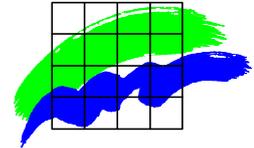


Time-and-space modelling of public transport systems using GIS

Per Thorlacius, M. Sc.
National Environmental Research Institute
Frederiksborgvej 399, P.O. Box 358
DK-4000 Roskilde



Abstract

Modelling the movement of vehicles and passengers in time and space in public transport systems is generally a rather complicated task. When finding the shortest path through a public transport network, not only the link impedance has to be taken into account, but also the time at which the link occurs, i.e. the time the bus actually rides the actual segment. This type of modelling requires an otherwise complex topological network. This paper describes how a geographic information system (GIS) can be used to create this kind of topology and how a model of a public transport system is implemented.

Introduction

When comparing public and individual transport systems one significant difference is the degree of dependence between the travel duration and the exact time one wants to travel. In public transport systems these parameters are closely related to each other, for individual transport this is much less the case.

This difference makes the movement of persons and vehicles in public transport systems more difficult to model than in individual systems.

The duration of a travel in a public transport system (the travelling time) is defined as the sum of the terms waiting time, in-vehicle travel time and interchange time (1).

$$t = t_w + t_v + t_i \quad (1)$$

These terms again, are defined as follows (with busses as an example):

- The waiting time is the time the passenger has to wait for the bus, i.e. the time from the moment he or she decides to travel to the time the bus is boarded;
- The in-vehicle travel time is the time the passenger has to ride the bus;
- The interchange time is the time the passenger has to wait for a connecting bus at another stop if it is necessary to change busses on the way.

When examining public transport the time aspect has to be taken into account both in a relative and in an absolute way: The three different types of times defined above are all relative, but they are determined by the absolute time at which the passenger wants to travel and the absolute time at which a connection between the origin and destination stops actually exists.

Conventional models of public transport systems do not take the exact absolute time into consideration and therefore only average values of travelling times can be calculated. Furthermore, these models can neither take connection correspondences nor connection frequency distribution in time into account. A conventional model can hence not be used to answer questions like *Which is the fastest way to go from A to B, Wednesday morning at 8:17?*

To do this, a model of a more complex topological nature is required. Complex networks of this kind have been rather difficult to represent in a GIS up until now. However, using the data structure proposed in this paper, new topological features of the geographic information system ARC/INFO provide a way of solving this problem.

1. Conventional Models for Public Transport Systems

In conventional models of public transport systems each transit line is typically represented by a set of links between the stops that that particular line visits. [Harris et al. (1992), Ortúzar et al. (1994)] This type of representation is shown in figure 1.

Figure 1 shows two transit lines serving seven stops in all. Line 1 starts at A, ends at E and visits the stops B, C and D. Line 2 starts at I, ends at H and also visits the stops B, C and D. The representation is not directed, each line on the figure and accordingly in the model represents both directions of that line. To represent the two transit lines in the figure, eight links are needed in the model.

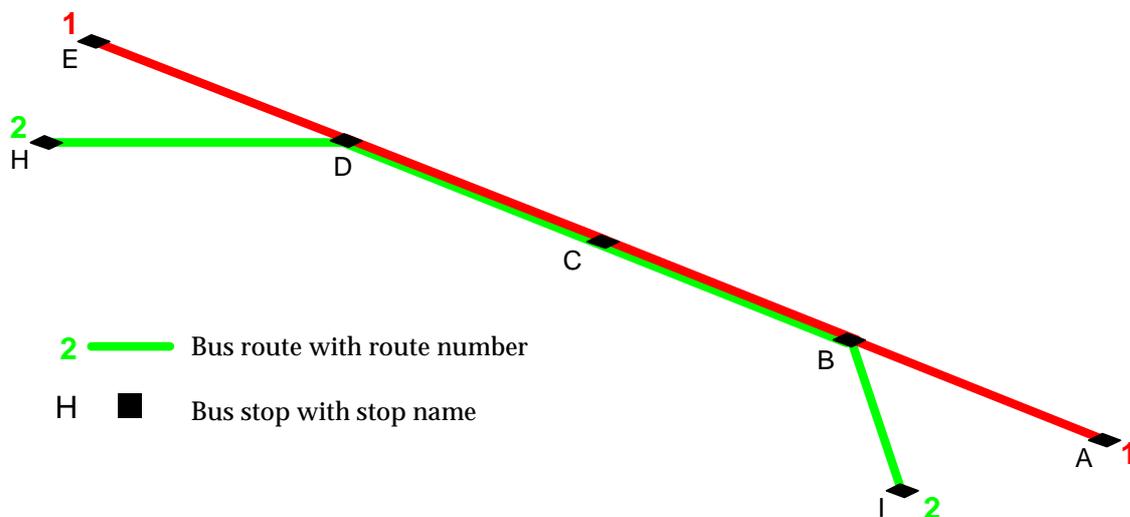


Figure 1: The traditional way of modelling public transport in a transport model

These links have feature attributes describing the in-vehicle travel time and the frequency by which the service of transport is provided. The frequency is calculated for the period of time the model is valid, e.g. the morning rush or the whole day. The commonly known transport models EMME/2 and TRIPS both follow this principle. [Ortúzar et al. (1994), HTM (1996)]

The in-vehicle travelling times are obtained directly from the link attributes, waiting and interchange times at the stops are calculated from the link frequencies. However, since the frequencies represent the number of connections between stops for a certain period of time,

the calculated waiting and interchange times are correspondingly average values for that specific period of time.

Neither can a model of this type take correspondences between transit connections into account. Therefore, models of this type often lead to particularly inaccurate results when used in transit systems with low frequencies, i.e. in rural areas where the few connections often are equally well co-ordinated to correspond with each other.

Furthermore, a model of this type can not take the frequency distribution in time into account. If the model covers the whole day, the travelling times calculated with the model will correspondingly be average travelling times over the day. Typically, the waiting and interchange times in the rush hour are shorter due to the high frequencies with which the services are provided than for example in the evening. Equally, the in-vehicle travel times are typically longer in the rush hour due to congestion.

2. The Proposed Public Transport Model

To calculate travelling times more accurately, a different type of model is hence needed. The data model proposed in this paper is called the Three Dimensional Vehicle Run Representation, and is developed by NERI as a part of the ALTRANS project.

2.1. The Principal Concept of the Proposed Model

The conceptual idea behind the proposed model is to look upon a transport model as a projection of streets, bus and train lines onto the plane. In this sense, conventional transport models are two dimensional models.

The idea behind the proposed model is to represent time as the third dimension, and to take each physical vehicle running on the transit lines into consideration. Each vehicle is then represented by a link between points not only in the plane but also in time. The concept of the proposed data model is shown in figure 2 below.

The lower part of figure 2 shows the map of the bus lines from figure 1. Above this map a grid of dotted lines is spanned. The perpendicular lines in this grid represent the same position in the plane, the horizontal lines the same position in time.

In this grid the links of three different vehicle runs are shown. A vehicle run is defined as the movement of a physical vehicle (train, bus, etc.) in one direction from the first stop the vehicle visits to the last. The links thus represent the physical movement of the individual vehicles in time and space.

Due to the fact that one link is needed for each vehicle run between two stops, a total of twelve links is needed to model these three vehicle runs. As opposed to the model shown in figure 1, the proposed model is a so called directed graph as all links are one-way links. [Cormen et al. (1989)]

For reasons of simplicity only the vehicle runs in the general direction of the X-axis are shown in the figure, in reality vehicles would be running in both directions on each particular line. Note that the direction of the time axis in figure 2 is down, so that the only possible direction of movement is downward.

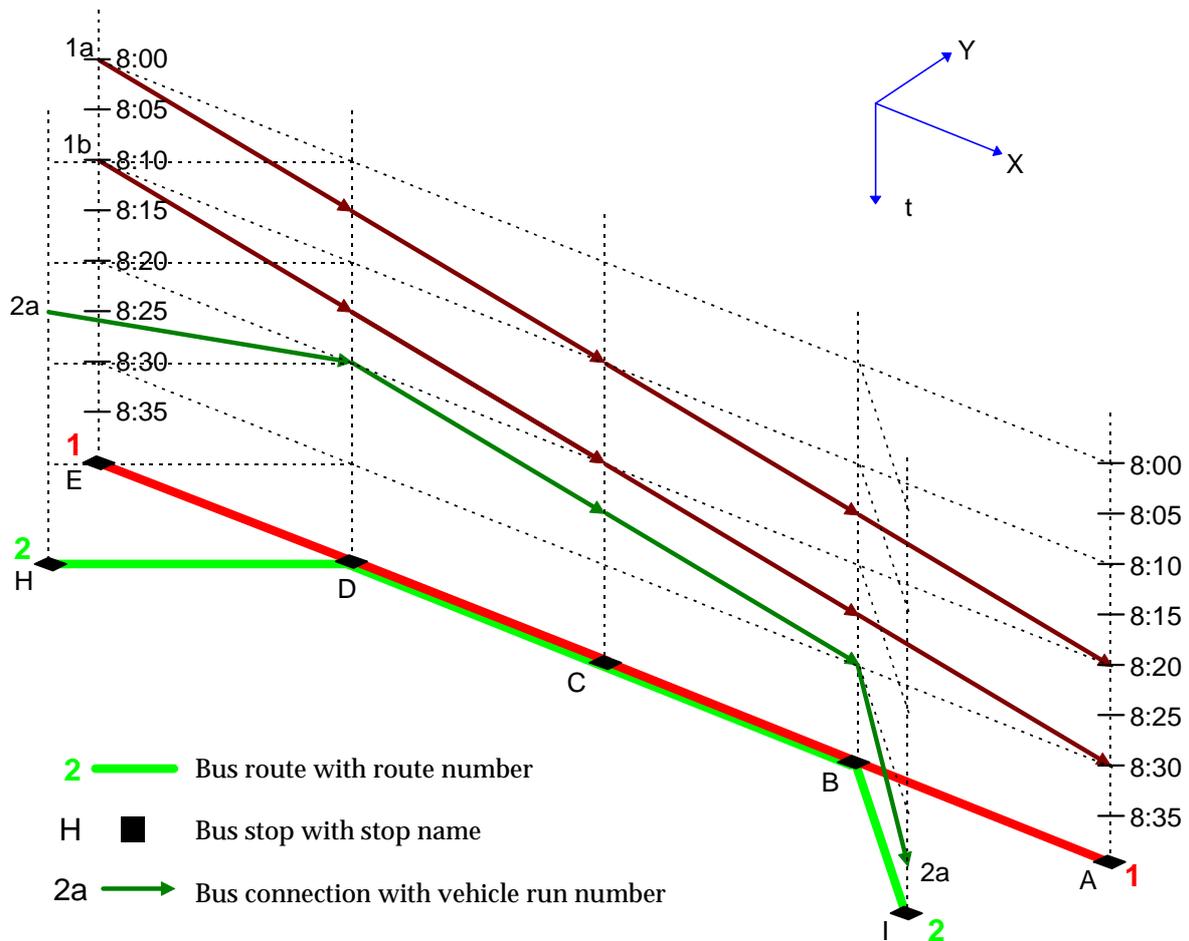


Figure 2: The proposed 3D vehicle run representation of modelling public transport

In figure 2 line 1 runs from A to E via B, C and D and vice versa. The bus connection with the vehicle run number 1a, departs from A at 8 o'clock and arrives at H at 8:20. On the way the stops D, C and B are visited at 8:05, 8:10 and 8:15 respectively.

If a passenger wants to travel from A to I, and arrives at the stop at A at 8:05, he or she has to wait five minutes before the first bus is departing from the stop at 8:10. This bus has vehicle run number 1b and arrives at B at 8:25. Here the passenger has to wait five minutes until a connection to I is provided at 8:30. This connection arrives at I at 8:35. The total travelling time is thus 30 minutes, of which five minutes are waiting time, five minutes interchange time and 20 minutes in-vehicle travel time.

2.2. The Representation of Changing Possibilities in the Proposed Model

A model as the one shown in figure 2 can be used to determine the movement of vehicles in time and space. However, to model the movement of passengers, changing possibilities at each stop must be provided in the model. These possibilities are represented by the so called interchange links.

In conventional models for public transport systems, each stop is represented by one node, in some models by one node for each line serving the stop. In the proposed model however, there is one node for each unique point of time at which a vehicle arrives or departs from the stop. A transit stop is then represented by a group of nodes as shown in figure 3.

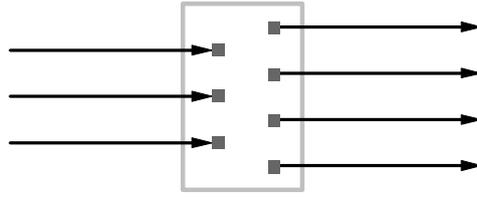


Figure 3: The representation of a transit stop in the proposed model as a group of nodes

As shown in figure 4, the representation of interchange links in a model like the one proposed can be accomplished using several principles.

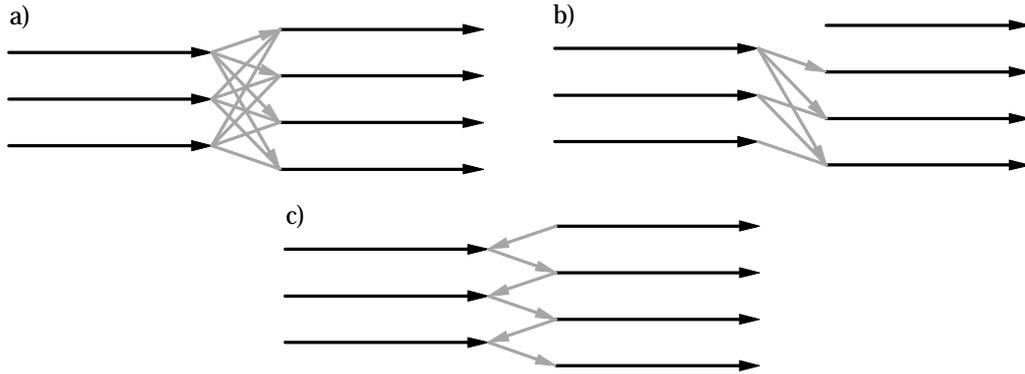


Figure 4: Different principles of representing the interchange links (grey) at each stop

a) The arriving and departing links can be coupled in a combinatorial manner, so that each arriving link is connected to all of the departing links. If n denotes the number of arriving links, and m the number of departing ones, this results in a total number of interchange links i given by (2).

$$i = n \cdot m \quad (2)$$

One can assume, that the number of arriving vehicles equals the number of departing ones, otherwise the vehicles would accumulate at the stop. Therefore it is reasonable to assume, that:

$$i \approx n^2 \quad (3)$$

b) It is not possible however, to change to a connection that has already left the stop. If this is taken into account, the number of interchange links needed can be reduced. Assuming that the number of departing vehicles approximately equals the number of arriving ones, and that the distribution of arrivals and departures is approximately equal, the approximate number of interchange links needed is then given by (4). [Spiegel (1968)]

$$i \approx 1 + 2 + 3 + \dots + (n - 1) + n = \frac{1}{2}n(n + 1) = \frac{1}{2}n^2 + \frac{1}{2}n \quad (4)$$

c) It is, however, possible to reduce the number of necessary links additionally by using a serial approach. If the arriving and departing nodes are ordered in time, each node can be coupled to the next one in time. This gives a total number of necessary interchange links given by (5).

$$i = n + m - 1 \approx 2n - 1 \quad (5)$$

As seen, it is associated with significant advantage to chose the last representation principle, as the number of necessary interchange links only increases proportionally to the number of

arrivals. Using the first two representation principles the number of necessary interchange links increases with the square of the number of arrivals.

When using the so called modified Dijkstra path finding algorithm, the algorithmic complexity for sparse graphs is given by (6), where P is the number of points (nodes) in the graph, E the number of edges (links) and O the order of complexity. [Cormen et al. (1989)]

$$O(E \log P) \tag{6}$$

It may therefore be concluded that both computational and storage advantages are obtained by choosing the interchange link representation principle c).

For practical reasons, the interchange links are represented as spirals, which when projected onto the plane as on figure 5 below, become circles.

2.3. The Calculation of Waiting Times in the Proposed Model

As with conventional transport models, it is necessary to connect the network with points at which travels (trips) have their origin and destination, i.e. where travels are generated and attracted. As with conventional models this is done with connection links to the centroids of the so called transport analysis zones, as shown in figure 5. In the proposed model however, these links are also used to calculate the waiting times at the stops, and to model the changes between services using separate networks, e.g. bus and train.

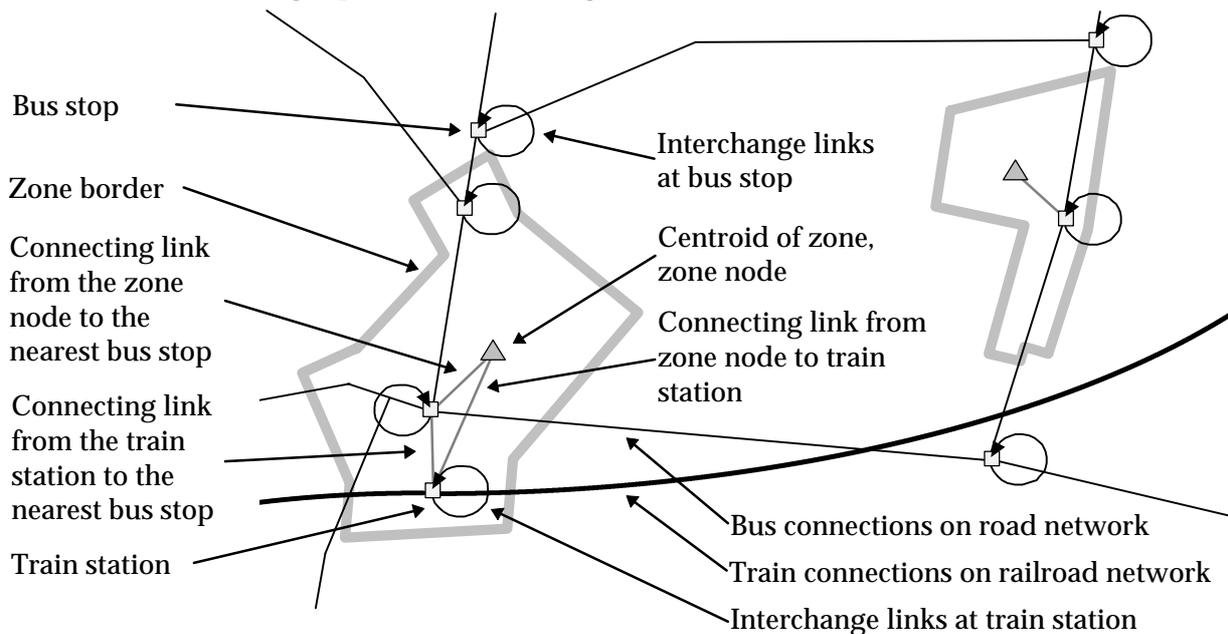


Figure 5: Objects in the implemented model

The link impedances on the connection links going from the zone node to the departing links are calculated dynamically according to the exact time at which the travel starts, i.e. the time the passenger wants to travel. The travel start time is subtracted from the departure time of the nodes of the departing links. The result of this subtraction is the time the passenger has to wait from the travel start time until each vehicle departs. The impedance of the connection links is thus the waiting time. If the waiting time for a particular departure is negative, that departure has already left the stop, and the link can not be traversed.

3. Implementation

The model described in chapter 2 has been implemented for the use in relation to the ALTRANS project, initiated by the Department of Policy Analysis at NERI.

The objective of the model is to calculate travelling times corresponding to a set of survey travel data (trips). These travelling times are then used to calibrate an econometric transport mode choice and trip length generation model to investigate political instruments that can lead to change of transport behaviour, and the environmental changes as a result hereof.

3.1. Topological Implementation

The proposed 3D representation is implemented using new topological features of the geographic information system ARC/INFO 7.1. With these features, it is possible to build line topology having more than one node at the same position in the plane. The third dimension is established through this topological capability.

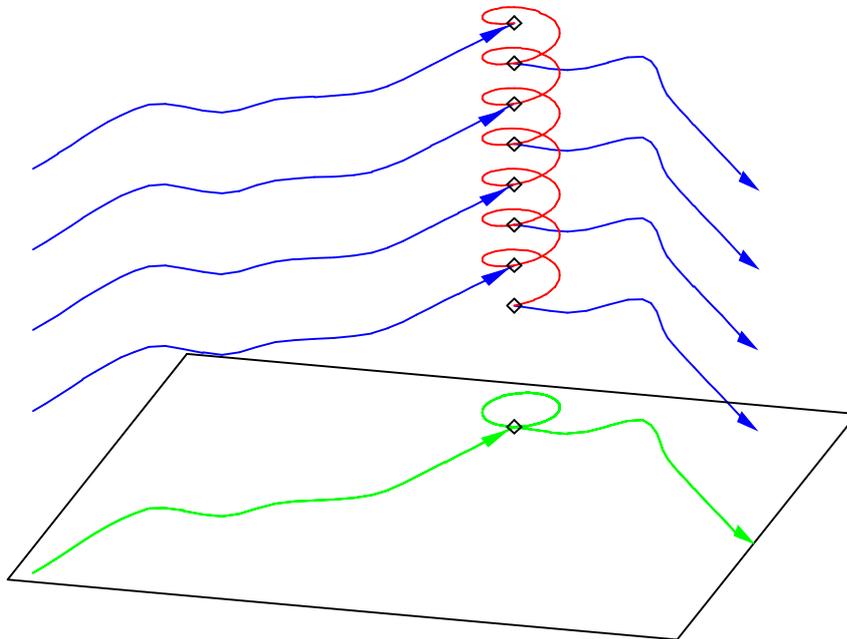


Figure 6: 3D view of the proposed model showing vehicle run and interchange links

Figure 6 shows a 3D view of the proposed model for a stop with four arriving vehicles and four departing ones. To model the interchange possibilities at the stop, a total of seven interchange links is needed. The lines at the bottom show the projection of the model onto the plane represented by the rectangle.

3.2. The Construction and Operation of the Model

The model is constructed in several steps using data from a variety of different sources. Certain data are stored in tabular form in a relational database, others are geographical data stored in the geographic information system. Figure 7 shows the data flow in both the construction and operation phases of the model. The different data sets seen on figure 7 are described in the following section 3.3.

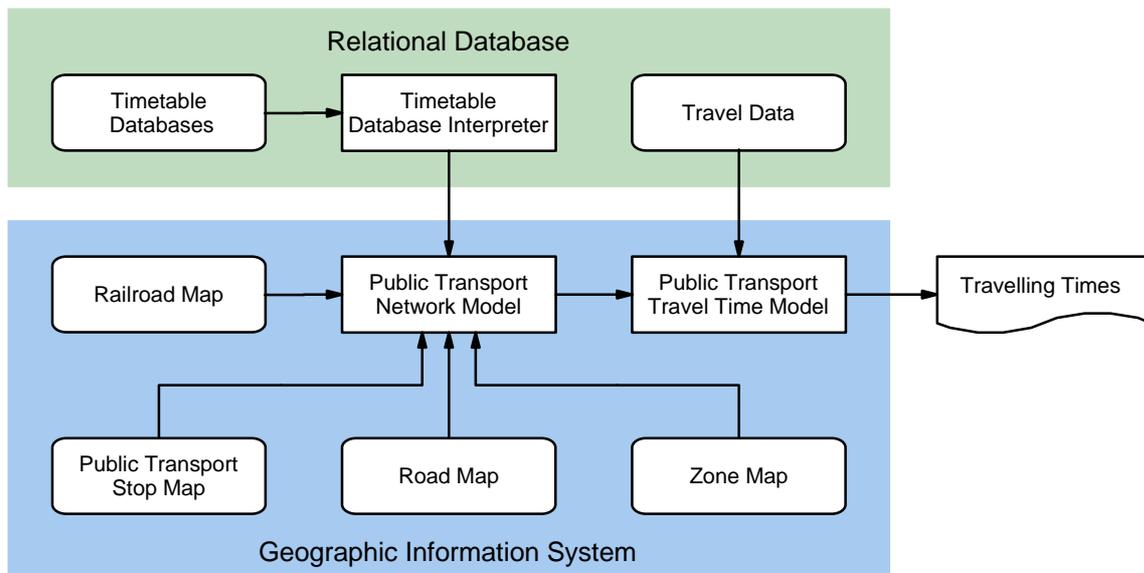


Figure 7: Data flow in the implemented model for the calculation of travelling times

The constructing of the public transport network model is as follows:

Through the timetable database interpreter a vehicle run is extracted from the timetable database. The vehicle run is then divided into segments, i.e. links between two consecutive stops. For each segment the shortest path on road or rail network is found according to vehicle type. It is assumed that this path is the route the vehicle follows in reality. The procedure is then repeated for all segments of the particular vehicle run, for all vehicle runs in the particular database and for all transport company databases. The interchange and connecting links are then constructed at each destination using the arrival and departure times from the links described above, and the model is completed.

In the operating phase the waiting times are calculated as described in section 2.3.

3.3. Data

Figure 8 shows an example of the timetable data used in the project. Totally, the five different database types contain about 150 MB of timetable data from 17 different transport companies.

TransportModeCode	Description	SerialNo	StopNo	ShortName	Arrival	Departure
0001	Inter City Train					
0002	High Speed Train					
0004	Regional Train	147	1	8600001	9999	440
0005	Private Rail Compa	147	2	8600007	501	501
0006	Local Train	147	3	8600009	511	512
0007	Inter Regional Train	147	4	8600013	520	521
0008	Night Train	147	5	8600015	528	529
		147	6	8600020	547	551
		147	7	8600027	606	607

VectorNo	Vector	ShortName	CountryCode	StationName
1	0111110011111001111100111110	8600001	0086	Frederikshavn
2	1000000100000010000001000000	8600005	0086	Kvissel
3	0111110011111001111100111110011	8600006	0086	Tolne
4	0111110011111001111100111110011	8600007	0086	Sindal
5	100000010000001000000100	8600009	0086	Hjørring

Figure 8: A simplified example of the time table data

Figure 9 shows the four different geographical data sets used for the implementation of the model. The upper left corner of the figure shows the road network, the upper right corner the railroad network. The two maps shown in the lower part of the figure are maps digitised by NERI itself. The public transport stop map (left) consists of about 4,500 destinations (bus stops and train stations) and the transport analysis zone centroid map (right) consists of about 1,500 zones.



Figure 9: The road map, the railroad map, the public transport stop map and the transport analysis zone centroid map. The extracts of maps belonging to The National Survey and Cadastre are reproduced according to the agreement G18/1997

Concluding Remarks

From the experiences gained during the course of the ALTRANS project, it may be concluded that considerable advantages are associated with modelling public transport systems using the proposed 3D approach in a GIS.

However, this type of modelling also has a number of disadvantages. The necessary amount of data and the rather specific demands for data quality are some of these. Furthermore, the complex nature of the model makes access to powerful hardware essential in order to keep the processing time at a level suitable for modelling purposes.

For a more detailed description of the model please refer to Thorlacius (1998).

References

- Cormen et al. (1990);** *Introduction to Algorithms*, Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, The MIT Press, Cambridge, Massachusetts 1990
- Harris et al. (1992);** *Transit Network Modelling - A New Approach*, Nigel Harris, London Underground Ltd., Martin Bach, MVA Systematica; In: Proceedings of the 4th International Conference on Microcomputers in Transportation, Baltimore 1992
- HTM (1996);** *Hovedstadstrafikmodel Version 3.0 - Beskrivelse af døgntrafikmodel*, Firma Anders Nyvig A/S, København 1996 (Danish)
- Ortúzar et al. (1994);** *Modelling Transport*, Second Edition, Juan de Dios Ortúzar, Luis G. Willumsen, John Wiley & Sons, Chichester, England 1994
- Spiegel (1968);** *Mathematical Handbook of Formulas and Tables*, Murray R. Spiegel, McGraw-Hill Publishing Company, New York 1968
- Thorlacius (1998);** *ALTRANS - Beregning af rejsetider for rejser med bil og kollektiv trafik*, Faglig rapport fra DMU, Per Thorlacius, Afd. for Systemanalyse, Danmarks Miljøundersøgelser, Roskilde 1998 (Danish)