South African Journal of Information Management
[ DoE accredited ]

Peer Reviewed Article Vol.6(2) June 2004

# Spatial information systems in managing public transport information

**C. Dondo**
chiedzad@yahoo.com

**U. Rivett**
Department of Civil Engineering
University of Cape Town
Rondebosch
Cape Town
ulrike@eng.uct.ac.za

**Contents**



## 1 Introduction

Traffic congestion is one of the major problems plaguing urban cities all over the world today. With traffic congestion, goods and services experience delays in accessing markets and employees are late for work. This results in reduced production. At the same time, idling vehicles waste fuel and cause air pollution. This situation is the same in South Africa, where millions of rand are lost each year on congestion costs (CERF 1999). Studies carried out in Cape Town, South Africa, project that if no immediate action is taken to reduce traffic congestion, by the year 2020, the car population will increase by 64%, thereby reducing highway traffic speeds to 29 kilometres per hour (South Africa. National Department of Transport 1999).

Worldwide, policies to reduce congestion levels include, among other things, the promotion of public transport (Climate Change 2001; Cullinane and Cullinane 2003). As public transport has a better cost and space effectiveness for transportation of volumes of people, having more people use public transport will lead to a reduction in the number of vehicles on the road. This will, in turn, reduce traffic congestion and its effects (De Saint Laurent 1998).

To dissuade people from private car use to public transport, public transport should be as flexible as the private car (Wendell Cox Consultants 1996) and some of the unattractive characteristics associated with public transport, ranging from lack of security to slowness of services, have to be removed. In Cape Town, South Africa, some of the reasons why public transport is unpopular with users include long travel times and long walking distances to the nearest bus stop.

To improve public transport services, planners need to determine what the unattractive characteristics are, how serious the situation is and their effects on the usage of public transport. This involves the collection of data on public transport usage and services.

Public transport data can be classified into two types, non-spatial data and spatial data. Non-spatial data are data with no location information and are mostly obtained from surveys, for example the ride check surveys carried out in Cape Town. The data also include the service timetables, fares and layouts of stops. Spatial data are data with location information (x and y co-ordinates according to a certain reference system) and can be represented by features on a map. Examples of spatial data include the road network, the stops and the routes.

Public transport services planners need to integrated these two types of data and analyse them in order to extract meaningful information for decision support on planning and development of public transport infrastructure and services. Considering the dynamic nature of the urban environment, it means that this data need frequent updating. Since most of the data are voluminous and spatial in nature, a spatial information system facilitates the storage, integration, display, update and analysis of the data (Zhong 1998; Papacostas and Prevedourous 1993; Etches, Parker, Ince and James. 2000; Miller and Storm 1996; Goodchild 1996).

The first stage in developing a spatial information system is the needs assessment and data discovery stage, during which the type of analysis that need to be performed and the data required or used are noted. The second stage, which is the most important stage, is the creation of the data model to accurately abstract data from the real world into the database. The last stage is the development of a prototype. In this research, the prototype was developed for Blauuwberg Municipality, in Cape Town. Some of the sample analyses that can be performed are discussed.

top

## 2 Public transport information management conceptual data model

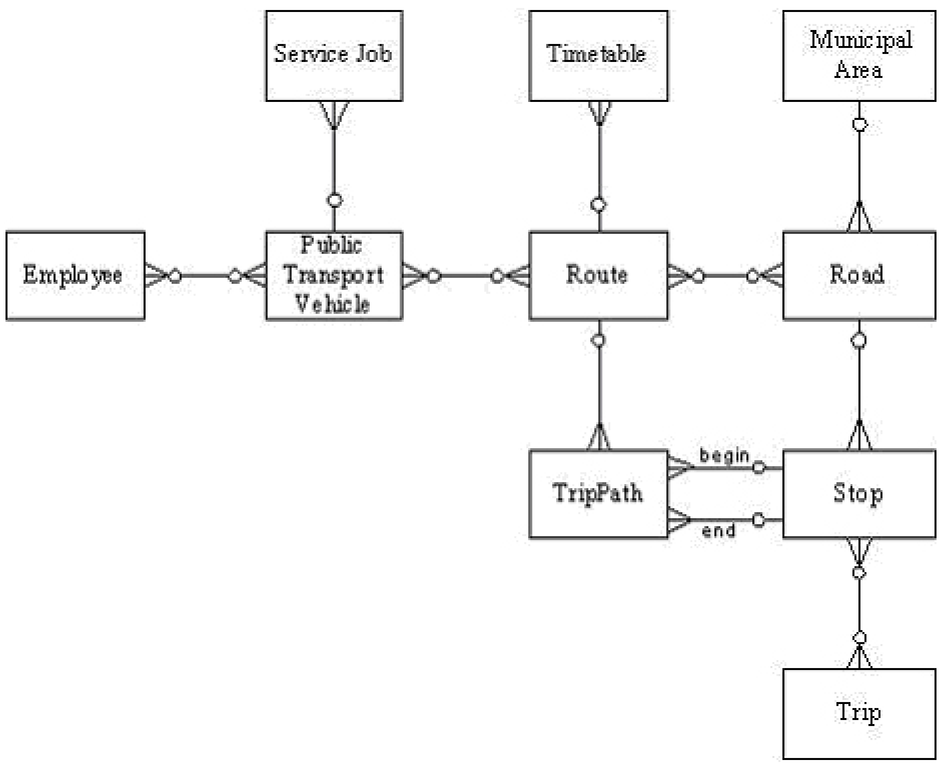
The conceptual data model defines what data are to be included in the spatial information system and how the data are connected and related. It is a model independent of the implementation software. In this article, the conceptual data model is represented using the entity-relationship model, which is the most widely used. This model views the world as being made up of entities and relationships. An entity is a thing that can be uniquely identified, for example a specific car or a specific person. Entities can be grouped into classes of similar things known as entity types. A relationship is an association among entities or entity types (Chen 1976). An important aspect of a relationship is its cardinality or degree. There are three kinds of relationship degrees, namely the:

- one-to-one relationship
- one-to-many relationship
- many-to-many relationship.

The only variations on these arise from how each of the entities is participating in the relationship. If a relationship is such that each entity must participate in the relationship, then the relationship is obligatory; otherwise it is optional or non-obligatory.

Before creating the conceptual data model, the designer notes the process performed, the data required or available for use in these processes and that have to be included in the database. From this data, the entities are identified and the relationships defined. In this article, the identified entities and relationships are illustrated in the conceptual data in Figure 1 below. The diagram is based on the Crow Foot entity-relationship diagramming notation (Connolly and Begg 2002).

**Figure 1** Conceptual data model for public transport information management




The municipal area represents the area in which a road is located. A town is divided into different municipal areas. In each area, the municipality is responsible for the supply and maintenance of infrastructure. A road represents the network that is used daily by different modes of transport. A road must be located in one municipal area, but one municipal area may have more than one road in its area.

The route is a physical feature representing a path taken by a public transport vehicle travelling along the transport network. A route may use one or more roads and a road may have one or more routes using it.

The timetable represents any information on the daily times and services available for each mode of public transport. A route may have one or more timetables but a timetable must be for one particular route.

Public transport vehicle represents all the different modes of public transport, that is, buses, trains and minibus taxis. Service job is a record of any maintenance or repair work performed on a public transport vehicle. A public transport vehicle may have one or more service jobs but a service job must be for only one public transport vehicle. The employee category contains information on the drivers of public transport vehicles. A driver may drive one or more public transport vehicles and a public transport vehicle may be driven by more than one driver.

The trip and trip path entities are included in the data model to cater for the inclusion of ride check survey data into the system. A trip represents the journey taken by a public transport vehicle from an origin stop to a destination stop. A trip path represents the section of a route. It is the path from one stop to the next stop. A trip path must begin at one stop and end at another stop, but a stop may have many trip paths passing through it. A route may have many trip paths but one trip path must be for one route.

Stop represents a place where the public transport vehicle stops to pick up and drop off passengers during any particular trip. Stops include bus stops, bus termini[1], train stations, minibus taxi ranks and trip stops. A trip stop, in the case of representing the ride check survey data, represents a point where the bus stopped to drop off or pick up passengers but which is not a designated stop.

The conceptual data model has to be transformed into a logical data model, which can be implemented in a GIS software product.

top

### 3 Public transport information management logical data model

Whereas the conceptual model discussed in Section 2 above is just a schema of the data to be included in the spatial information system and how the data are related, the logical data model is a model that can be implemented in a particular software product of choice. The conceptual model has to be transformed into a logical data model. Various models can be used in achieving this, but the object-relational model was used in this research.

The object-relational model captures information from the conceptual data model and represents it as a collection of two-dimensional tables in the database. Each table is given a name. It is made up of columns or named attribute types and rows or records. Each row contains information relating to one entity and must therefore have a unique identifier. The intersection of each row and column in the table contains an attribute value (Codd 1970; Worboys 1995; Laurini and Thompson 1996).

The creation of the logical data model involves removing redundant data and restructuring some relations that cannot be handled by the software (known as normalization). Figures 2, 3, 4 and 5 below illustrate, with examples, how this is achieved for the type of relationships in the conceptual data model.

### 3.1 Route- timetable relationship

Figure 2 below illustrates the structure of the one-to-many relationship between route and timetable from the conceptual data model.

**Figure 2** One-to-many relationship between route and timetable

| Timetable Name | Timetable ID | RouteName | RouteID |
|---|---|---|---|
| B-CT1 | T1 | Blauuwberg to Cape Town | RT1 |
| B-CT2 | T2 | Blauuwberg to Cape Town | RT1 |
| S-CT1 | T3 | Sea Point to Cape Town | RT2 |
| S-CT2 | T4 | Sea Point to Cape Town | RT2 |
| S-CT2 | T5 | Sea Point to Cape Town | RT2 |

From Figure 2 above, the relationship between route and timetable leads to the existence of some redundant data in the database. For example, the route names, ‘Blauuwberg to Cape Town’ and ‘Sea Point to Cape Town’ have been repeated. This increases the size of the database. Another problem is that each time a new timetable is added into the database for one of these routes, the route name has to be retyped. If there are any spelling mistakes, then some of the information is not retrieved when queries are performed. To structure this relationship properly, two tables are created, one for route and one for timetable, and the RouteID is posted into the timetable table (see Figure 3 below). The same procedure is performed for the other one-to-many relationships in the conceptual data model.

**Figure 3** Resulting tables after structuring the route-timetable relationship

| Timetable Name | Timetable ID | RouteID |
|---|---|---|
| B-CT1 | T1 | RT1 |
| B-CT2 | T2 | RT1 |
| S-CT1 | T3 | RT2 |
| S-CT2 | T4 | RT2 |
| S-CT2 | T5 | RT2 |

| RouteName | RouteID |
|---|---|
| Blauuwberg to Cape Town | RT1 |
| Sea Point to Cape Town | RT2 |

### 3.2 Route-road relationship

Figure 4 below illustrates the structure of the route-road relationship from the conceptual data model.

**Figure 4** Many-to-many relationship between route and road

| RouteName | RouteID | RoadName | RoadID |
|---|---|---|---|
| Blauuwberg to Cape Town | RT1 | Blauuwberg Road | RD1 |
| Blauuwberg to Cape Town | RT1 | Koeberg Road | RD2 |
| Atlantis to Cape Town | RT2 | Blauuwberg Road | RD1 |
| Atlantis to Cape Town | RT2 | Koeberg Road | RD2 |

If the relationship were to be represented as illustrated in Figure 4, there would be occurrences of redundant data in the database. For one or more routes sharing the same roads, the road name is repeatedly entered. Also, for roads sharing similar routes, the route name is repeatedly entered.

Unlike the one-to-many relationship, which can be structured by posting the RouteID into the Timetable table, the many-to-many relationship can only be structured well by constructing a third table to represent the relationship (see Figure 5). A combination of the RouteID and RoadID uniquely identifies a row and this type of identifier is called a composite identifier.

**Figure 5** Three tables constructed to structure the many-to-many relationship between the road and the route

| RouteName | RouteID |
|---|---|
| Blauuwberg to Cape Town | RT1 |
| Atlantis to Cape Town | RT2 |

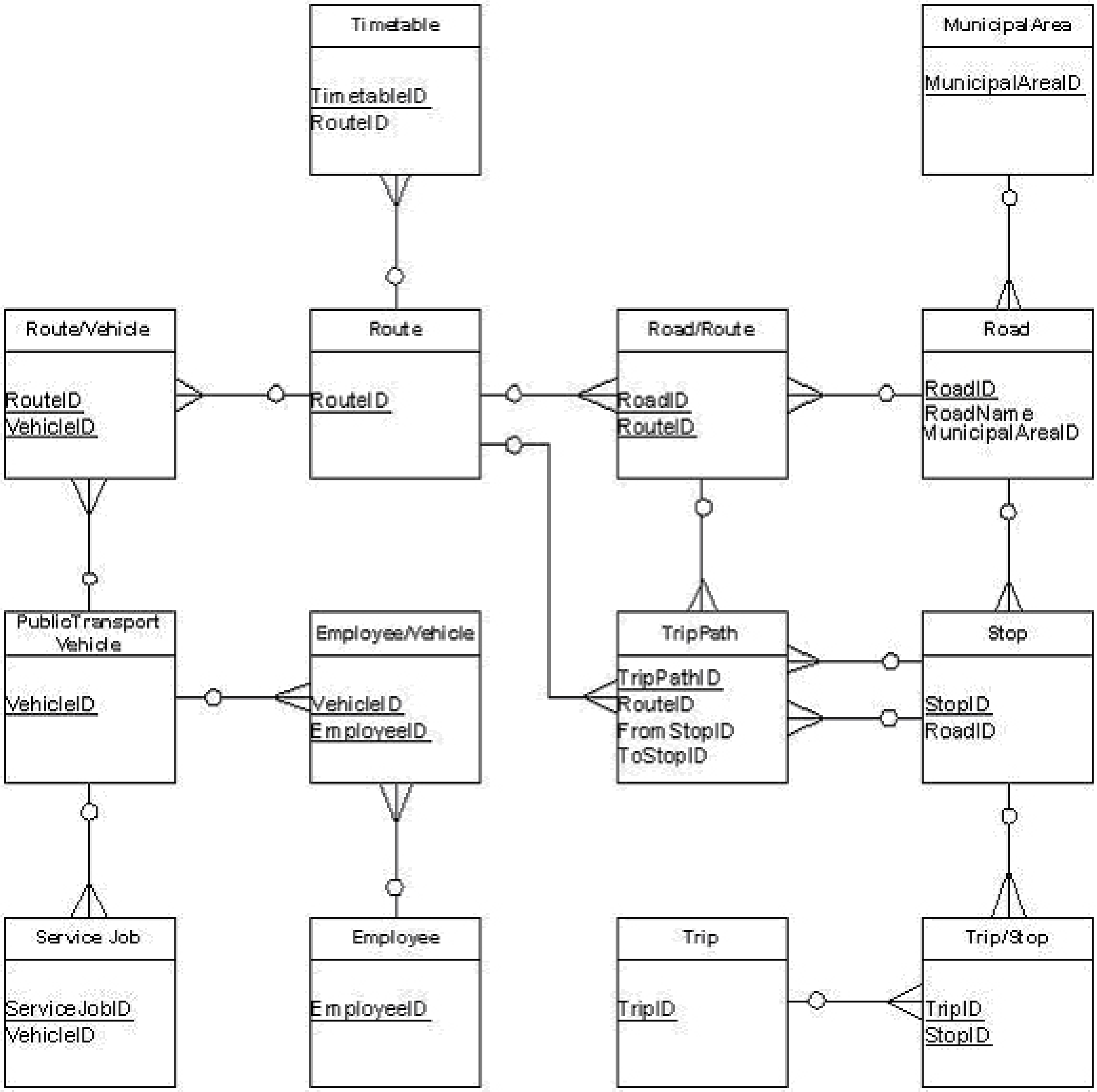
| RoadName | RoadID |
|---|---|
| Blauuwberg Road | RD1 |
| Koeberg Road | RD2 |

| RouteID | RoadID |
|---|---|
| RT1 | RD1 |
| RT1 | RD2 |
| RT2 | RD1 |
| RT2 | RD2 |

Relationship table

After removing occurrences of redundant data and restructuring the relationships, the resulting logical data model shown in Figure 6 is ready for implementation in the GIS software of choice.

**Figure 6** Logical data model for public transport information management

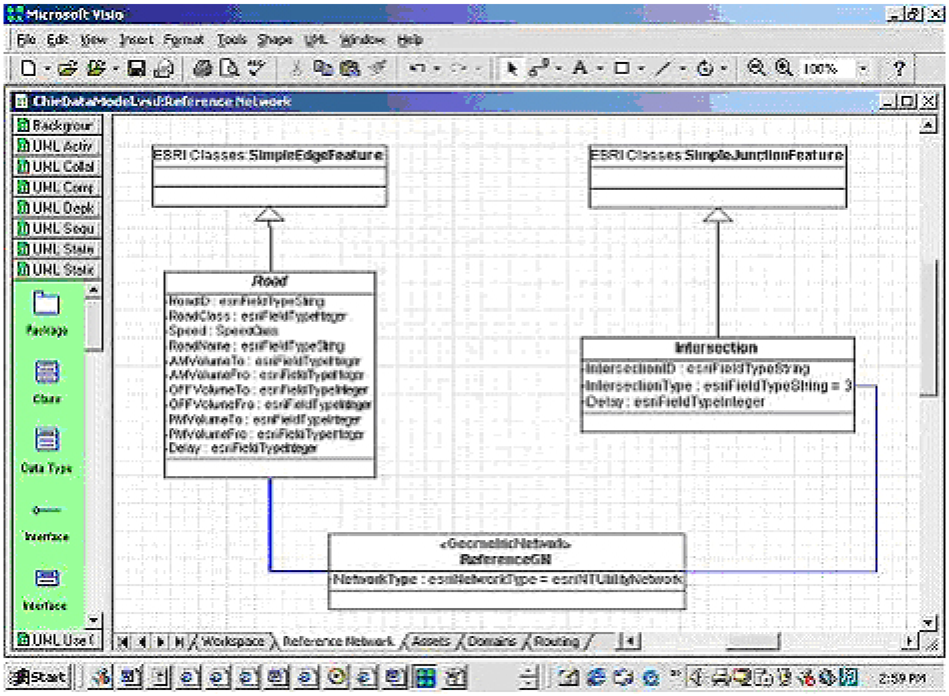



top

## 4 Creation of the prototype

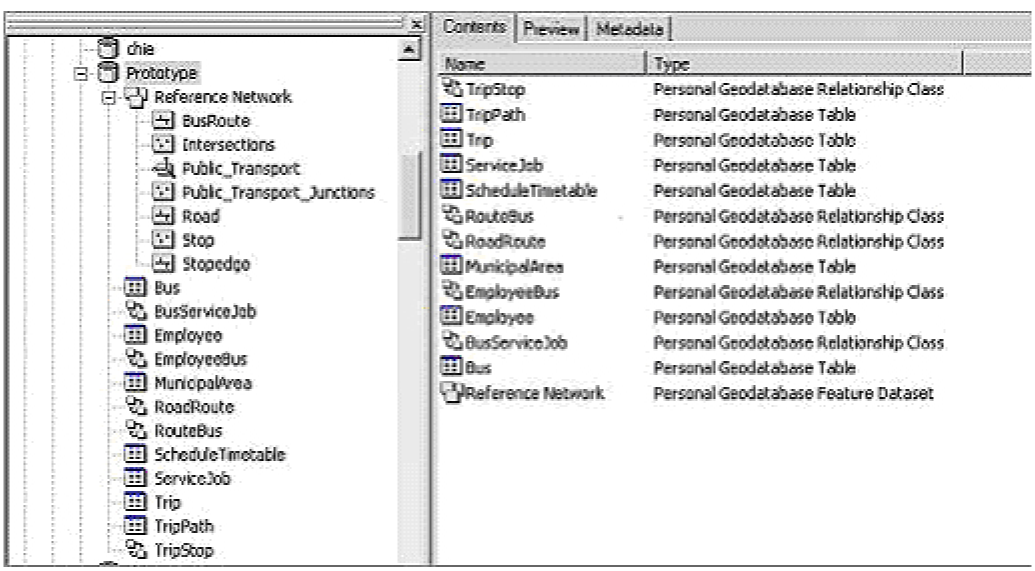
The prototype was developed for Blauuwberg Municipality in Cape Town and the public transport mode selected was buses. For designing the database using the object-oriented geodatabase model in ArcInfo 8.2, the Computer Aided Software Engineering (CASE) tools available in the Unified Modelling Language (UML) in Microsoft Visio 2000 Enterprise were used. Using UML, first, a schema, like the one illustrated in Figure 7 below was created. This schema was based on the ArcGIS Transportation Model, which was developed by ESRI (Zeiler 1999; Curtin, Noronha, Goodchild and Grise 2001).

**Figure 7** Created UML schema




After the schema was created using UML, it was exported to the Microsoft repository. The created Microsoft repository was then imported into ArcCatalog, where a personal geodatabase (the one called Prototype, shown highlighted on the left-hand side of the box in Figure 8) was generated.

**Figure 8** Geodatabase created in ArcInfo 8.2




In the geodatabase, the representations of the objects were as follows:

- The road was modelled as a network feature. A network is made up of edges and junctions. The edges correspond to road sections and the junctions to road intersections.
- The route was created through a process called dynamic segmentation. During dynamic segmentation in ArcInfo, the user selects sections of the road that are required to represent

the route. A section can be made up of the entire arc or a fraction of each arc defining the road. The user then instructs the program to create a route based on the selected sections.
- The stops were represented as a separate layer of point features with x and y co-ordinates. As explained above, a stop can either be a bus stop, a bus terminus, or an undesignated place where the bus stops to drop off and pick up passengers during the ride check survey. These three variations of stop were represented using subtypes. The three subtypes inherited the attributes and properties of the stop object. The subtypes could be represented on the map using different symbols (see Figure 10 below).
- The non-spatial objects were modelled as tables. Tables were created in the geodatabase for employee, service job, timetable, trip path and municipal area. The relationships between the objects were modelled as relationship classes.

The geodatabase was then populated with the available data. A stand-alone application was developed using ESRI MapObjects Version 2 and Microsoft Visual Basic 6.0 to enable easy information retrieval. The created system was then available for querying. Some of the typical queries and analyses are discussed below.

---

top

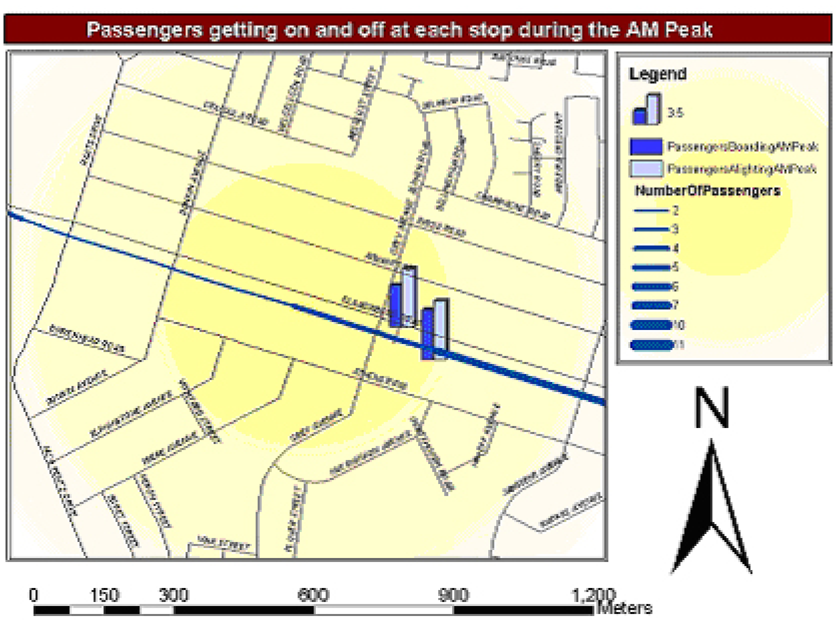
## 5 Functions of the system

Some of the queries and types of analysis that can be performed in the designed system are the following:

### 5.1 Passenger analysis

These are performed to display and output information on the amount of passengers travelling along a section of a particular route or getting on and off at each stop (see Figure 9).

**Figure 9** Graphs show the number of passengers boarding and alighting at each stop during the am peak[2]



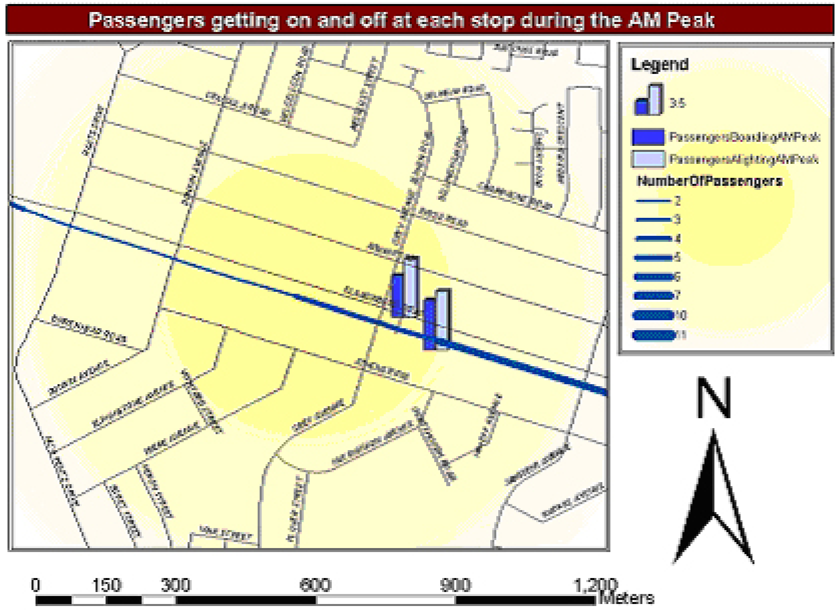
**5.2 Accessibility analysis**
This type of analysis is usually performed when it is required to determine the number of people living within a distance of a stop or a terminus. The census data are used to provide the numbers. This kind of information is used to assess how well the stops are serving the customers. It can also be used to estimate the probable utilization statistics for a proposed stop location. Other additional information that may be used is statistics on car ownership in each area.

**5.3 Shortest path analysis**
This is one of the most important applications of spatial information systems in transport. It is the calculation of the optimal path between any two points, marking the start and end of a proposed route. This assists public transport operators when reassigning public transport vehicles to more efficient routes, that is, routes that are less congested and will take less time to traverse through. The shortest path is calculated by minimizing the weights assigned by the user to each route section. The shortest path can be the shortest distance or the shortest time, depending on the weights assigned. The following weights are considered in this system:

- The volume of traffic in each segment of the road for the three different peak periods, that is the morning peak (am peak), during the day (off peak)[3] and the evening peak (pm peak)[4]
- time delays along each segment of the road based on the maximum speed allowable for the road segment. This time is calculated by dividing the length of each road section by the maximum allowable speed along that road, that is, Time = Length (m) / Speed ($ms^{-1}$); and
- Time delays incurred when a vehicle waits for its turn to cross an intersection, either a signalized intersection (one with traffic lights) or an unsignalized (one without traffic lights) intersection.

**5.4 Ride check survey analysis**
Figure 10 presents the results of the ride check survey carried out in Cape Town in the year 2000. Information can be obtained per section of a route, per particular stop or aggregated for the whole route, depending on the kind of analysis being performed. The information can be presented as text in reports or graphs on maps. Figure 10 below illustrates the amount of passengers who were travelling along a section of a route during the ride check survey.

**Figure 10** Amount of passengers who were travelling along a section of the route during the survey




top

## 6 Conclusion

This article documents the creation of a spatial information system for public transport information management. The designed spatial information system assists in analysing the data and enables the output of most of the information in the form of maps, graphs and reports, which are easy to interpret. This is important when presenting information on any service improvements to policy makers or potential project fund providers.

The developed system was distributed to potential users, namely transport professionals, together with a manual and a questionnaire. Most of the professionals selected did not have prior experience of using a GIS software package. They all agreed that the system and the user manual were user-friendly.

Another advantage of the system is that most of the queries are predefined and ready for output; the user does not see the underlying database. The stand-alone application can run on any Windows Operating System. It was tested on Windows 2000 Professional and Windows 98 with the same results. No special software is required to run it.

The data model used to design the database for the system is based on the object-relational model, and can therefore be used in relational databases. For example, anyone using Microsoft Access or ArcView can use the same data model. The data model is general to public transport and so other modes of public transport besides buses can also be added into the system. Any other survey data available, for example data from passenger counting at termini or the data from the bus transit surveys are also supported by the system.

As mentioned above, shortest path analysis can be based on time delays incurred when a vehicle waits for its turn to cross an intersection, at either a signalized intersection (one with traffic lights) or an unsignalized (one without traffic lights) intersection. The formula used for calculating the delay requires detailed data on the volumes of traffic going in the various directions at an intersection. The data are obtainable through traffic counts performed for an intersection. For the traffic count to be truly representative, the count should be done for at least three days. This requires a lot of financial resources and manpower and hence such data might not be readily available in the transport department. Such detailed data were not available when the system was developed. If someone wants to use this component of the system, the correct data will have to be collected first.

The deduction from the developed system is that GIS software is effective in managing, analysing and displaying information. One of the arguments against GIS software is its inability to support transport modelling and forecasting. Although such a module can easily be created using the Visual Basic for Applications programming language available, this requires programming expertise. Ultimately, spending resources on purchasing the software and hiring employees when ready-made dedicated modelling software is available on the market cannot be justified. In addition, most departments that have been using dedicated transport modelling and forecasting software have had to consider GIS software in order to overcome one of the main shortfalls of dedicated planning or modelling software, namely the inability to output analysis and modelling results in graphical form. A possible solution would be that a number of government departments pool their resources and purchase GIS software products that can be shared by all.

top

**7 References**

CERF. 1999. *Efficiency in action: saving time, money and the environment*. [Online]. Available WWW: http://www.cerf.org/iiec/news/121997/i8.htm (Accessed 20 February 2002).

Chen, P.P.S. 1976. The entity relationship model – towards a unified view of data. *Association of Computing Machinery Transactions on Database Systems* 1(1):9-36.

Codd, E.F. 1970. A relational model of data for large shared data banks. *Communications of the ACM* 13(6).

Connolly, T. and Begg, C. 2002 *Database systems – a practical approach to design, implementation, and management*. 3rd edition. United States of America: Addison Wesley.

*Climate Change. Mitigation*. 2001. Intergovernmental Panel on Climate Change [Online]. Available WWW: http://www.ipcc.ch/pub/tar/wg3/368.htm#928 (Accessed 17 April 2002).

Cullinane, S. and Cullinane. K. 2003. Car dependence in a public transport dominated city: evidence from Hong Kong. *Transportation Research D: Transport and Environment* 8(2).

Curtin, K., Noronha, V., Goodchild, M.F. and Grise, S. 2001. ArcGIS transportation data model (draft). Environmental Systems Research Institute, Inc.

De Saint Laurent, B. 1998. Overview of Urban Transport in South Africa. In: *Proceedings of the 8th CODATU Conference. Cape Town, September 1998*. [Online]. Available WWW: http://www.dbsa.org/privatesector/ppinote6.htm (Accessed 15 March 2002).

Etches, A., Parker, D., Ince, S. and James, P. 2000. UTIS (urban transportation information system). A geo-spatial transport database. In: *ACM Symposium on GIS, Washington D.C., November 2000*.

Goodchild, M.F. 1996. *Geographical information systems and disaggregate transportation modeling*. [Online]. Available on the World Wide Web. http://www.ncgia.ucsb.edu/vital/research/pubs/mfg9607.pdf (Accessed 15 October 2001)

Laurini, R. and Thompson, D. 1996. *Fundamentals of spatial information systems*. Academic Press Limited.

Miller, H.J. and Storm, J.D.1996. Geographical information system design for network equilibrium-based travel demand models. *Transportation Research Part C* 4(6):373-389.

Papacostas, C.S. and Prevedourous, P.D. 1993. *Transportation engineering and planning*. 2nd edition. Prentice Hall Inc.

South Africa. National Department of Transport. 1999. *Moving South Africa*. [Online]. Available WWW: http://www.transport.gov.za/projects/msa/index.html (Accessed 15 January 2001).

Wendel Cox Consultants. 1996. Urban transport: from theory to reality. *The Public Purpose* (11). [Online]. Available WWW: http://www.publicpurpose.com//pp-nz91.htm (Accessed 10 April 2002).

Worboys, M.F. 1995. *GIS: a computing perspective*. Taylor and Francis Limited.

Zeiler, M. 1999. *Modelling our world*. The ESRI guide to geodatabase design. Environmental Systems Research Institute Inc.

Zhong, R.P. 1998. Internet GIS and its applications in transportation. *TR News* (March-April).



top

ISSN 1560-683X
Published by InterWord Communications for the Centre for Research in Web-based Applications,
Rand Afrikaans University