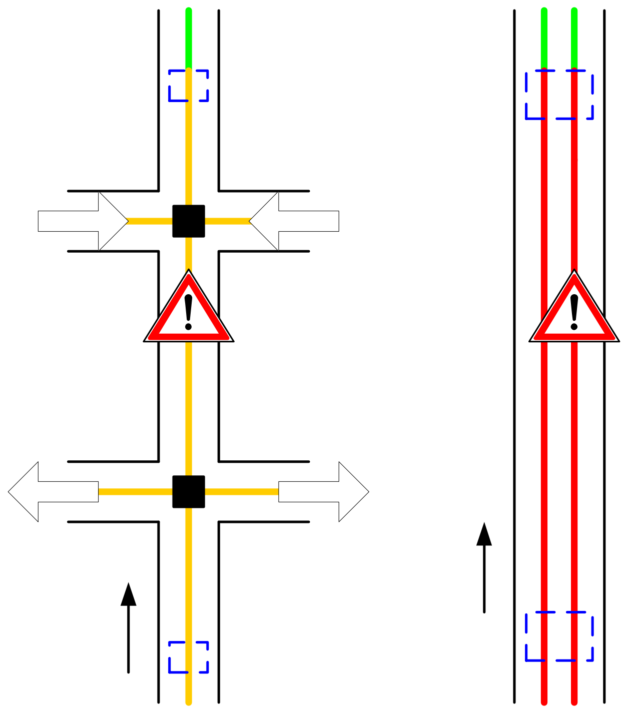
I would theorise that Loop detectors may give more accurate readings over roads where there are limited entry and access points, like in Fig 6.4.4(b); but in a city like London there is a high density of junctions for entry and exit to parts of the road network, thus making the traffic conditions local to every road segment – this can be modelled as follows in Fig 6.4.3(a).



(a) Fig 6.4.3[7] (b)

[7] For diagrammatic clarity the curvature that forms the different link geometries have been straightened.

If we were to use the GPS location data to model the congestion, we would expect to see the above results in Fig 6.4.3(a) & (b). The road layout diagram Fig 6.4.3(a), allows traffic to freely exit and divert around the incident. Traffic in the affected link is stationary and unable to turn-around which is indicated by the link colour, but other links, although marginally affected by the diverting traffic, have movement and are indicated accordingly.

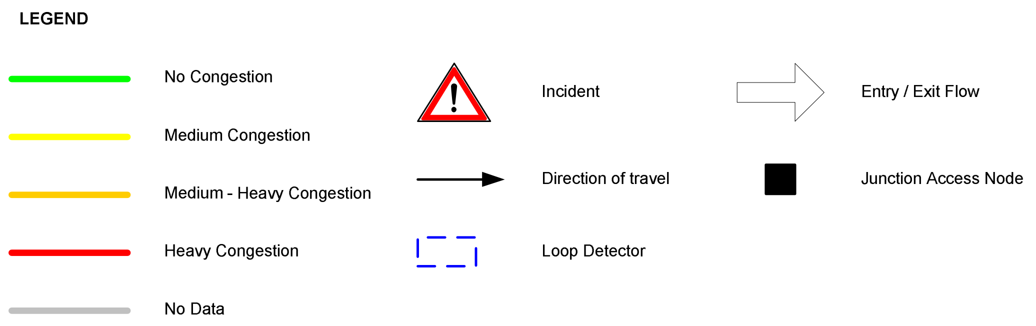


(a) Fig 6.4.4[8] (b)

Compare this to use of fixed loop detectors – the status of each link is determined by the sensors in the road, assigned to a particular stretch[9] of road, detecting the presence of vehicles and measuring their stationary periods.

[8] For diagrammatic clarity the curvature that forms the different link geometries have been straightened.

[9] The assignment is not down to the transit link level, but at a higher grouping.



If we use our real-world example to demonstrate this point, table 6.4.1 shows the number of loop detectors along the road network used in this research.

Link Sequence	Link ID (OS TOID)	Length (m)	No. of Loop Detector Sets	Lane Coverage	Spatial Matches[10]
1	430217380	97.12	1	1 / 3 (Bus Lane)	50
2	430217379	104.04	2	3 /3	53
3	430407949	12.71	0	-	4
4	430217386	62.93	1	3 / 3	13
5	430473662	4.12	0	-	4
6	430324906	67.05	0	-	68
7	430217372	70.21	4	3 / 3	16
8	430217373	98.70	1	3 / 3	13
9	430135703	151.52	2	3 / 3	17
10	430217156	79.69	2	3 / 3	65
11	430324655	53.16	2	3 / 3	61
12	430217163	71.56	0	-	22
13	430418631	18.02	1	3 / 3	5
14	430343528	36.89	2	3 / 3	6
15	430418630	17.22	0	-	3
16	430343534	39.39	0	-	4
17	430343529	40.52	1	3 / 3	4
18	430459115	15.24	1	3 / 3	2
19	430147230	117.35	0	-	1
	Totals	1157.44	20		411

Table 6.4.1

At the transit link level – link TOID 430473662 and 430324906 have no loop detectors over their entire length thus it will either present no data or is reliant on loop detectors from the previous link (430217386) to attempt to represent

[10] The spatial occurrences matched are from one trip

the traffic conditions – but given there is two junction nodes for this link, there is the highest probability that conditions in the TOID links are unlikely to be the same as that in the previous link(s). Especially as from Fig 6.4.5 we can see that the northern junction node with Bickenhall Street / Porter Street allows traffic to flow into and out the link segments undetected, just as the southern junction node with York Street allows traffic to flow out of the link segment undetected. [Loop detectors represented in Fig 6.4.5 by "X"]

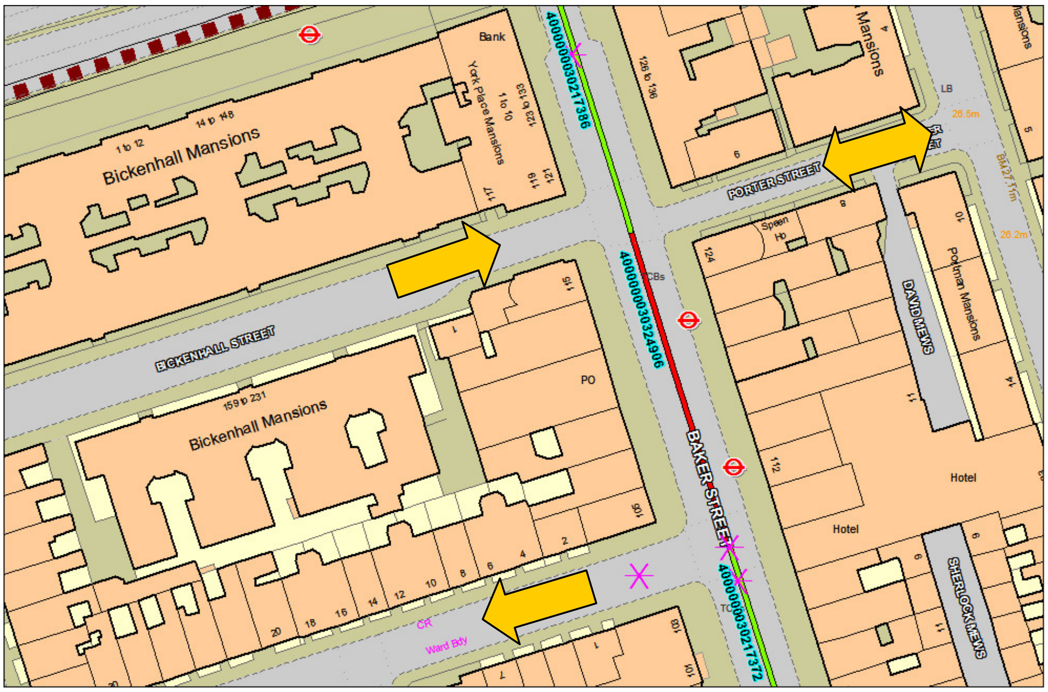


Fig 6.4.5

Yet when we superimpose our results (Table 6.3.2) on the map in Fig 6.4.5 we can get a fairly accurate representation of the traffic conditions, as GPS spatial matching approach is not reliant on fixed loop detectors being present and takes into account traffic entering or exiting the segment. It should be noted that the link speed representation for link 430324906 is below 5mph – which if the trial data was conducted using a service Bus – could be down to either of the Bus Stops located on that link. However, the trials were carried out using a normal road vehicle, but even if we were modelling the dwell time at the bus stop within link segments – it is still a useful dimension (information facts) to have in a decision support application for the bus industry. This is discussed further, as a benefit, in the Issues and Solutions section.

And the final conclusion on the use of the loop operation compared to the suggested GPS operation as a monitoring approach would be the fundamental drawback in that they are easily damaged by road-works and require constant checking to ensure efficiency of the traffic control and monitoring at the site [40].

Temporal

The provision of temporal spatial data at non-fixed points across the road network gives rise to measurement of delays (congestion / time lost) on an even greater local scale. The temporal dimension is very important in our model. Conventional stores of business data are very much static and have no real requirement for the complex modelling of time. The standard database concept from a temporal perspective is for the record to be time-stamped once at a single point in time and/or for the attribute to be continually overwritten – erasing any historical temporal data charting changes through the records life span. In order for the application of real-time analysis there needs to be a temporal history of each object (bus) to attempt accurate calculation and representation.

Introducing the temporal element of granularities to the spatial data model not only enhances the accuracy of the information but also improves database performance. Query optimisation is important from a performance perspective, so the sooner larger amounts of data can be disregarded, the less data that has to be searched through to meet the query results. The temporal granularities, in this application, validate what are potentially 23 million rows of data transactions down to a maximum of 2.8 million transactions during any one period. This is calculated by:

t = (s(M)m(H)h(D)/f)n

= 60 * 60 * 24
= 86,400 / 30
= 2,800 * 8,000
= **23, 040, 000**

Where

t = no. of tuples; s(M) = seconds in a minute; m(H) = minutes in an hour; h(D) = hours in a day; f = polling interval; n = maximum no. of buses;

If we introduce a temporal granularity interval of a maximum of 3 hours, we can significantly reduce the amount of database transactions:

t = (s(M)m(H)g(i)/f)n

= 60 * 60 * 3
= 10,800 / 30
= 360 * 8,000
= **2, 880, 000**

Where

t = no. of tuples; s(M) = seconds in a minute; m(H) = minutes in an hour; f = polling interval; n = maximum no. of buses; g(i) = temporal granularity interval (hours)

6.5. Issues and Solutions

Any fully working application will need to overcome the highlighted issues and problems identified here.

Data Availability and Validation

Obviously the lower the count of occurrences per link and large time interval the less accurate the resulting indication will be – and this is compounded by the length of the link which is also a contributing factor.

- Long links and a low count can cause the occurrences to have greater spatial intervals and thus may not be reflective of the actual road conditions. (Ref. Table 6.3.1)
- Short links and low counts decrease the odds of an occurrence actually matching to that link, and again may not be reflective of the actual road conditions. (Ref. Table 6.3.1)

As the number of trips operated increase so does the availability of data thus increases and negates this issue. The data used in this research consisted of one trip (Baker Street to Norwood). Application of this approach on a London wide scale would generate in excess of 140,000 trips daily on a normal Monday to Friday.

But there is obviously an issue at the start of the change of a temporal granularity and interval in our model where the amount of data occurrences will be low - this requires us to re-think how to model temporal validation.

We could validate on a rolling interval basis, say 30mins, yet although this may significantly lower the risk of having a low number of occurrences at any time, it still throws up another issue. A 30 minute capture window will yield more data during the day as opposed to the night, due to the nature of bus operations (fewer buses run during the night).

We need to be careful to ensure outliers do not skew the calculated statistics of our model and that the % for error is within an agreed tolerance. Outliers and exceptions can be classified into the areas of: passenger behaviour, bus stopping, urban canyoning and signal drop-out – all can potentially affect the calculations for links.

Urban canyoning, where the signal can reflect off of buildings, may give false location data which may match it to an incorrect spatial transport link.

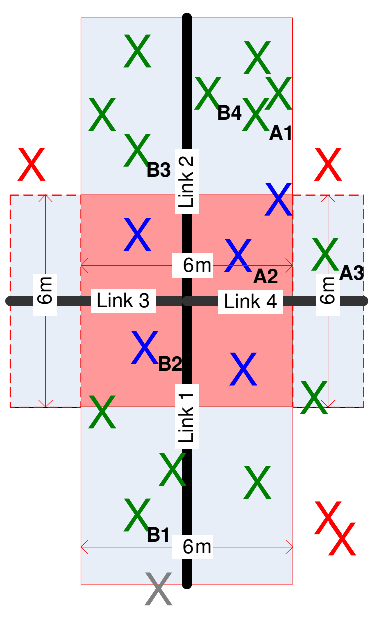
Signal drop out, where the line of sight to three or more satellites is obscured by tall buildings, will mean the spatial distribution of data for a given trip is increased, thus the saturation of matching per ITN link is less.

The prudent solution to solve these issues would be to build in an algorithm using the median. The median is primarily used for skewed distributions, which it represents more accurately than the arithmetic mean (average) to handle outliers and false information (to look for patterns and trends). Calculation of medians is a popular technique in summary statistics and summarizing statistical data, since it is simple to understand and easy to calculate, while also giving a measure that is more robust in the presence of outlier values than is the mean [41].

Given that passenger loadings and demand is highly dynamic across all buses, the distributed and unpredictable nature negates any skew in the data (all buses are subject to the same approximate delays) and given that this is a bus decision tool may even add another element into the statistics, e.g. if a concert has just finished, passenger loadings is of high demand and this is reflected in the reduced speed times along a particular stretch of road.

Direction can be very difficult to represent as the majority of road network links represent bi-directional. The matching of GPS location points from journeys in alternate directions, using the spatial operators 'nearest neighbour' or 'within distance', tends to consolidate to the same transit link. This can be an issue especially when delays and congestion are only in one direction, one way to determine the heading is to acknowledge the start node and end node along with using a linear referencing system (LRS). A second method would be to use the vehicle heading sent back from the bus.

But, is direction important in our application? Delays in one direction on a single carriageway may affect the speed of travel in the opposite direction – the carriageways in either direction are not closed entities. Thus the information presented could be quite useful when assessing congestion along a single carriageway stretch of road. And anyway, where dual carriageways do exist, the carriageways are modelled with separate links; the directional carriageways are isolated and therefore less likely to be affected by each other. Although the term 'rubber-necking' would tend to counter claim that statement, the traffic condition has affectively changed and this would be reflected in the average speed of the link.



As has already been discussed, false readings can cause matching to an incorrect link. Yet the GPS reading does not necessarily need to be false to be incorrectly matched.

If we take the example on the left – regardless whether or not direction is important – the accuracy of spatial matching to the correct transit link is. The location fixes highlighted in blue could belong to either of the enclosing links in each of the 3m^2 quads.

This is where Map Matching (MM) algorithms have been applied in transport applications to aid in correctly identifying a vehicle location against a road network. The prognostic approach of probability would be to use the directional heading of the vehicle at each location fix, yet this attribute may not always be a feature of all location systems. Therefore the pragmatic approach is to undertake a two-stage validation process, this being (1) the initial matching process (IMP), and (2) the subsequent matching process (SMP). Using our example, the process works as follows:

Initial Matching Process		**Subsequent Matching Process**	
A1	Link 2	–	–
A2	Link 2 or Link 4	A3	Link 4
A3	Link 4	–	–
B1	Link 1	–	–
B2	Link 1 or Link 3	B3	Link 1
B3	Link 2	–	–
B4	Link 2	–	–

The linear referencing system, which can help resolve the issue of direction, works in a similar way to the subsequent matching process, but instead uses a process of preceding matching. In our example, the location fix at B4 could be mistaken for going in a southerly direction, however the preceding initial matching of fix B3 was south of B4, so the direction being travelled must therefore be north.

Research by authors (Quddus et al., 2003; Ochieng et al., 2003) has developed improved Map Matching algorithms. And further research by authors (Noland et al., 2004), submitted to the 84th Annual Meeting of the Transportation Research Board, goes a stage further in assessing the performance of map matching algorithms by describing a generic validation strategy and the results for the previously developed map matching algorithms [43].

Our application dictates that it should reflect as closely as possible to the real world operating time periods in order to gain maximum benefits from the location data used to model congestion as accurately as is feasible. The Linear model of time supported by most relational temporal models would not be sufficient for a fully working version of our application if we desired absolute accuracy due to the branching nature of the temporal primitives ([non]-working days) and the multiple granularities (days / hours) involved for the time span of data validity. If we were to ignore the different day types as an additional time granularity then there would not be an issue. But it is perhaps sensible that we do consider the Object Orientated (OO) concept if we really want to get the best results out of the data. Research has shown that Object Orientated techniques are ideally suited for the infrastructure design of diverse notions of time under a single framework which capture the temporal needs of real-world applications [11]. The OO approach is not constrained in the same way as the relational model, where the data is normalised into different relations that are only susceptible to one temporal validation constraint. The Object Orientated approach for temporal modelling is similar to the spatial object in the Oracle relational database – temporal objects have their own attributes and behaviours imposed on them which can be programmed to interact as required – and more specifically adhere to multi temporal granularity constraints. Further research in establishing a comprehensive Object Orientated approach to the representation of multi-granular spatio-temporal data, as opposed to just OO concepts, was recently submitted at the Advance Information Systems Engineering 17th International Conference in 2005 by the authors (Bertion et al. 2005).

The spatial operator I used, 'within distance', within the Oracle platform was sufficient to prove the feasibility of the research project - but use of the 'nearest neighbour' operator for a full application would be more efficient, especially in situations of map matching probabilities. I was unable in the time to use the 'nearest neighbour' operator due to issues enabling the required spatial indexing on the Oracle spatial tables. The 'within distance' operator requires a set distance which applied to the entire road network would not capture all GPS occurrences due to the varied width of road carriageways, and there would be overlap in link segment buffers causing duplication in potentially one or more location fixes being matched to more than one transit link.

6.6. Limitations

There are, however, limitations that can not easily be solved, for instance not all roads in Greater London are served by buses, only 27.4% (3,730 km) of the 13,600 km of roads are served by bus routes, but these are mainly main roads [8]. This coverage may be sufficient, but if we take for example the Embankment (road) in London – it is a major trunk route through the capital, yet no buses run along the entire length of it. This is a major limitation if we wanted to use this form of application for monitoring all vehicle movements – but is not an issue in solving this problem domain.

Representation of the average / median speed over the network is a good indicator, although the inclusion of deviation from schedule could yield even greater beneficial results, especially for low frequency routes. Some form of linear interpolation would have to be generated for each trip based on distance and time in order for the deviation from the GPS location timestamp to be established at the matching transit link. This combination of using both average speed and schedule deviation together is likely to prove too difficult and must be deemed as a limitation of the current application data model.

The University Oracle Database version lacks support of the *spatial* object data type or the *timestamp* data type required for this research application. It was therefore reliant on me to install Oracle 10g (Express Edition) on my own home PC to build the database.

It was not possible to link the GIS directly into the DBMS due to the nature of the GIS software residing on my employer's network and the Oracle 10g DBMS being on my personal computer, the connection was not possible due to security limitations. This is not impossible, however, given the right setup.

6.7. Conclusions

I would like to have compared the processed GPS data against the processed loop detector along with a visual record of the traffic conditions for the day of the GPS pilot, 15th March, 2005. This would have been a useful validation exercise to prove the integrity of the processed GPS results. This has however not been possible due to data from loop detectors only being archived for up to 4 months. This is not to say that this cannot be done in future as the iBus project rolls out.

The use of Oracle 10g Locator module was easy to learn and use, the results yielded were to an acceptable accuracy for this prototype. The semantics of the data structure for spatial objects can easily be accessed and manipulated in the DBMS, as was the functionality to make spatial comparisons and indexing. The functionality to create triggers automates the database (Active Database) for use in a real-time environment, and even where it is difficult to implement temporal modelling the triggers can help to validate data.

The use of Oracle Map Viewer should be explored further as an alternative to GIS for rendering the spatio-temporal analysis. Map Viewer has the advantages of being fully integrated into the Oracle platform, and the ability to be web enabled.

There is still a disparate identity of GIS platforms. The vendors of GIS products should either announce their allegiance to one particular application (i.e. spatial planning, analysis, storage) or at least demonstrate progress in development of issues pertaining to extended functionality (i.e. temporal, different data structures, GML etc.).

A future project for TfL would be to combine the traffic flow data sources of Buses, Loops and Cameras to create an integrated and versatile Intelligent Transport System (ITS) for London. If the modelling of the spatial data from these various sources can perform to accurate temporal and spatial constraints then there are definite integration possibilities.

The use of this (research) application, in particular, for the scope of road conditions affecting *all* vehicles may not be entirely accurate – bus lanes and other bus priority measures would mean its use as a measure of congestion for normal vehicular traffic could be debateable in some places.

However, using GPS location data from buses for this type of application is demonstrable in this research in providing the desired outcome of the project. And as long as the issues and the possible solutions raised here are implemented then use of the application as a decision support tool for London Buses will be highly beneficial to its business and a strong basis for future development in this field.

6.8. Further Work

A first suggestion of further work would be to investigate how the Object Orientated approach can be used to the application's advantage in further enhancing the accuracy of the results by validating data against multi temporal granularities that mirror the operation of the bus network in real-time.

There is also great potential to take development one stage further than this research application. The next obvious evolution of this type of application is to integrate a "what if" capability – a projected simulation that responds to the different real-time behaviours of the bus network and predicts the likely scenarios of the road network in near future periods of time. There has already been research conducted in the area of Adaptive Prediction of Incident Delay [43], the research attempts to implement a sense of fuzzy logic in the unpredictability that surround delays caused by incidents, in order to predict the likely degree of future delays on the road network. This research was conducted with the intended use for advertising accurate delay messages on Variable Message Signs (VMS); however, there is a similar opportunity for this to be adapted to enhance the real-time research application discussed in this project.

7. References

[1] Reggiani, A. *Evolutionary Approaches to Transport and Spatial Systems*. In Hensher, D.A. (ed) Handbook of Transport Geography and Spatial Systems. Oxford: Elsvier, 2004, Chapter 14, pp.237-253.

[2] Golledge, R.G., Garling, T. *Cognitive Maps and Urban Travel*. In Hensher, D.A. (ed) Handbook of Transport Geography and Spatial Systems. Oxford: Elsvier, 2004, Chapter 28, pp.501-512.

[3] Stern, E., Sinuany- Stern, Z., Spharadi, Z., Holm, E. Congestion-related information and road network performance. *Journal of Transport Geography*, 1996, 4(3), pp.169-178.

[4] Janelle, D.G. Sustainable transportation and information technology: suggested research issues. *Journal of Transport Geography*, 1997, 5(1), pp.39-40.

[5] Wolf, J. *Defining GPS and its Capabilities*. In Hensher, D.A. (ed) Handbook of Transport Geography and Spatial Systems. Oxford: Elsvier, 2004, Chapter 23, pp.411-431.

[6] Transport for London. *iBus Initiative (Automatic Vehicle Location and Voice Radio Services)*. Available: http://www.tfl.gov.uk/buses/ini-ibus.asp

[7] Manolopoulos, Y., Papadopoulos, A.N., Vassilakopoulos, M.G. *Spatial Databases – Technologies, Techniques and Trends*. London: Hershey, 2005.

[8] Kothuri, R., Godfrind, A., Beinat, E. *Pro Oracle Spatial*. USA: Apress, 2004.

[9] OGC (Open GIS Consortium) White Paper. *Data Models and Interoperability*. 2003.

[10] Elmasri, R., Navathe, S.B. *Fundamentals of Database Systems (Fourth Edition)*. UK: Addison Wesley, 2004.

[11] Goralwalla, I.A., Ozsu, M.T., Szafron, D. *An Object-Orientated Framework for Temporal Data Models*. In Etzion, O., Jajodia, S., Sripada, S. (ed) Temporal Databases: Research and Practice. Germany: Springer, 1998, Part 1, pp.1-35.

[12] Slavin, H.L. *The Role of GIS in Land Use and Transport Planning*. In Hensher, D.A. (ed) Handbook of Transport Geography and Spatial Systems. Oxford: Elsvier, 2004, Chapter 19, pp.355-356.

[13]

[14] Lee, J-Y. Integrating Spatial and Temporal Relationship Operators into SQL3 for Historical Data Management. *ETRI Journal*, 2002, 24(3), pp.226-238.

[15] Spear, B.D. *Linking Spatial and Transportation Data*. In Hensher, D.A. (ed) Handbook of Transport Geography and Spatial Systems. Oxford: Elsvier, 2004, Chapter 18, pp.309-326.

[16] Eisenberg, A., Melton, J. SQL: 1999, formerly known as SQL3, 1999.

[17] Stopher, P.R. *Spatial Data Issues: A Historical Perspective*. In Hensher, D.A. (ed) Handbook of Transport Geography and Spatial Systems. Oxford: Elsvier, 2004, Chapter 17, pp.293-308.

[18] Fischer, M.M. *GIS and Network Analysis*. In Hensher, D.A. (ed) Handbook of Transport Geography and Spatial Systems. Oxford: Elsvier, 2004, Chapter 22, pp.391-408.

[19] Sutton, J.C., Visser, J. *The Role of GIS in Routing and Logistics*. In Hensher, D.A. (ed) Handbook of Transport Geography and Spatial Systems. Oxford: Elsvier, 2004, Chapter 20, pp.357-374.

[20] Greaves, S. *GIS and the collection of Travel Survey Data*. In Hensher, D.A. (ed) Handbook of Transport Geography and Spatial Systems. Oxford: Elsvier, 2004, Chapter 21, pp.375-390.

[21] Bettinin, C., Jajodia, S., Wang, S.X. *Time Granularities in Databases, Data Mining, and Temporal Reasoning*. Germany: Springer, 2005.

[22] OpenGIS Geography Markup Language (GML) Implementation Specification ISO 19136. Version 3.1.0. 2004. Editors: Simon Cox, Paul Daisey, Ron Lake, Clemens Portele, Arliss Whiteside.

[23] Seligman, L., Rosenthal, A. *XML's Impact on Databases and Data Sharing*. IEEE. June 2001.

[24] Open GeoSpatial Consortium, Inc. *OpenGIS FAQs - OGC and "Openness"*. 2006. Available: http://www.opengeospatial.org/

[25] Gould, Prof. M., Hecht Jr., L. OGC: A Framework for Geospatial and Statistical Information Integration. An Open GIS Consortium (OGC) White Paper. In *Proceedings Joint UNECE/Eurostat Work Session on Methodological Issues Involving the Integration of Statistics and Geography*. Tallinn, Estonia 25-28 Sept 2001.

[26] Lake, R. *Geography Markup Language Presentation*. Available: http://www.w3.org/Mobile/posdep/Presentations/GMLIntro/sld001.htm W3.org. 2000.

[27] OpenGIS Temporal Recommendations. Version 0.0.9. 2006. Editor: James Resler.

[28] Torrens, P.M. *Geosimulation, Automata, and Traffic Modelling*. In Hensher, D.A. (ed) Handbook of Transport Geography and Spatial Systems. Oxford: Elsvier, 2004, Chapter 31, pp.549-564.

[29] Qian, L.J. Oracle Database 10g: Developing Spatial Applications Using Oracle Spatial and MapViewer. An Oracle Technical White Paper. 2004. Available: http://www.oracle.com

[30] Sharma, Dr. J. Oracle Spatial. An Oracle Technical White Paper. 2002. Available: http://www.oracle.com

[31] Thrall, G.I. *[MapInfo] User-Friendly GIS*. In Geospatial Solutions. USA: Advanstar, 2005.

[32] Eastaway, R., Wyndham, J. *Why Do Buses Come In Threes? The hidden mathematics of everyday life*. Great Britain: Robson Books Ltd, 1998.

[33] Camilletti, P. *High Frequency Calculations Explained*. London Buses. 2002. Available: TfL Intranet.

[34] Avison, D., Fitzgerald, G. *Information Systems Development: Methodologies, Techniques and Tools, Third Edition*. London: McGraw-Hill Education, 2003.

[35] Ordnance Survey. *OS Mastermap® Integrated Transport Network™(ITN) Layer*. 2006. Available: http://www.ordnancesurvey.co.uk/oswebsite/products/osmastermap/layers/itn/index.html

[36] Oxford Economic Forecasting. *The Economic Effects of Transport Delays on the City of London*. London: Corporation of London, 2003.

[37] Khatri, V., Vessey, I., Ram, S., Ramesh, V. Cognitive Fit Between Conceptual Schemas and Internal Problem Representations: The Case of Geospatio–Temporal Conceptual Schema Comprehension. *IEEE Transactions on Professional Communication*, 2006, 49(2), pp.109-127.

[38] Transport for London. *Traffic Operations – Signals*. Available: http://www.tfl.gov.uk/streets/dtm/signals.shtml

[39] Transport for London. *Traffic Operations – COMET*. Available: http://www.tfl.gov.uk/streets/dtm/comet.shtml

[40] Transport for London. *TfL Junction Design Guide*. Available: TfL Intranet.

[41] Rousseeuw, P.J., Leroy, A.M. *Robust Regression and Outlier Detection*. Hoboken, New Jersey: John Wiley & Sons, Inc, 2003.

[42] Quddus, M.A., Noland, R.B., Ochieng, W.Y. Validation of Map Matching Algorithms using High Precision Positioning with GPS. In *Proceedings 84th Annual Meeting of the Transportation Research Board*. Sept 2004.

[43] Fu, L. A Fuzzy Queuing Model for Real-time, Adaptive Prediction of Incident Delay for ATMS/ATIS. *Transport Planning and Technology*, 2004, 27(1), pp.1-23.

8. Appendix

APPENDIX A

GPS Location Survey Data

DATE	TIME	X_COORD	Y_COORD	SPEED
15-Mar-05	15:31:47	527912	181968	28
15-Mar-05	15:36:04	527862	182163	11
15-Mar-05	15:36:05	527862	182163	17
15-Mar-05	15:36:06	527862	182163	23
15-Mar-05	15:36:07	527862	182163	23
15-Mar-05	15:36:08	527863	182160	28
15-Mar-05	15:36:09	527865	182155	28
15-Mar-05	15:36:10	527866	182149	28
15-Mar-05	15:36:11	527868	182144	23
15-Mar-05	15:36:12	527870	182137	23
15-Mar-05	15:36:13	527872	182130	23
15-Mar-05	15:36:14	527874	182124	23
15-Mar-05	15:36:15	527876	182117	23
15-Mar-05	15:36:16	527878	182112	18
15-Mar-05	15:36:17	527879	182106	17
15-Mar-05	15:36:18	527881	182100	11
15-Mar-05	15:36:19	527883	182095	11
15-Mar-05	15:36:20	527884	182091	12
15-Mar-05	15:36:21	527885	182088	6
15-Mar-05	15:36:22	527886	182085	0
15-Mar-05	15:36:23	527887	182082	0
15-Mar-05	15:36:24	527887	182080	6
15-Mar-05	15:36:25	527887	182079	0
15-Mar-05	15:36:26	527888	182079	0
15-Mar-05	15:36:27	527888	182079	0
15-Mar-05	15:36:28	527888	182079	0
15-Mar-05	15:36:29	527888	182079	0
15-Mar-05	15:36:30	527888	182079	0
15-Mar-05	15:36:31	527888	182079	0
15-Mar-05	15:36:32	527888	182079	0
15-Mar-05	15:36:33	527888	182079	0
15-Mar-05	15:36:34	527888	182079	0
15-Mar-05	15:36:35	527888	182079	0
15-Mar-05	15:36:36	527888	182079	0
15-Mar-05	15:36:37	527888	182079	0
15-Mar-05	15:36:38	527888	182079	0

15-Mar-05	15:36:39	527888	182079	0
15-Mar-05	15:36:40	527888	182079	0
15-Mar-05	15:36:41	527888	182079	0
15-Mar-05	15:36:42	527888	182079	0
15-Mar-05	15:36:43	527888	182079	0
15-Mar-05	15:36:44	527888	182079	0
15-Mar-05	15:36:45	527888	182079	0
15-Mar-05	15:36:46	527888	182079	0
15-Mar-05	15:36:47	527888	182079	0
15-Mar-05	15:36:48	527888	182079	0
15-Mar-05	15:36:49	527888	182079	0
15-Mar-05	15:36:50	527888	182079	6
15-Mar-05	15:36:51	527888	182079	6
15-Mar-05	15:36:52	527888	182078	11
15-Mar-05	15:36:53	527889	182075	17
15-Mar-05	15:36:54	527890	182072	17
15-Mar-05	15:36:55	527891	182067	17
15-Mar-05	15:36:56	527893	182062	12
15-Mar-05	15:36:57	527894	182057	11
15-Mar-05	15:36:58	527896	182053	17
15-Mar-05	15:36:59	527897	182048	11
15-Mar-05	15:37:00	527899	182044	12
15-Mar-05	15:37:01	527900	182040	11
15-Mar-05	15:37:02	527901	182036	11
15-Mar-05	15:37:03	527902	182033	11
15-Mar-05	15:37:04	527903	182030	12
15-Mar-05	15:37:05	527904	182028	6
15-Mar-05	15:37:06	527904	182026	6
15-Mar-05	15:37:07	527905	182024	0
15-Mar-05	15:37:08	527906	182022	0
15-Mar-05	15:37:09	527906	182022	0
15-Mar-05	15:37:10	527906	182022	0
15-Mar-05	15:37:11	527906	182022	0
15-Mar-05	15:37:12	527906	182022	0
15-Mar-05	15:37:13	527906	182022	0
15-Mar-05	15:37:14	527906	182022	0
15-Mar-05	15:37:15	527906	182022	0
15-Mar-05	15:37:16	527906	182022	0

15-Mar-05	15:37:17	527906	182022	0
15-Mar-05	15:37:18	527906	182022	0
15-Mar-05	15:37:19	527906	182022	0
15-Mar-05	15:37:20	527906	182022	0
15-Mar-05	15:37:21	527906	182022	0
15-Mar-05	15:37:22	527906	182022	0
15-Mar-05	15:37:23	527906	182022	0
15-Mar-05	15:37:24	527906	182022	0
15-Mar-05	15:37:25	527906	182022	0
15-Mar-05	15:37:26	527906	182022	0
15-Mar-05	15:37:27	527906	182022	0
15-Mar-05	15:37:28	527906	182022	0
15-Mar-05	15:37:29	527906	182022	0
15-Mar-05	15:37:30	527906	182022	0
15-Mar-05	15:37:31	527906	182022	0
15-Mar-05	15:37:32	527906	182022	0
15-Mar-05	15:37:33	527906	182022	6
15-Mar-05	15:37:34	527906	182022	11
15-Mar-05	15:37:35	527906	182021	17
15-Mar-05	15:37:36	527907	182019	17
15-Mar-05	15:37:37	527908	182014	23
15-Mar-05	15:37:38	527910	182009	23
15-Mar-05	15:37:39	527912	182003	23
15-Mar-05	15:37:40	527914	181997	23
15-Mar-05	15:37:40	527918	181984	23
15-Mar-05	15:37:41	527920	181978	23
15-Mar-05	15:37:43	527918	181984	23
15-Mar-05	15:37:44	527922	181971	23
15-Mar-05	15:37:45	527924	181964	23
15-Mar-05	15:37:46	527927	181958	23
15-Mar-05	15:37:47	527929	181951	23
15-Mar-05	15:37:48	527930	181945	23
15-Mar-05	15:37:49	527932	181938	23
15-Mar-05	15:37:49	527936	181926	23
15-Mar-05	15:37:50	527939	181919	23
15-Mar-05	15:37:52	527936	181926	17
15-Mar-05	15:37:53	527940	181913	18
15-Mar-05	15:37:54	527942	181907	17

15-Mar-05	15:37:55	527944	181902	17
15-Mar-05	15:37:56	527945	181898	17
15-Mar-05	15:37:57	527946	181894	17
15-Mar-05	15:37:58	527947	181890	12
15-Mar-05	15:37:59	527949	181886	11
15-Mar-05	15:38:00	527950	181883	6
15-Mar-05	15:38:01	527951	181880	0
15-Mar-05	15:38:02	527951	181879	0
15-Mar-05	15:38:03	527951	181878	0
15-Mar-05	15:38:04	527951	181878	0
15-Mar-05	15:38:05	527951	181878	0
15-Mar-05	15:38:06	527951	181878	0
15-Mar-05	15:38:07	527951	181878	0
15-Mar-05	15:38:08	527951	181878	0
15-Mar-05	15:38:09	527951	181878	0
15-Mar-05	15:38:10	527951	181878	0
15-Mar-05	15:38:11	527951	181878	0
15-Mar-05	15:38:12	527951	181878	0
15-Mar-05	15:38:13	527951	181878	0
15-Mar-05	15:38:14	527951	181878	0
15-Mar-05	15:38:15	527951	181878	0
15-Mar-05	15:38:16	527951	181878	0
15-Mar-05	15:38:17	527951	181878	0
15-Mar-05	15:38:18	527951	181878	0
15-Mar-05	15:38:19	527951	181878	0
15-Mar-05	15:38:20	527951	181878	0
15-Mar-05	15:38:21	527951	181878	0
15-Mar-05	15:38:22	527951	181878	0
15-Mar-05	15:38:23	527951	181878	0
15-Mar-05	15:38:24	527951	181878	0
15-Mar-05	15:38:25	527951	181878	0
15-Mar-05	15:38:26	527951	181878	0
15-Mar-05	15:38:27	527951	181878	0
15-Mar-05	15:38:28	527951	181878	0
15-Mar-05	15:38:29	527951	181878	0
15-Mar-05	15:38:30	527951	181878	0
15-Mar-05	15:38:31	527951	181878	0
15-Mar-05	15:38:32	527951	181878	0

15-Mar-05	15:38:33	527951	181878	0
15-Mar-05	15:38:34	527951	181878	0
15-Mar-05	15:38:35	527951	181878	0
15-Mar-05	15:38:36	527951	181878	0
15-Mar-05	15:38:37	527951	181878	0
15-Mar-05	15:38:38	527951	181878	0
15-Mar-05	15:38:39	527951	181878	0
15-Mar-05	15:38:40	527951	181878	0
15-Mar-05	15:38:41	527951	181878	0
15-Mar-05	15:38:42	527951	181878	0
15-Mar-05	15:38:43	527951	181878	0
15-Mar-05	15:38:44	527951	181878	0
15-Mar-05	15:38:45	527951	181878	0
15-Mar-05	15:38:46	527951	181878	0
15-Mar-05	15:38:47	527951	181878	6
15-Mar-05	15:38:48	527951	181878	6
15-Mar-05	15:38:49	527952	181877	6
15-Mar-05	15:38:50	527952	181876	6
15-Mar-05	15:38:51	527952	181875	12
15-Mar-05	15:38:52	527953	181873	11
15-Mar-05	15:38:53	527953	181871	6
15-Mar-05	15:38:54	527954	181869	11
15-Mar-05	15:38:55	527955	181867	18
15-Mar-05	15:38:56	527956	181864	17
15-Mar-05	15:38:57	527957	181860	17
15-Mar-05	15:38:58	527959	181855	17
15-Mar-05	15:38:59	527960	181850	17
15-Mar-05	15:39:00	527962	181845	23
15-Mar-05	15:39:01	527963	181840	23
15-Mar-05	15:39:02	527965	181835	23
15-Mar-05	15:39:03	527967	181829	17
15-Mar-05	15:39:04	527968	181824	17
15-Mar-05	15:39:05	527970	181819	17
15-Mar-05	15:39:06	527972	181813	17
15-Mar-05	15:39:07	527973	181809	17
15-Mar-05	15:39:08	527975	181804	17
15-Mar-05	15:39:09	527976	181800	17
15-Mar-05	15:39:09	527979	181791	17

15-Mar-05	15:39:10	527980	181787	17
15-Mar-05	15:39:12	527979	181791	17
15-Mar-05	15:39:13	527981	181783	23
15-Mar-05	15:39:14	527983	181779	23
15-Mar-05	15:39:15	527984	181774	23
15-Mar-05	15:39:16	527985	181770	23
15-Mar-05	15:39:17	527987	181764	23
15-Mar-05	15:39:18	527989	181758	28
15-Mar-05	15:39:19	527991	181752	28
15-Mar-05	15:39:20	527993	181745	29
15-Mar-05	15:39:21	527996	181737	34
15-Mar-05	15:39:22	527998	181729	29
15-Mar-05	15:39:23	528000	181723	34
15-Mar-05	15:39:24	528003	181715	34
15-Mar-05	15:39:25	528005	181707	34
15-Mar-05	15:39:26	528008	181699	35
15-Mar-05	15:39:27	528011	181690	28
15-Mar-05	15:39:28	528014	181681	34
15-Mar-05	15:39:29	528016	181672	34
15-Mar-05	15:39:30	528019	181664	29
15-Mar-05	15:39:31	528022	181655	29
15-Mar-05	15:39:32	528025	181646	34
15-Mar-05	15:39:33	528027	181637	34
15-Mar-05	15:39:34	528030	181628	34
15-Mar-05	15:39:35	528033	181619	35
15-Mar-05	15:39:36	528035	181611	34
15-Mar-05	15:39:37	528038	181602	29
15-Mar-05	15:39:38	528040	181594	34
15-Mar-05	15:39:39	528043	181585	29
15-Mar-05	15:39:40	528045	181577	29
15-Mar-05	15:39:41	528048	181569	34
15-Mar-05	15:39:42	528050	181561	28
15-Mar-05	15:39:43	528053	181552	29
15-Mar-05	15:39:44	528055	181544	35
15-Mar-05	15:39:45	528058	181535	34
15-Mar-05	15:39:46	528061	181527	34
15-Mar-05	15:39:47	528063	181518	34
15-Mar-05	15:39:48	528066	181509	29

15-Mar-05	15:39:49	528069	181501	34
15-Mar-05	15:39:50	528071	181492	28
15-Mar-05	15:39:51	528074	181484	28
15-Mar-05	15:39:52	528077	181475	23
15-Mar-05	15:39:53	528079	181467	12
15-Mar-05	15:39:54	528081	181461	6
15-Mar-05	15:39:55	528083	181456	0
15-Mar-05	15:39:56	528084	181454	0
15-Mar-05	15:39:57	528084	181453	0
15-Mar-05	15:39:58	528084	181453	0
15-Mar-05	15:39:59	528084	181453	0
15-Mar-05	15:40:00	528084	181453	0
15-Mar-05	15:40:01	528084	181453	0
15-Mar-05	15:40:02	528084	181453	0
15-Mar-05	15:40:03	528084	181453	0
15-Mar-05	15:40:04	528084	181453	0
15-Mar-05	15:40:05	528084	181453	0
15-Mar-05	15:40:06	528084	181453	0
15-Mar-05	15:40:07	528084	181453	0
15-Mar-05	15:40:08	528084	181453	0
15-Mar-05	15:40:09	528084	181453	0
15-Mar-05	15:40:10	528084	181453	0
15-Mar-05	15:40:11	528084	181453	0
15-Mar-05	15:40:12	528084	181453	0
15-Mar-05	15:40:13	528084	181453	0
15-Mar-05	15:40:14	528084	181453	0
15-Mar-05	15:40:15	528084	181453	0
15-Mar-05	15:40:16	528084	181453	0
15-Mar-05	15:40:17	528084	181453	0
15-Mar-05	15:40:18	528084	181453	0
15-Mar-05	15:40:19	528084	181453	0
15-Mar-05	15:40:20	528084	181453	0
15-Mar-05	15:40:21	528084	181453	0
15-Mar-05	15:40:22	528084	181453	0
15-Mar-05	15:40:23	528084	181453	0
15-Mar-05	15:40:24	528084	181453	0
15-Mar-05	15:40:25	528084	181453	0
15-Mar-05	15:40:26	528084	181453	0

15-Mar-05	15:40:27	528084	181453	0
15-Mar-05	15:40:28	528084	181453	0
15-Mar-05	15:40:29	528084	181453	0
15-Mar-05	15:40:30	528084	181453	0
15-Mar-05	15:40:31	528084	181453	0
15-Mar-05	15:40:32	528084	181453	0
15-Mar-05	15:40:33	528084	181453	0
15-Mar-05	15:40:34	528084	181453	0
15-Mar-05	15:40:35	528084	181453	0
15-Mar-05	15:40:36	528084	181453	0
15-Mar-05	15:40:37	528084	181453	0
15-Mar-05	15:40:38	528084	181453	0
15-Mar-05	15:40:39	528084	181453	0
15-Mar-05	15:40:40	528084	181453	0
15-Mar-05	15:40:41	528084	181453	0
15-Mar-05	15:40:42	528084	181453	0
15-Mar-05	15:40:43	528084	181453	0
15-Mar-05	15:40:44	528084	181453	0
15-Mar-05	15:40:45	528084	181453	0
15-Mar-05	15:40:46	528084	181453	0
15-Mar-05	15:40:47	528084	181453	0
15-Mar-05	15:40:48	528084	181453	6
15-Mar-05	15:40:49	528084	181453	17
15-Mar-05	15:40:50	528084	181452	17
15-Mar-05	15:40:51	528085	181449	23
15-Mar-05	15:40:52	528087	181444	28
15-Mar-05	15:40:53	528088	181438	28
15-Mar-05	15:40:54	528091	181430	29
15-Mar-05	15:40:55	528093	181422	29
15-Mar-05	15:40:56	528096	181413	28
15-Mar-05	15:40:57	528098	181405	23
15-Mar-05	15:40:58	528100	181397	28
15-Mar-05	15:40:59	528102	181390	23
15-Mar-05	15:41:00	528104	181383	23
15-Mar-05	15:41:01	528106	181376	23
15-Mar-05	15:41:02	528108	181370	23
15-Mar-05	15:41:03	528110	181364	23
15-Mar-05	15:41:04	528111	181358	12

15-Mar-05	15:41:05	528113	181351	11
15-Mar-05	15:41:06	528114	181347	6
15-Mar-05	15:41:07	528115	181343	6
15-Mar-05	15:41:08	528116	181341	0
15-Mar-05	15:41:09	528116	181339	0
15-Mar-05	15:41:10	528117	181338	0
15-Mar-05	15:41:11	528117	181337	0
15-Mar-05	15:41:12	528117	181337	0
15-Mar-05	15:41:13	528117	181337	0
15-Mar-05	15:41:14	528117	181336	0
15-Mar-05	15:41:15	528117	181336	6
15-Mar-05	15:41:16	528117	181336	17
15-Mar-05	15:41:17	528118	181335	17
15-Mar-05	15:41:18	528119	181331	23
15-Mar-05	15:41:19	528120	181326	29
15-Mar-05	15:41:20	528122	181322	28
15-Mar-05	15:41:21	528124	181315	35
15-Mar-05	15:41:22	528126	181308	35
15-Mar-05	15:41:23	528128	181300	34
15-Mar-05	15:41:24	528131	181291	40
15-Mar-05	15:41:25	528133	181282	40
15-Mar-05	15:41:26	528136	181272	41
15-Mar-05	15:41:26	528144	181251	41
15-Mar-05	15:41:27	528149	181240	40
15-Mar-05	15:41:29	528144	181251	45
15-Mar-05	15:41:30	528154	181229	47
15-Mar-05	15:41:31	528158	181217	40
15-Mar-05	15:41:32	528161	181206	40
15-Mar-05	15:41:33	528164	181195	40
15-Mar-05	15:41:34	528167	181183	41
15-Mar-05	15:41:35	528169	181171	41
15-Mar-05	15:41:36	528172	181162	40
15-Mar-05	15:41:37	528175	181152	34
15-Mar-05	15:41:38	528178	181142	34
15-Mar-05	15:41:39	528181	181132	35
15-Mar-05	15:41:40	528184	181123	34
15-Mar-05	15:41:41	528187	181114	34
15-Mar-05	15:41:42	528190	181105	28

15-Mar-05	15:41:43	528193	181096	23
15-Mar-05	15:41:44	528195	181088	23

APPENDIX B

DML and DDL SQL Statements

create_tables.sql

```
CREATE TABLE link
(
	link_toid		NUMBER CONSTRAINT pk_link_toid PRIMARY KEY,
	link_road_name		VARCHAR2(80),
	link_geom		SDO_GEOMETRY
);

CREATE TABLE location
(
	loc_id			NUMBER CONSTRAINT pk_id PRIMARY KEY,
	loc_poll_time		TIMESTAMP(0),
	loc_geom		SDO_GEOMETRY,
	loc_speed		NUMBER(2)
);

CREATE TABLE link_loc
(
	link_loc_toid		NUMBER CONSTRAINT fk_link_loc_toid REFERENCES link(link_toid),
	link_loc_id		NUMBER CONSTRAINT fk_link_loc_id REFERENCES location(loc_id)
);

CREATE TABLE link_speed
(
	link_speed_toid		NUMBER CONSTRAINT link_speed_toid REFERENCES link(link_toid),
	link_speed_avg_speed	FLOAT,
	link_speed_geom		SDO_GEOMETRY
);
```

insert_metadata_link.sql

```
INSERT INTO user_sdo_geom_metadata(table_name, column_name, srid, diminfo)VALUES
('link','link_geom',81989,
SDO_DIM_ARRAY(
SDO_DIM_ELEMENT('LONGITUDE',500,600,1),
SDO_DIM_ELEMENT('LATITUDE',150,205,1)));
```

insert_metadata_loc.sql

```
INSERT INTO user_sdo_geom_metadata(table_name, column_name, srid, diminfo)VALUES
('location','loc_geom',81989,
SDO_DIM_ARRAY(
SDO_DIM_ELEMENT('LONGITUDE',500,600,1),
SDO_DIM_ELEMENT('LATITUDE',150,205,1)));
```

create_index_link.sql

```
CREATE INDEX link_idx ON link(link_geom) INDEXTYPE IS MDSYS.SPATIAL_INDEX;
```

create_index_loc.sql

```
CREATE INDEX loc_idx ON location(loc_geom) INDEXTYPE IS MDSYS.SPATIAL_INDEX;
```

create_package.sql

```
CREATE OR REPLACE PACKAGE state_pkg
AS
	type ridArray is table of rowid index by binary_integer;

	newRows	ridArray;
	empty	ridArray;
END;
/

CREATE OR REPLACE TRIGGER location_bi
BEFORE INSERT OR UPDATE ON location
BEGIN
	state_pkg.newRows := state_pkg.empty;
END;
/

CREATE OR REPLACE TRIGGER location_aifer
AFTER INSERT OR UPDATE OF loc_id ON location FOR EACH ROW
BEGIN
	state_pkg.newRows( state_pkg.newRows.count+1 ) := :new.rowid;
END;
/

CREATE OR REPLACE TRIGGER location_ai
	AFTER INSERT OR UPDATE OF loc_id ON location
DECLARE
	BEGIN
	FOR i in 1 .. state_pkg.newRows.count loop
	INSERT INTO link_loc (link_loc_toid,link_loc_id)
	SELECT  K.link_toid, L.loc_id
		FROM location L, link K
		WHERE	SDO_WITHIN_DISTANCE(K.link_geom,	L.loc_geom,	'DISTANCE=6
UNIT=METER')='TRUE'
		AND L.rowid = state_pkg.newRows(i);
	END loop;
END;
/
```

create_package_2.sql

```
CREATE OR REPLACE PACKAGE state2_pkg
AS
        type ridArray is table of rowid index by binary_integer;

        newRows ridArray;
        empty   ridArray;
END;
/

CREATE OR REPLACE TRIGGER link_loc_bi
BEFORE INSERT OR UPDATE ON link_loc
BEGIN
        state2_pkg.newRows := state2_pkg.empty;
END;
/

CREATE OR REPLACE TRIGGER link_loc_aifer
AFTER INSERT OR UPDATE OF link_loc_id ON link_loc FOR EACH ROW
BEGIN
        state2_pkg.newRows( state2_pkg.newRows.count+1 ) := :new.rowid;
END;
/

CREATE OR REPLACE TRIGGER link_loc_ai
        AFTER INSERT OR UPDATE OF link_loc_id ON link_loc
DECLARE
        BEGIN
        FOR i in 1 .. state2_pkg.newRows.count loop
        INSERT INTO link_speed(link_speed_toid,link_speed_avg_speed)
        SELECT k.link_toid, avg(l.loc_speed)
        FROM link_loc c, location l, link k
        WHERE c.link_loc_id = l.loc_id
        AND c.link_loc_toid = k.link_toid
        AND c.rowid = state2_pkg.newRows(i)
        GROUP BY k.link_toid;

        END loop;
END;
/
```

insert_values_link.sql

```
INSERT INTO LINK VALUES (430217380,'BAKER STREET',SDO_GEOMETRY(2002,81989,NULL,SDO_ELEM_INFO_ARRAY(1,2,1),SDO_ORDINATE_ARRAY(527861,182162,527890,182070)));

INSERT INTO LINK VALUES (430217379,'BAKER STREET',SDO_GEOMETRY(2002,81989,NULL,SDO_ELEM_INFO_ARRAY(1,2,1),SDO_ORDINATE_ARRAY(527890,182070,527922,181971)));

INSERT INTO LINK VALUES (430407949,'BAKER STREET',SDO_GEOMETRY(2002,81989,NULL,SDO_ELEM_INFO_ARRAY(1,2,1),SDO_ORDINATE_ARRAY(527922,181971,527926,181959)));

INSERT INTO LINK VALUES (430217386,'BAKER STREET',SDO_GEOMETRY(2002,81989,NULL,SDO_ELEM_INFO_ARRAY(1,2,1),SDO_ORDINATE_ARRAY(527926,181959,527945,181899)));

INSERT INTO LINK VALUES (430473662,'BAKER STREET',SDO_GEOMETRY(2002,81989,NULL,SDO_ELEM_INFO_ARRAY(1,2,1),SDO_ORDINATE_ARRAY(527945,181899,527946,181895)));

INSERT INTO LINK VALUES (430324906,'BAKER STREET',SDO_GEOMETRY(2002,81989,NULL,SDO_ELEM_INFO_ARRAY(1,2,1),SDO_ORDINATE_ARRAY(527946,181895,527966,181831)));

INSERT INTO LINK VALUES (430217372,'BAKER STREET',SDO_GEOMETRY(2002,81989,NULL,SDO_ELEM_INFO_ARRAY(1,2,1),SDO_ORDINATE_ARRAY(527966,181831,527987,181763)));

INSERT INTO LINK VALUES (430217373,'BAKER STREET',SDO_GEOMETRY(2002,81989,NULL,SDO_ELEM_INFO_ARRAY(1,2,1),SDO_ORDINATE_ARRAY(527987,181763,528016,181669)));

INSERT INTO LINK VALUES (430135703,'BAKER STREET',SDO_GEOMETRY(2002,81989,NULL,SDO_ELEM_INFO_ARRAY(1,2,1),SDO_ORDINATE_ARRAY(528016,181669,528061,181524)));

INSERT INTO LINK VALUES (430217156,'BAKER STREET',SDO_GEOMETRY(2002,81989,NULL,SDO_ELEM_INFO_ARRAY(1,2,1),SDO_ORDINATE_ARRAY(528061,181524,528084,181449)));

INSERT INTO LINK VALUES (430324655,'BAKER STREET',SDO_GEOMETRY(2002,81989,NULL,SDO_ELEM_INFO_ARRAY(1,2,1),SDO_ORDINATE_ARRAY(528084,181449,528100,181398)));

INSERT INTO LINK VALUES (430217163,'BAKER STREET',SDO_GEOMETRY(2002,81989,NULL,SDO_ELEM_INFO_ARRAY(1,2,1),SDO_ORDINATE_ARRAY(528100,181398,528118,181329)));
```

```
INSERT INTO LINK VALUES (430418631,'PORTMAN SQUARE',SDO_GEOMETRY(2002,81989,NULL,SDO_ELEM_INFO_ARRAY(1,2,1),SDO_ORDINATE_ARRAY(528118,181329,528125,181311)));

INSERT INTO LINK VALUES (430343528,'PORTMAN SQUARE',SDO_GEOMETRY(2002,81989,NULL,SDO_ELEM_INFO_ARRAY(1,2,1),SDO_ORDINATE_ARRAY(528125,181311,528135,181276)));

INSERT INTO LINK VALUES (430418630,'PORTMAN SQUARE',SDO_GEOMETRY(2002,81989,NULL,SDO_ELEM_INFO_ARRAY(1,2,1),SDO_ORDINATE_ARRAY(528135,181276,528139,181260)));

INSERT INTO LINK VALUES (430343534,'PORTMAN SQUARE',SDO_GEOMETRY(2002,81989,NULL,SDO_ELEM_INFO_ARRAY(1,2,1),SDO_ORDINATE_ARRAY(528139,181260,528156,181224)));

INSERT INTO LINK VALUES (430343529,'ORCHARD STREET',SDO_GEOMETRY(2002,81989,NULL,SDO_ELEM_INFO_ARRAY(1,2,1),SDO_ORDINATE_ARRAY(528156,181224,528167,181184)));

INSERT INTO LINK VALUES (430459115,'ORCHARD STREET',SDO_GEOMETRY(2002,81989,NULL,SDO_ELEM_INFO_ARRAY(1,2,1),SDO_ORDINATE_ARRAY(528167,181184,528170,181170)));

INSERT INTO LINK VALUES (430147230,'ORCHARD STREET',SDO_GEOMETRY(2002,81989,NULL,SDO_ELEM_INFO_ARRAY(1,2,1),SDO_ORDINATE_ARRAY(528170,181170,528205,181058)));
```

insert_values_loc.sql (approx. 340 records)

insert_values_loc_2.sql (approx. 340 records)

Survey data from Appendix A in the following format:

```
INSERT INTO LOCATION VALUES ( 1 , to_timestamp(' 15-Mar-05 15:31:47
    ','DD-MM-YYYY HH24:MI:SS'), SDO_GEOMETRY(2001,81989,SDO_POINT_TYPE(
    527912 , 181968 , NULL),NULL,NULL), 28 );
```

APPENDIX C

SQL*Plus Application Screenshots

Run SQL Command Line

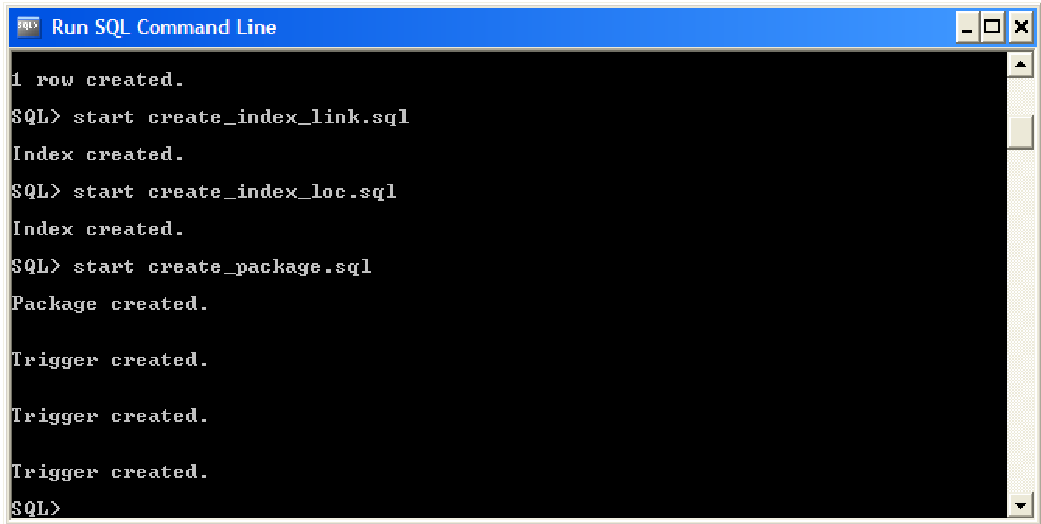
```
SQL*Plus: Release 10.2.0.1.0 - Production on Sun Oct 1 23:12:49 2006

Copyright (c) 1982, 2005, Oracle.  All rights reserved.

SQL> connect
Enter user-name: hr
Enter password:
Connected.
SQL> _
```

Run SQL Command Line

```
SQL*Plus: Release 10.2.0.1.0 - Production on Sun Oct 1 23:18:22 2006

Copyright (c) 1982, 2005, Oracle.  All rights reserved.

SQL> connect
Enter user-name: hr
Enter password:
Connected.
SQL> start create_tables.sql

Table created.


Table created.


Table created.


Table created.

SQL>
```

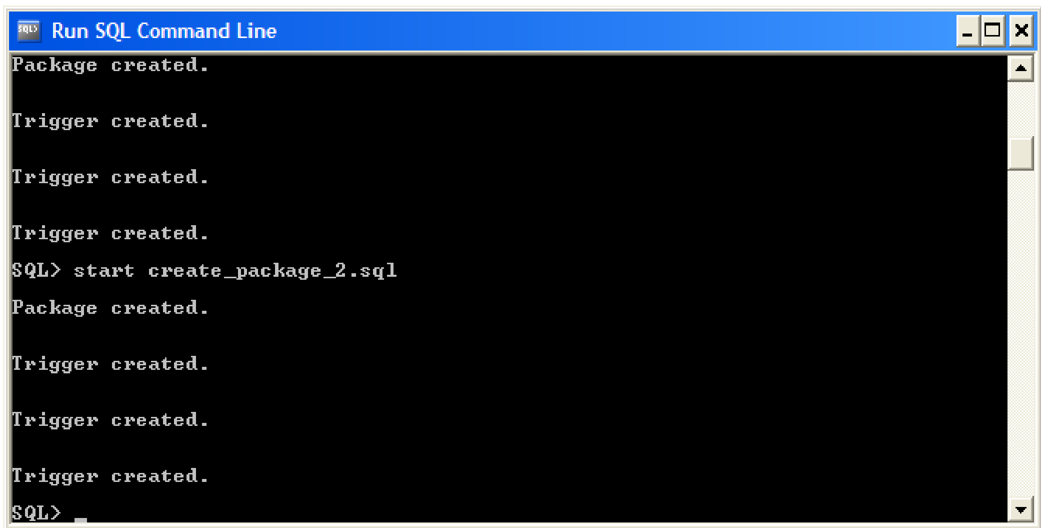
Run SQL Command Line

```
Table created.


Table created.


Table created.

SQL> start insert_metadata_link.sql

1 row created.

SQL> start insert_metadata_loc.sql

1 row created.

SQL> start create_index_link.sql

Index created.

SQL> start create_index_loc.sql

Index created.

SQL>
```



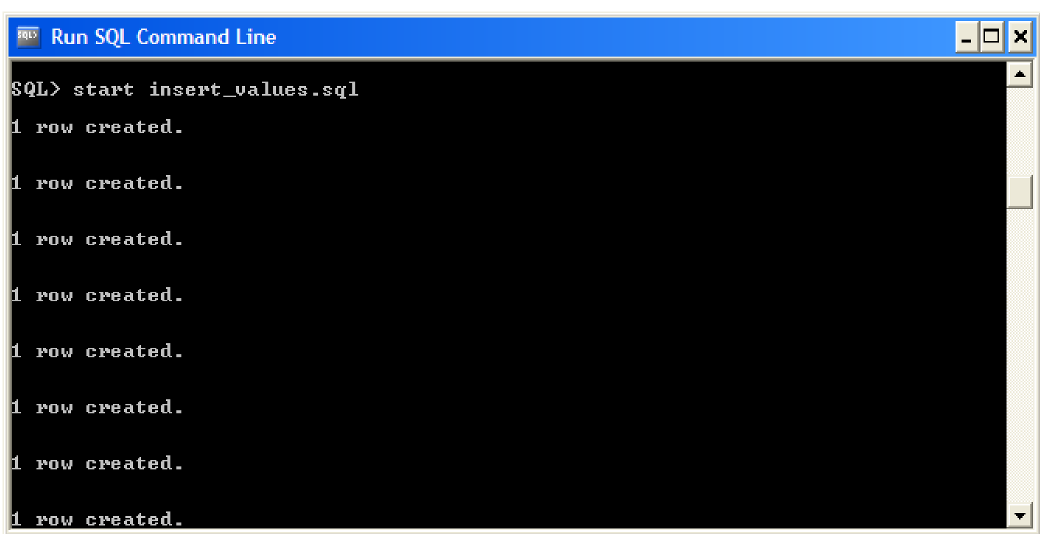
Run SQL Command Line

```
1 row created.
SQL> start create_index_link.sql
Index created.
SQL> start create_index_loc.sql
Index created.
SQL> start create_package.sql
Package created.

Trigger created.

Trigger created.

Trigger created.
SQL>
```



Run SQL Command Line

```
Package created.

Trigger created.

Trigger created.

Trigger created.
SQL> start create_package_2.sql
Package created.

Trigger created.

Trigger created.

Trigger created.
SQL> _
```



Run SQL Command Line

```
SQL> start insert_values.sql
1 row created.

1 row created.

1 row created.

1 row created.

1 row created.

1 row created.

1 row created.

1 row created.
```

Run SQL Command Line

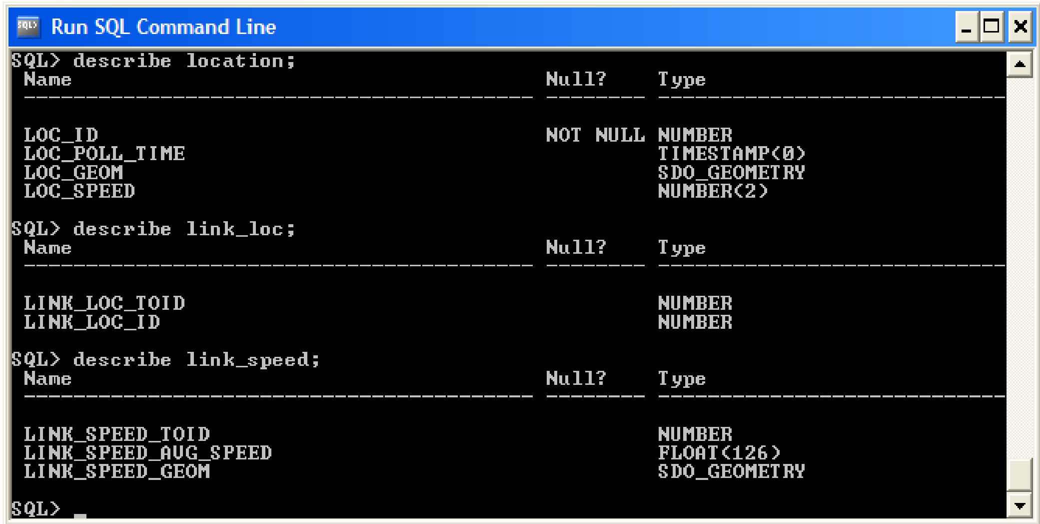
```
SQL> select count(*) from link;

  COUNT(*)
----------
        19

SQL> select count(*) from location;

  COUNT(*)
----------
       342

SQL> select count(*) from link_loc;

  COUNT(*)
----------
       353

SQL> select count(*) from link_speed;

  COUNT(*)
----------
       353

SQL>
```

Run SQL Command Line

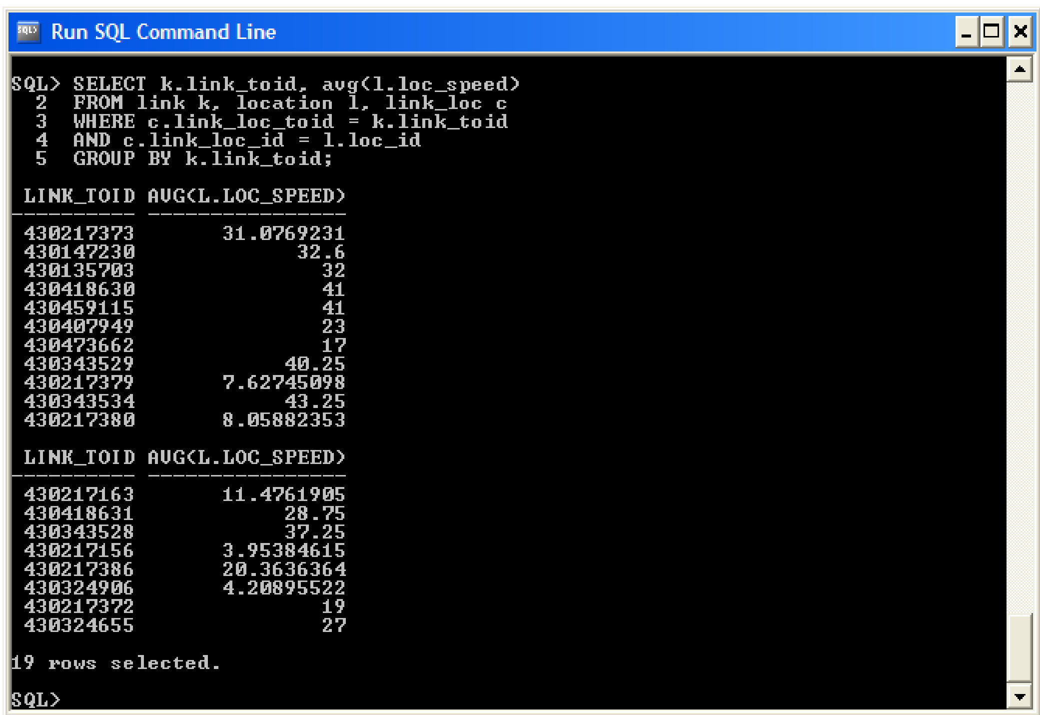
```
BIN$ZwbFRJKXTjG+fuLUgimtww==$0 TABLE
BIN$5DrrKSepSIuVeJz84BMSag==$0 TABLE
BIN$bSKU4JeuQ4KS8XCNejg//g==$0 TABLE
BIN$rVUzbZv2S+W576Eno9pTSg==$0 TABLE
BIN$BbUcyNakRu+SIqQcQs43xw==$0 TABLE
BIN$Z95Nrh0pTw2obcGAZq+gUA==$0 TABLE
BIN$WwASAWphTvmbWDabwqeTHg==$0 TABLE
BIN$zKfTFGmBQFeBnxau9r1tVA==$0 TABLE
BIN$GKuIzEn3Tf+h+PBBo0DqiA==$0 TABLE

TNAME                          TABTYPE  CLUSTERID
------------------------------ ------- ----------
BIN$PvIsncPhR5OjOULGf12LJA==$0 TABLE
BIN$e8CQ3KFgSXacvc4SUkujWw==$0 TABLE
BIN$jexTN0U8SNmZ68CK2re/gA==$0 TABLE
LINK                           TABLE
LOCATION                       TABLE
LINK_LOC                       TABLE
LINK_SPEED                     TABLE
MDRT_37A8$                     TABLE
MDRT_37AF$                     TABLE

53 rows selected.

SQL> _
```

Run SQL Command Line

```
SQL> describe link;
 Name                                      Null?    Type
 ----------------------------------------- -------- ----------------------------

 LINK_TOID                                 NOT NULL NUMBER
 LINK_ROAD_NAME                                     VARCHAR2(80)
 LINK_GEOM                                          SDO_GEOMETRY

SQL> describe location;
 Name                                      Null?    Type
 ----------------------------------------- -------- ----------------------------

 LOC_ID                                    NOT NULL NUMBER
 LOC_POLL_TIME                                      TIMESTAMP(0)
 LOC_GEOM                                           SDO_GEOMETRY
 LOC_SPEED                                          NUMBER(2)

SQL> describe link_loc;
 Name                                      Null?    Type
 ----------------------------------------- -------- ----------------------------

 LINK_LOC_TOID                                      NUMBER
 LINK_LOC_ID                                        NUMBER

SQL> _
```



```
Run SQL Command Line
SQL> describe location;
 Name                                      Null?    Type
 ----------------------------------------- -------- ----------------------------

 LOC_ID                                    NOT NULL NUMBER
 LOC_POLL_TIME                                      TIMESTAMP(0)
 LOC_GEOM                                           SDO_GEOMETRY
 LOC_SPEED                                          NUMBER(2)

SQL> describe link_loc;
 Name                                      Null?    Type
 ----------------------------------------- -------- ----------------------------

 LINK_LOC_TOID                                      NUMBER
 LINK_LOC_ID                                        NUMBER

SQL> describe link_speed;
 Name                                      Null?    Type
 ----------------------------------------- -------- ----------------------------

 LINK_SPEED_TOID                                    NUMBER
 LINK_SPEED_AVG_SPEED                               FLOAT(126)
 LINK_SPEED_GEOM                                    SDO_GEOMETRY

SQL> _
```



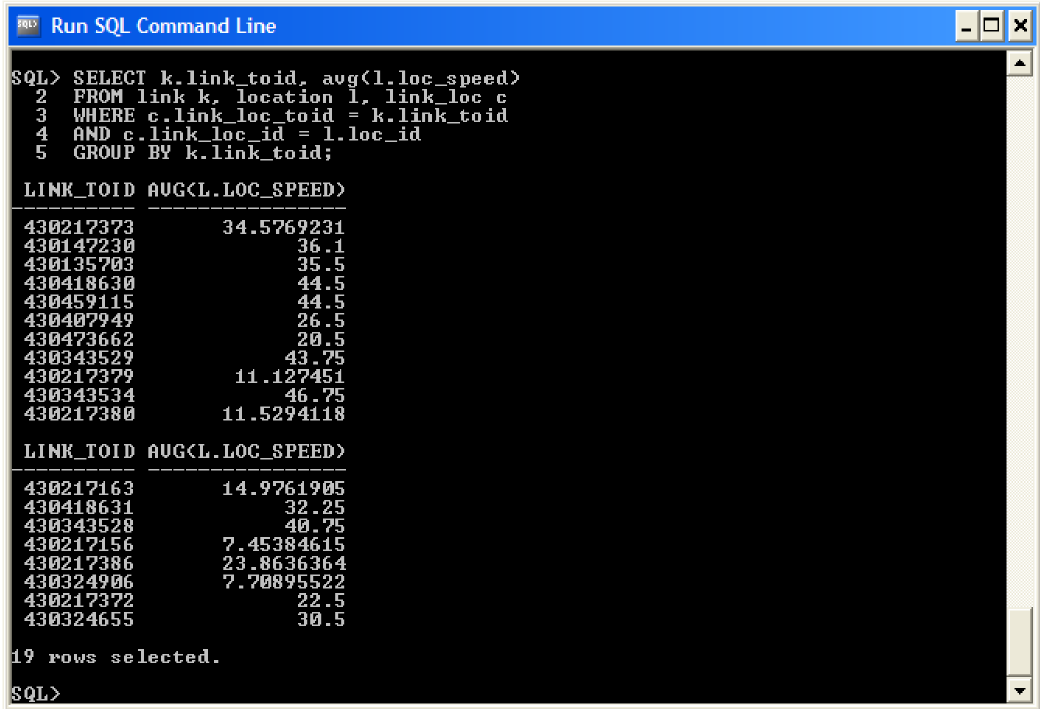
```
Run SQL Command Line

SQL> SELECT k.link_toid, avg(l.loc_speed)
  2  FROM link k, location l, link_loc c
  3  WHERE c.link_loc_toid = k.link_toid
  4  AND c.link_loc_id = l.loc_id
  5  GROUP BY k.link_toid;

 LINK_TOID AVG(L.LOC_SPEED)
---------- ----------------
 430217373       31.0769231
 430147230             32.6
 430135703               32
 430418630               41
 430459115               41
 430407949               23
 430473662               17
 430343529            40.25
 430217379       7.62745098
 430343534            43.25
 430217380       8.05882353

 LINK_TOID AVG(L.LOC_SPEED)
---------- ----------------
 430217163       11.4761905
 430418631            28.75
 430343528            37.25
 430217156       3.95384615
 430217386       20.3636364
 430324906       4.20895522
 430217372               19
 430324655               27

19 rows selected.

SQL>
```

Run SQL Command Line

```
  2  FROM link k, location l, link_loc c
  3  WHERE c.link_loc_toid = k.link_toid
  4  AND c.link_loc_id = l.loc_id
  5  GROUP BY k.link_toid;

 LINK_TOID AVG(L.LOC_SPEED)
---------- ----------------
 430217373       31.0769231
 430147230             32.6
 430135703               32
 430418630               41
 430459115               41
 430407949               23
 430473662               17
 430343529            40.25
 430217379       7.62745098
 430343534            43.25
 430217380       8.05882353

 LINK_TOID AVG(L.LOC_SPEED)
---------- ----------------
 430217163       11.4761905
 430418631            28.75
 430343528            37.25
 430217156       3.95384615
 430217386       20.3636364
 430324906       4.20895522
 430217372               19
 430324655               27

19 rows selected.

SQL> edit insert_loc_2.sql

SQL> start insert_loc_2.sql
```

Run SQL Command Line

```
1 row created.


1 row created.


1 row created.

SQL> select count(*) from location;

  COUNT(*)
----------
       684

SQL> select count(*) from link_loc;

  COUNT(*)
----------
       706

SQL> select count(*) from link_speed;

  COUNT(*)
----------
       706

SQL>
```



Run SQL Command Line

```
SQL> SELECT k.link_toid, avg(l.loc_speed)
  2  FROM link k, location l, link_loc c
  3  WHERE c.link_loc_toid = k.link_toid
  4  AND c.link_loc_id = l.loc_id
  5  GROUP BY k.link_toid;

 LINK_TOID AVG(L.LOC_SPEED)
---------- ----------------
 430217373       34.5769231
 430147230             36.1
 430135703             35.5
 430418630             44.5
 430459115             44.5
 430407949             26.5
 430473662             20.5
 430343529            43.75
 430217379        11.127451
 430343534            46.75
 430217380       11.5294118

 LINK_TOID AVG(L.LOC_SPEED)
---------- ----------------
 430217163       14.9761905
 430418631            32.25
 430343528            40.75
 430217156       7.45384615
 430217386       23.8636364
 430324906       7.70895522
 430217372             22.5
 430324655             30.5

19 rows selected.

SQL>
```

Run SQL Command Line

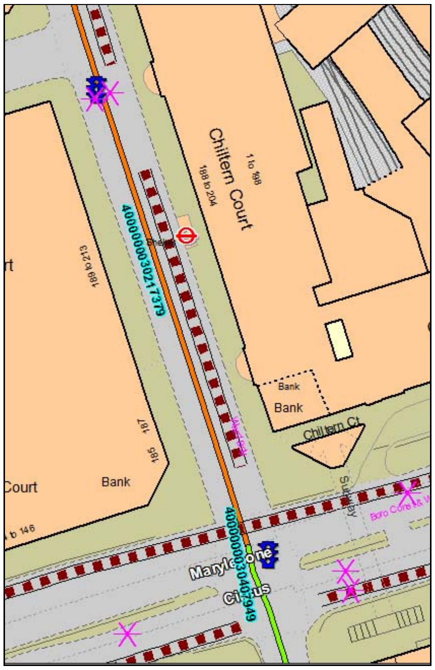
```
LOC_POLL_TIME
---------------------------------------------------------------------------
LOC_GEOM(SDO_GTYPE, SDO_SRID, SDO_POINT(X, Y, Z), SDO_ELEM_INFO, SDO_ORDINATES)
--------------------------------------------------------------------------------

 LOC_SPEED
----------
       332
15-MAR-05 15.41.34
SDO_GEOMETRY(2001, 81989, SDO_POINT_TYPE(528167, 181183, NULL), NULL, NULL)
        41


    LOC_ID
----------
LOC_POLL_TIME
---------------------------------------------------------------------------
LOC_GEOM(SDO_GTYPE, SDO_SRID, SDO_POINT(X, Y, Z), SDO_ELEM_INFO, SDO_ORDINATES)
--------------------------------------------------------------------------------

 LOC_SPEED
----------
       333
15-MAR-05 15.41.35
SDO_GEOMETRY(2001, 81989, SDO_POINT_TYPE(528169, 181171, NULL), NULL, NULL)
        41


    LOC_ID
----------
LOC_POLL_TIME
---------------------------------------------------------------------------
LOC_GEOM(SDO_GTYPE, SDO_SRID, SDO_POINT(X, Y, Z), SDO_ELEM_INFO, SDO_ORDINATES)
--------------------------------------------------------------------------------

 LOC_SPEED
----------
       334
15-MAR-05 15.41.36
SDO_GEOMETRY(2001, 81989, SDO_POINT_TYPE(528172, 181162, NULL), NULL, NULL)
        40


    LOC_ID
----------
LOC_POLL_TIME
---------------------------------------------------------------------------
LOC_GEOM(SDO_GTYPE, SDO_SRID, SDO_POINT(X, Y, Z), SDO_ELEM_INFO, SDO_ORDINATES)
--------------------------------------------------------------------------------

 LOC_SPEED
----------
       335
15-MAR-05 15.41.37
SDO_GEOMETRY(2001, 81989, SDO_POINT_TYPE(528175, 181152, NULL), NULL, NULL)
        34


684 rows selected.

SQL>
```

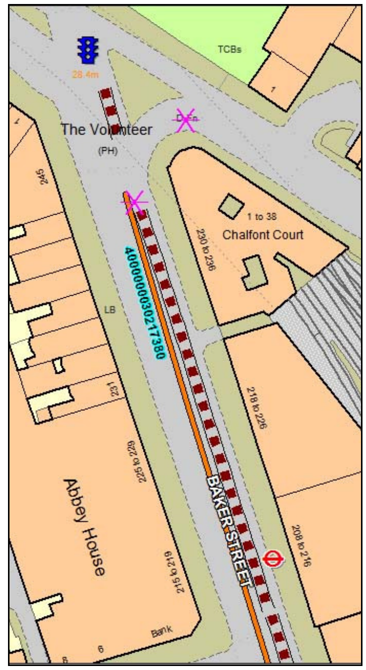
Run SQL Command Line

```
     430217163         643
     430217163         644
     430217163         645
     430217163         646
     430217163         647
     430217163         648
     430217163         649
     430217163         650
     430217163         651
     430217163         652

LINK_LOC_TOID LINK_LOC_ID
------------- -----------
     430217163         653
     430217163         654
     430217163         655
     430217163         656
     430217163         657
     430217163         658
     430418631         658
     430418631         659
     430418631         660
     430418631         661
     430343528         662

LINK_LOC_TOID LINK_LOC_ID
------------- -----------
     430343528         663
     430343528         664
     430343528         665
     430418630         666
     430343534         667
     430343534         668
     430343534         669
     430343534         670
     430343529         671
     430343529         672
     430343529         673

LINK_LOC_TOID LINK_LOC_ID
------------- -----------
     430343529         674
     430459115         674
     430459115         675
     430147230         675
     430147230         676
     430147230         677
     430147230         678
     430147230         679
     430147230         680
     430147230         681
     430147230         682

LINK_LOC_TOID LINK_LOC_ID
------------- -----------
     430147230         683
     430147230         684

706 rows selected.

SQL> _
```

APPENDIX D

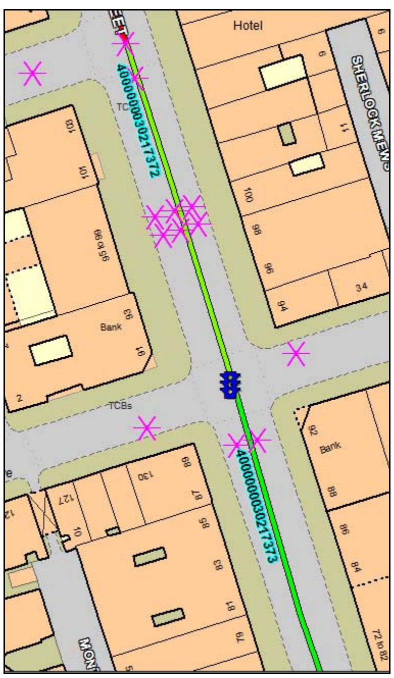
Mapped Results

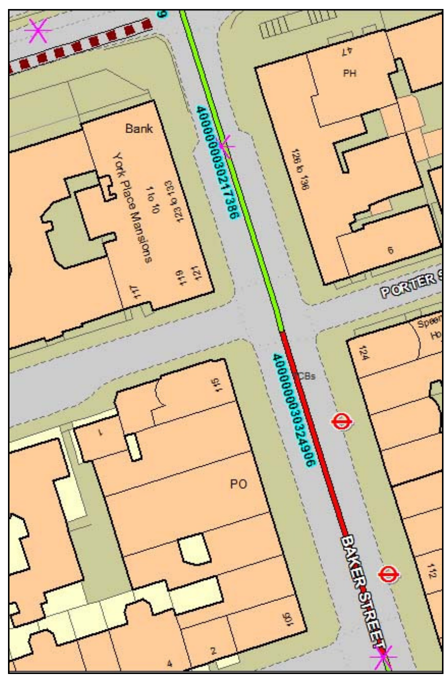

Chiltern Court
1 to 188
188 to 204
189 to 213
4000000030217379
Bank
Chiltern Ct
Court
Bank
187
185
Marylebone Circus
4000000030402949

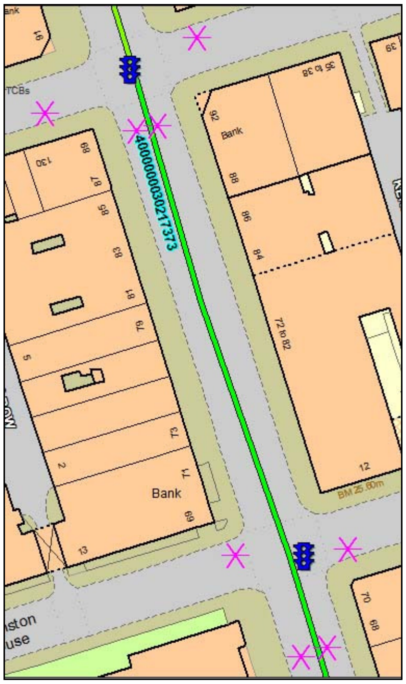

TCBs
The Volunteer
(PH)
1 to 38
Chalfont Court
230 to 236
4000000030217380
LB
231
218 to 226
225 to 229
Abbey House
BAKER STREET
215 to 219
208 to 216
Bank

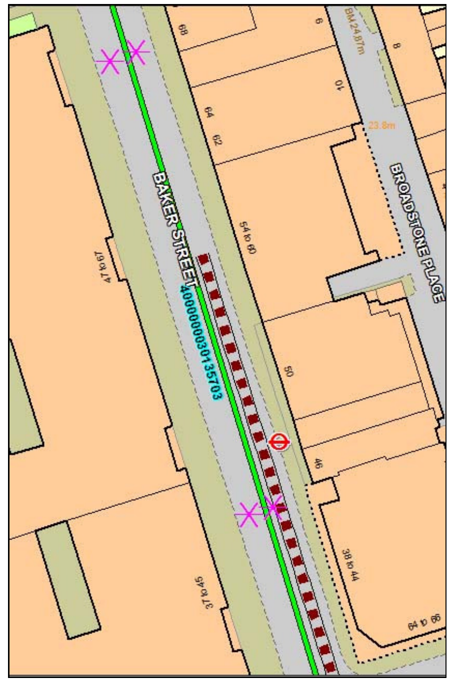
(1) (2)

(3) (4)


Hotel
4000000030217372
SHERLOCK MEWS
Bank
TCBs
Bank
4000000030217373

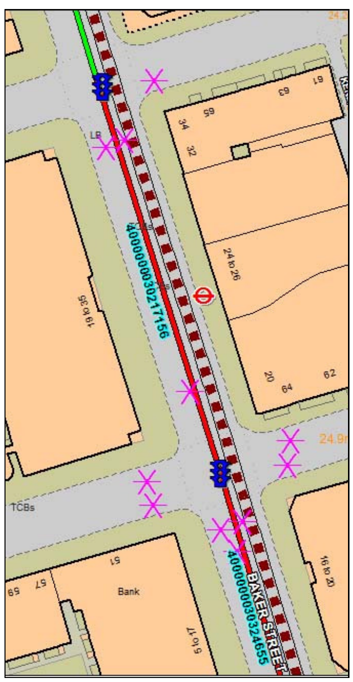

Bank
York Place Mansions
1 to 10
4000000030217386
PH
128 to 136
PORTER S
4000000030324906
TCBs
PO
BAKER STREET

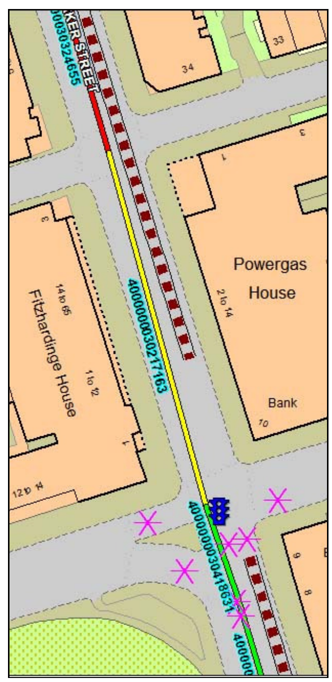

4000000030217373
Bank
72 to 82
BM 25.60m

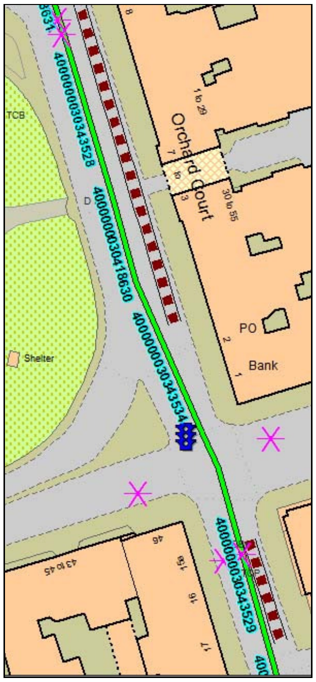

BAKER STREET
4000000030135703
BROADSTONE PLACE
54 to 60
47 to 57
37 to 45
38 to 44
64 to 66

(5) (6)

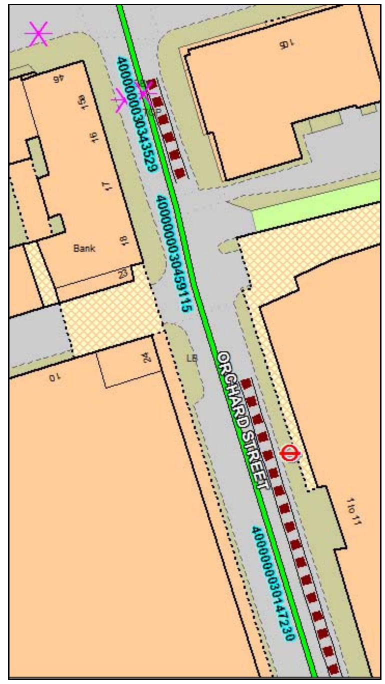
(7) (8)

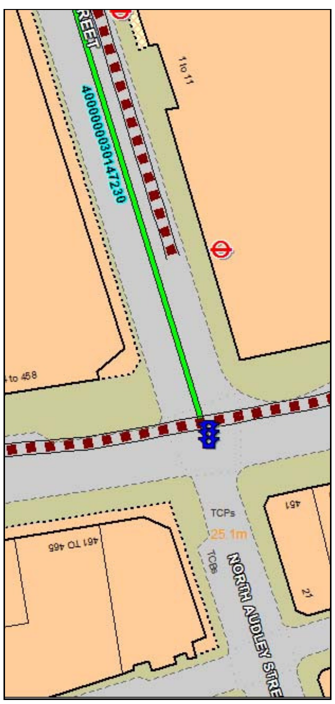

4000000030217156
24 to 26
18 to 35
Bank
TCBs
4000000030324655
BAKER STREET


4000000030217163
Powergas
House
2 to 14
Bank
Fitzhardinge House
14 to 65



(9) (10)





(11)

Symbols

Bus Stop

Traffic Lights

Loop Detectors

Bus Lane

Legend

Link Speeds:

25 to 50 mph

15 to 25 mph

10 to 15 mph

5 to 10 mph

0 to 5 mph

< This page is intentionally blank >