

# Development of network data model for urban transportation system in GIS environment for the city of Plovdiv, Bulgaria.

**Lubomir Filipov, Boris Davidkov**

Department of Cartography and GIS,  
Sofia University “St. Kliment Ohridski”, Bulgaria.

## *Abstract*

*The urban transportation systems are very complex in nature, due to the fact that they combine different modes of transportation over a limited space in high-density areas with increasing transport demand. With variety of information needed in the field of Transportation, Geographic Information Systems have long been recognized as a valuable tool for the representation and analysis of transportation networks and related activities.*

*In this paper we make a brief review of the major GIS data models, related to the transportation. Based on that, we propose a methodology for development of Network data model in GIS environment for the city of Plovdiv, the second major city in Bulgaria. The GIS database presents all the components related to routing and analysis of multimodal transport systems, such as various modes of transportation, restrictions on the network and complex turns. This model is obligatory base for development of complex process models, such as 4-Step travel demand models, Accessibility models and others.*

Phd. Student, Lyubomir Filipov,  
Sofia University “St. Kliment Ohridski”, Faculty of Geology and Geography, Department of  
Cartography and GIS.  
Sofia, 1504; 15, Tzar Osvoboditel Blvd.,  
Mobile phone: (+359) 888 972 277 : Fax : (+359 2) 944 6487, e-mail: [lubomirfilipov@gmail.com](mailto:lubomirfilipov@gmail.com)

Assoc. Prof., Ph. D. Boris Davidkov,  
Sofia University “St. Kliment Ohridski”, Faculty of Geology and Geography, Department of  
Cartography and GIS.  
Sofia, 1504; 15, Tzar Osvoboditel Blvd.,  
Tel.:(+3592) 930 84 41; Fax: (+3592) 944 6487; e-mail: [boby@gea.uni-sofia.bg](mailto:boby@gea.uni-sofia.bg);

## **INTRODUCTION**

Transportation data encompasses a wide range of spatial data entities that are fundamental to many Geographic Information System (GIS) and cartographic applications. Transportation data is normally considered a key element of base maps and serves as essential reference data in this context. Besides its role as reference data, transportation data is at the core of applications, such as emergency response, routing, urban and regional planning, public transport, municipal service provision and general purpose mapping. Transport models are used for a wide range of purposes, from local traffic planning to managing infrastructure at a national level, as parts of other models (for example land-use models), and for providing the basis for evaluating the environmental impacts of traffic.

In general, topics related to GIS studies in the field of transportation can be grouped into three categories (Rodrigue, 2005):

- **Data representations.** How can various components of transport systems be represented in a GIS?
- **Analysis and modelling.** How can transport methodologies be used in a GIS?
- **Applications.** What types of applications are particularly suitable for GIS?

Data representation is a core research topic of GIS. Before a GIS can be used to tackle real world problems, data must be properly represented in a digital computing environment.

Except some logistics applications there is a lack of use of GIS functionality in the field of transportation in Bulgaria. Many government agencies and private companies still don't realize the potential and the importance of GIS for the transportation. The main goal of this paper is to provide the methodology for development of a sample network data model, which is the essential part of the transportation data models. In order to achieve this goal we have made a brief review of the existing transportation data models, we have clarified the required data and have build a functional network data model for the city of Plovdiv, the second major city in Bulgaria. The result is a basic network structure from which multiple views of the transportation network could be generated to accommodate multiple transportation modes for routing and analysis.

## **OVERVIEW OF TRANSPORTATION DATA MODELS AND STANDARDS**

GIS stores spatial data with logically linked attribute information in a GIS storage database where analytical functions are controlled interactively by a human operator to generate the needed information products (Tomlinson, 2003). This definition clearly describes the five basic components of any GIS system: data, GIS software, trained personal, specific procedures and off course the underlying hardware to run the whole system. Within these components the data is probably the most time consuming and expensive part. It is obligatory to determine what data is needed and what information products should be generated from this data. The goals and objectives of a certain application should be clearly stated and defined prior to the physical development of the data model.

With the growing development of GIS in the field of transportation there are quite a few data models and standards. A standard is "A set of criteria (some of which may be mandatory), voluntary guidelines, and best practices" (Interoperability Clearinghouse Glossary of Terms reference). Some of most prolific data standards in the filed of transportation are Geographic Data files, NCHRP and TIGER. The Geographic Data Files standard was developed as a draft submitted to the European Committee for Standardization. A consortium of public and private entities developed GDF to "improve the efficiency of the capture, the production and handling of road related geographic information" (GDF Overall Data Specifications, 2002). The National Cooperative Highway Research Program (NCHRP) in USA is designed to facilitate research to develop practical solutions to problems facing transportation agencies. One significant area of research is that of location referencing for transportation. A network structure for applications of GIS in transportation is the Topologically Integrated Geographic Encoding and Referencing (TIGER) files, accepted by the U.S. Bureau of the Census. TIGER files provide national coverage of the transportation network and other features.

A model is "a representation of a set of components of a process, system, or subject area, generally developed for understanding, analysis, improvement, and/or replacement of the process" ((Interoperability Clearinghouse Glossary of Terms reference). All GIS systems are build using formal models that describe how things are located in space. A formal model is an abstract and well-defined system of concepts (Zeiler, 1999). A geographic data model defines the vocabulary for describing and reasoning about the things that are located on the earth. The transport model is typically a mathematical model of a certain geographical area and the transportation system in this area (Israelsen and Frederiksen, 2005). The size of the model area, the level of detail of the model, and the complexity of the mathematics involved can vary substantially, depending on the purpose of the specific transport model. The most widely used transport model is the Travel Demand Model, which aims to assess the amount of transportation use there will be in a certain area and how that use will impact the transportation system. There are a large number of procedures for gathering and analyzing this information, all of which exist under the generic name of Travel Demand Modeling. Many travel demand models are variants of the Urban Transportation Planning System (UTPS) four-step travel demand model. UTPS models include a great deal of information on populations and socioeconomic conditions. A generic and practical orientated is the UNETRANS data model, which is a representation of the essential elements for a broad range of transportation-related functions. The UNETRANS model is intended as a tool to be used in support of the range of users who participate in transportation science, planning, or management and provides a suggested base from which to build a transportation-oriented geodatabase.

All these data models and standards are reviewed an taken into consideration for the developing of the network data model for the city of Plovdiv.

## **DESIGNING A NETWORK DATA MODEL**

The actual structure of the data is done by using some of the several logical data models: relational, object-oriented or object-relational (Tomlinson, 2003). Taking advantage of the object-relational data model a new data structure for network in geodatabases is being developed which allow for much more realistic modeling of network connectivity for

the purposes of routing and tracing. This data structure fully supports multimodal networks, such as found in transportation (car, bus, bike, rail, shipping). Network connectivity can be based on geometry using a very rich connecting model, as well as database relationships, such as bus-route relationships to bus stops. These database relationships can be thought of as “virtual pathways” (Tomlinson, 2003) and allow for multimodal routing, where people, commodities, or information are transported by road, rail, ship, or airplane, transferring between each mode at known transfer points. Turns are modeled as features, allowing them to be conflated between different feature classes. Turns can also be modeled as “implicit” turning movements, taking into account turning angles.

Since the transportation network is a central element to so many GIS applications in Transportation, the underlying network representation is of primary importance. Networks are generally represented as a set of points and a set of lines that represents connections between those points. Although the points may be referred to as nodes or vertices, and the lines may be called arcs or links, the idea of connectivity is independent of the terminology used. The connectivity of a network is often referred to as its topology.

The main steps of the development of the data model are shown in the figure below.

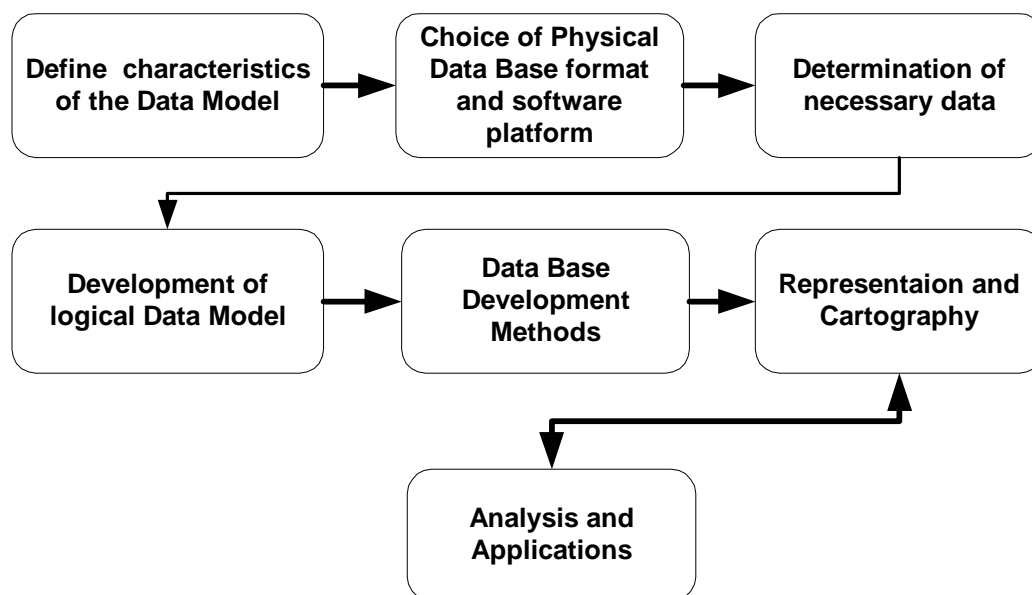


Figure 1. Steps for Development of the Network Data Model.

### Define characteristics of the Data Model

These characteristics include the scale of the data, the resolution, map projection and error tolerance. We have used different sources of information to acquire the existing data: existing digital data and paper maps. The scales vary from 1:5000 to 1:10 000 for city level information and 1:200 000 for national level of information. All data is defined in a common coordinate system: WGS 1984 UTM Zone 35 N. The resolution and the error tolerance are specified differently according to the scale and types of features.

### Choice of Physical Data Base format and software platform

For the data storage format we have choose the ESRI geodatabase, due to the rich functionality for storing network elements and features. The geodatabase is a logical extension of the object-relational databases that have been long accepted in GIS applications. Geodatabases can store feature classes and tables. Feature classes can exist as stand-alone entities with a geodatabase or can be grouped together into a feature data set. Feature classes store information that has a spatial representation. Tables can contain attributes that are related to a feature class but do not have an explicit spatial representation of their own. The geodatabase structure also allows relationships between feature classes and tables. Relationships are kind of behaviour that can be defined for a feature class within the geodatabase structure. Additional behaviours can be defined such as validation rules, connectivity rules, subtypes and domains.

The software platform used is ArcGIS 9.1 and the additional extension Network Analyst 9.1.

## Determination of necessary data

Generally speaking, this step includes the determination of what data is necessary and available for a particular application and then deciding how that data should be organized. Although this process is simply stated, it is not so easily implemented. To determine what data is needed, we must determine who will be using the application and under what conditions. An analysis of the existing data sets should be carried out with an eye toward what additional data will be required.

Transportation data has widely varying views and uses at different levels of geography. At the national level view transportation data is usually used in generalized forms. At the regional and municipal levels, transportation data is usually used by displaying highways and major roads while omitting details, such as local streets. The transportation maps use features from other categories, such as parks, water features and administrative boundaries. At the city and neighbourhood level view transportation data is displayed in detail including major highways, local streets, public transportation routes and additional cadastral data (parcels, buildings, etc.)

The proposed network data model is not to be considered as a generic transportation model. The accent is on the methods for the development of the most complicated part of the transportation data models – the network data set, which will serve for routing and analysis of multimodal transport systems. As a base for the model we have evaluated the existing data standards and models, as was mentioned in the previous section of this paper. We have used many ideas from these models to define and structure the required data. A valuable source of information was also the Preliminary project proposal of General City Plan of Plovdiv, which describes in basic terms the needed data to be acquired and analyzed and the needed information products to be derived from this data. Our network data model is orientated for the needs of the city of Plovdiv, but with additional development could be used for any other Urban Transportation System.

The elements of the network data model are divided into two main themes or groups:

- Elements, spatially related to the transportation network. The elements in this group are shown in the figure below:

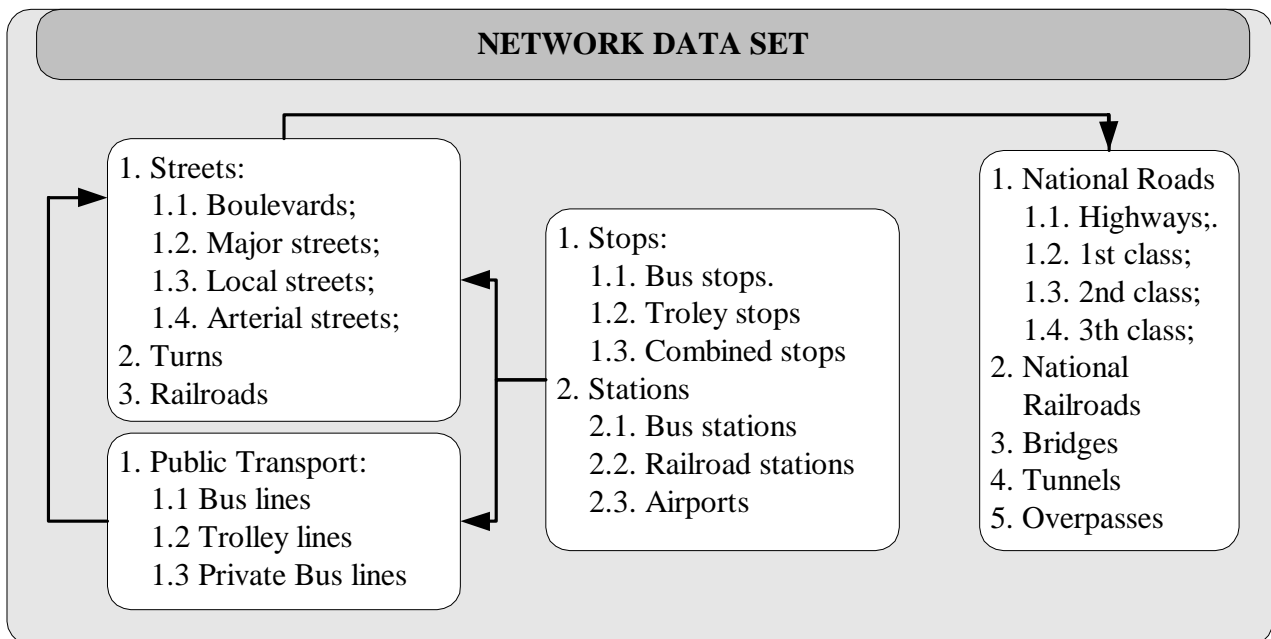


Figure 2: Network Data Model Set Elements.

The elements in this group include data for Streets, Railroads, Turns, Public Transportation lines, Public Transport stops, Stations for the city level. There are also additional data from a national transport network – Roads, Railroads, Bridges, Tunnels and Overpasses.

Additional elements could be added to this Network Data Model, but there are not present for the city of Plovdiv. Examples of these are: metro lines, tramway lines, bicycle lines, etc. Related to the network data model is the development of Linear referencing system for storing route-based information, based on dynamic segmentation. At this

time no use of such a system is determined, due to the lack of data. An example of this data is the information for the pavement conditions.

- Elements, logically related to the transportation network.

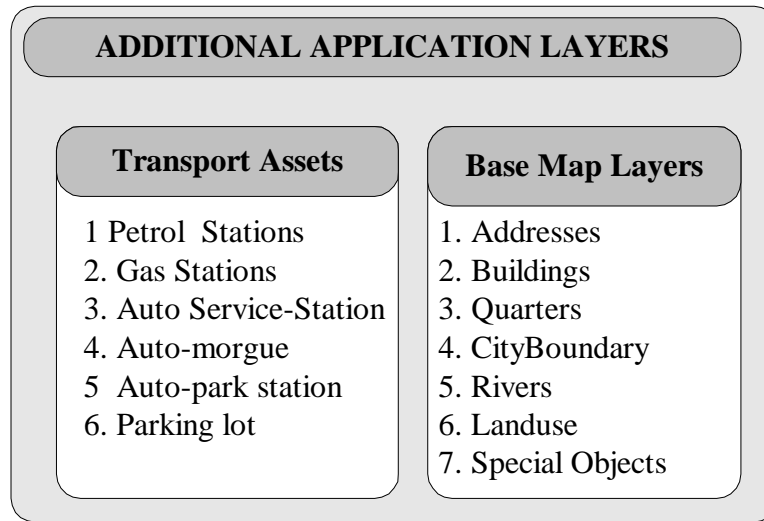


Figure 3: Additional elements of the Network Data Model.

The elements in this group include additional Transportation Assets, such as petrol and gas stations, auto-park stations, etc. These elements have logical connection with the network data set and could be used in additional analytical procedure. The second group of base map layers is used for visualization purposes only, but it's an obligatory base for extending the Network Data model in to more comprehensive Transportation Data Model.

Additional information in these groups could be added in order to implement this model as a generic transportation data model. The choice of these elements will vary accordingly with the specific application. Examples of these elements are: traffic signs, traffic lights, incidents, mobile objects, etc.

### Development of logical Data Model

After the data needs assessment has been completed, the physical design of the data can begin. The data that has been deemed necessary can be described and grouped by function or type. This may create themes in much the same way the geographic layers can be grouped into themes. These themes can be arranged into an Analysis Diagram that exposes the relationships between features and objects.

Networks used by ArcGIS Geodatabase are stored as network datasets. The physical implementation of a network is made of junction, edge, and turn elements, derived from the geodatabase feature classes and object relationships. Edges are elements that connect to other elements (junctions) and are the links over which resources flow. Junctions connect edges and facilitate navigation from one edge to another. Turn elements record information about movement between two or more edges. All these elements incorporate an advanced connectivity model that can represent complex scenarios, such as multimodal transportation networks. It also possesses a rich network attribute model that helps model impedances, restrictions, and hierarchy for the network. Elements can have any number of attributes, such as travel time for cars, buses, emergency vehicles, heavy trucks, cost, restrictions, slope, number of lanes, etc. These attributes can be calculated from field values found in geodatabase tables or from a script.

The main linear transportation features, identified for the city of Plovdiv are streets and railroads. Turns linear feature class is also required element for the network analysis. Turns can be made at any junction where edges connect. There are  $n^2$  possible turns at every network junction, where  $n$  is the number of edges connected at that junction. Even at a junction with a single edge, it is possible to make one U-turn. The Streets feature class is divided by a subtype in four classes: Boulevards, Major streets, Local streets and Arterial streets. The Railroad feature class has no subtypes. Based on the Streets feature class is build the Public transport feature class, which is divided in three subtypes: Bus Lines, Trolley lines and Private Bus lines. Another group of linear features are the National Roads and National Railroads feature classes. They represent the basic transportation infrastructure outside the city limit. The intersections of the city transport infrastructure with the national transport infrastructure are important to be monitored by the city authorities.

The main point feature classes are: Stops (Bus Stops, Trolley Stops and Combined Stops); Stations (Bus Stations, Railroad stations and Airports). Bridges, Tunnels and Overpasses are represented as point feature classes due to the fact that they belong to the National Transportation Network and are derived from scale 1: 200 000. Each of this feature classes have a rich set of attributes to allow different types of analysis. An example for the attributes calculations are the applied Cost attribute to the Streets feature class for calculation of travel time of a certain edge, measured in minutes. The Streets has also additional attributes set as restrictions: one-way streets and pedestrian streets. For a specific application additional attributes could be defined.

To model this diverse set of features we have used connectivity rules, which define the relationships between features. Connectivity is important in order to travel over the network. Network elements, such as edges (lines) and junctions (points), must be interconnected to allow navigation over the network. Additionally, these elements have properties that control navigation on the network. Connectivity in a network dataset is based on geometric coincidences of line endpoints, line vertices, and points and applying connectivity rules that we set as properties of the network dataset. Connectivity begins with the definition of connectivity groups. Each edge source is assigned to exactly one connectivity group and each junction source can be assigned to one or more connectivity groups. Junctions that are assigned to two or more connectivity groups are the only way that edges in different connectivity groups can connect. Connectivity groups are used to model multimodal transportation systems.

There are 30 connectivity rules for the whole network model. For example the point feature classes Stops must be spatially coincident with the Streets feature class and also each Stops subtype should connect to the specific subtype from the Public Transport feature class. Most of the point feature classes participate in more than one connectivity group.

The figure bellow gives an overview of the connectivity policy for some of the network elements.

Source	Connectivity Policy	1	2	3	4	5	6
BG_railroads	Any Vertex	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
BG_Roads	Any Vertex	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
CityStreet	Any Vertex	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Public_Transport : Автобуси	End Point	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Public_Transport : Маршру...	End Point	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Public_Transport : Тролей...	End Point	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Bridges	Honor	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Overpasses	Honor	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Stations : Автогара	Override	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Stations : Аерогара	Override	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Figure 4. Connectivity Rules.

Additional layers of information from Transport Assets and Base Map thematic groups are defined as a feature classes within the same geodatabase. These layers of information are defined as point, line and polygon classes and do not participate in the network. Many of these layers are used for spatial and attribute queries, network analysis and complex process models.

## Data Base Development Methods

After the definition of the database schema the next step is to populate the model with data. This step is typically accomplished using one or more of the following methods:

- Digitizing
- Scanning
- Keyboard input
- File input
- File transfer

For the different types of data within the database we have used all of these methods.

## Representation and Cartography

Transportation networks can be stored and displayed in a variety of ways. Commonly, road and rail networks are represented by single centerlines. This line symbol can theoretically be infinitely thin yet still represent the entire navigable way. In a multimodal transport system there are a great number of feature types and combinations of those to be presented. The presentation itself depends of the type of application and the function required by the needed information product.

The figure shows an overview of multiple line representations of the Public Transport System and the Streets feature class.

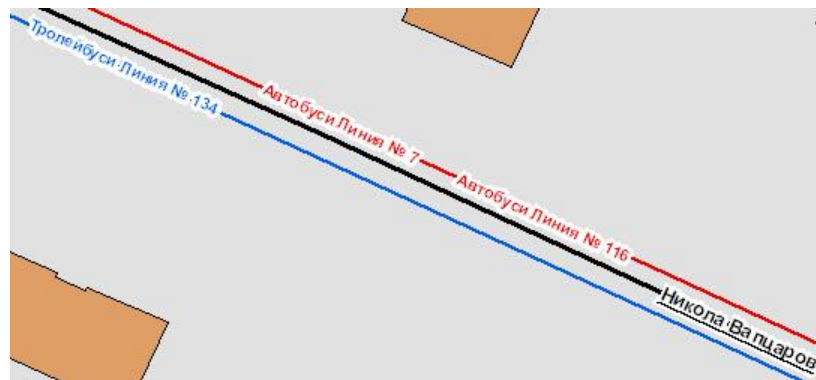


Figure 5. Multiple line representations

## Analysis and Applications

The true value of the network structure is that it allows a flexible and extended query capability for the production of information products, based on the network. These capabilities include:

- Shortest path. Find the least-cost path (time, distance, etc.) through a series of stops. (Stops are network point locations).
- Closest facility. Given one location, find the closest (time, distance, etc.) facility (ATM, hospital, etc.).
- Traveling salesman. First find the optimal sequence to visit set of stops, and then run a shortest path.
- Allocate. Build a shortest-path tree to define a “service area” or a “space-time defined space”.
- Origin/destination matrix (OD Matrix). Build a matrix of cost between a set of origins (O) and destinations (D). OD Matrixes are used extensively in many network analyses (such as Tour, Location/Allocation, etc.).
- Vehicle routing. The core route optimizer. Given a fleet of heterogeneous vehicles, in terms of capacity, available hours of service, cost to deploy, overtime cost, etc., and a set of heterogeneous customers, both vehicles and customers with time window constraints, find the optimal vehicle and customer assignment and route for the vehicle.
- Location/allocation. Simultaneously locate facilities and assign (allocate) demand to the facilities.
- “Chinese postman”. Find the optimal path through a set of connected edges. Garbage trucks, newspaper delivery.

The developed network data model could be used for all of the above mentioned function, which are generic for a lot of applications, related to the transportation: Bus Route Development, Emergency Dispatch and Route Planning, Evacuation Planning, Fleet Management, Hazardous Waste Transport, Traffic Management Package and Service, Delivery Sand and Salt Spreading, Snow Plowing and Removal Sanitation Services, Toll Collection, etc.

The figure below shows some basic routing functionality, based on all properties defined in the network data model (turns, restriction, etc.). The points are randomly selected and the optimal route is generated automatically.

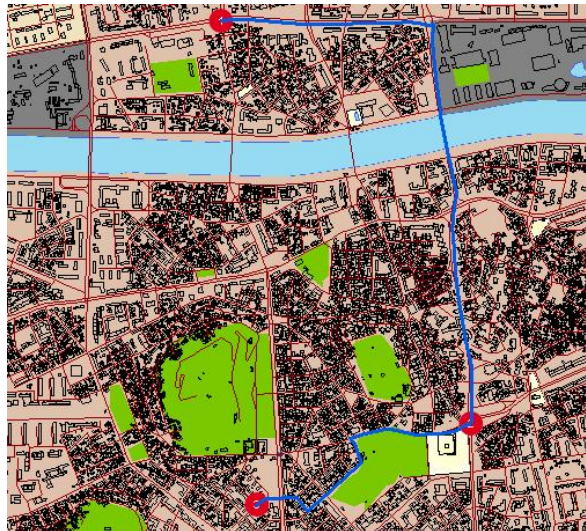


Figure 6. Routing functionality of the network model

A general map of the city of Plovdiv is provided on the figure below. The map includes the city streets classification, the national road data and the intersection points between two layers.

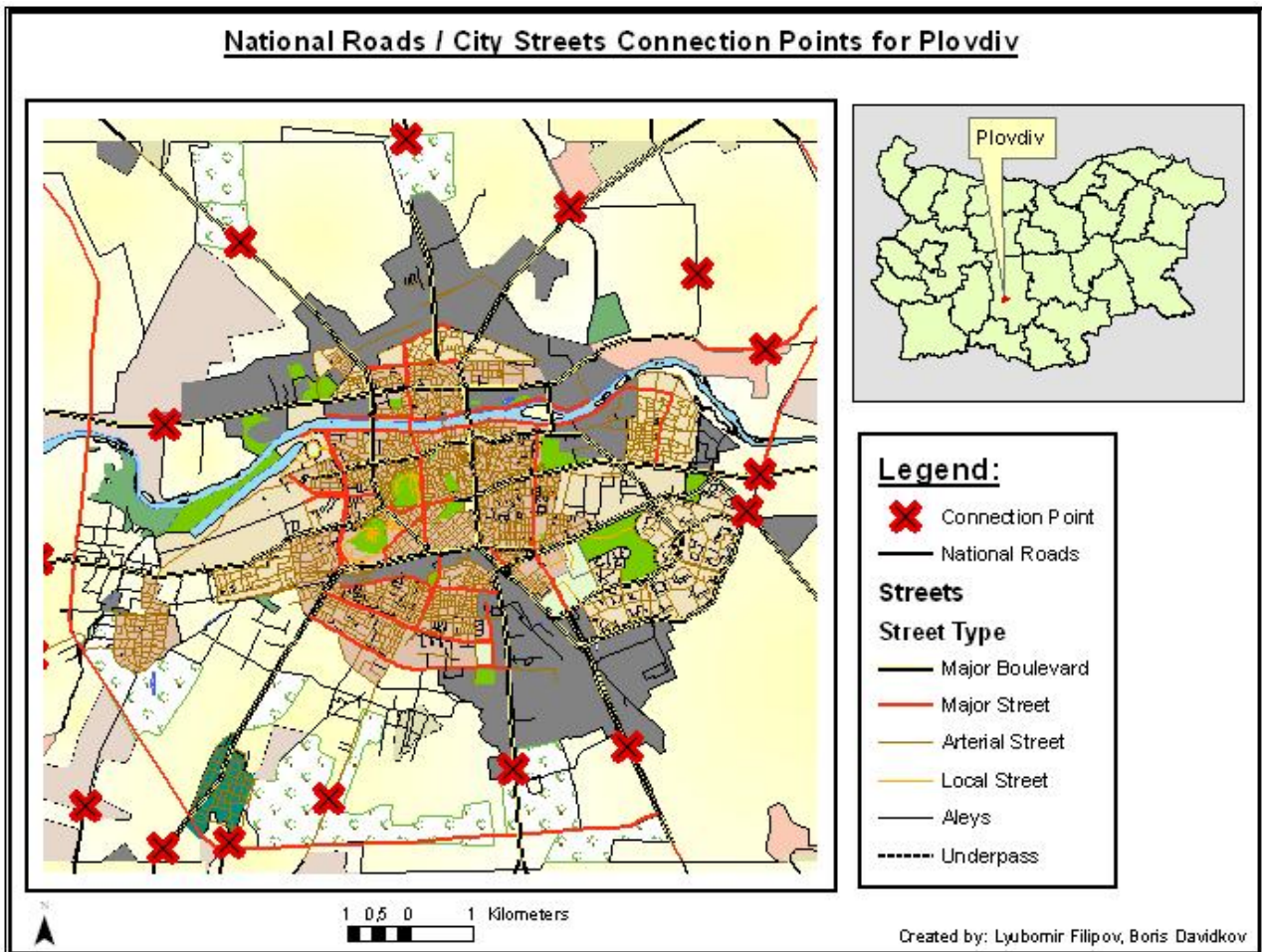


Figure 7. Map of Plovdiv.

The connection points of the city street network with the national road network are constantly monitored by the city authorities in order to collect data (traffic volumes, environmental parameters) and to analyze the potential problems of external traffic flows to the transport environment.

## CONCLUSIONS

The proposed data model is orientated towards routing and analysis of multimodal networks. There are some disadvantages of the model, related to the lack of transport data in Bulgaria – pour data for city level traffic volumes, environmental parameters, demography, etc. Apart from that the model is a good base for starting a comprehensive transport data model, which will cover the needs of transport planning and management of the city authorities. There are a lot of application areas for this model, mentioned in the previous section, which are not only limited to the competences of the city authorities, but to all private and public companies, whose activities involve some kind of transportation data or analysis. The difference and the great advantage of the model, from the traditional models for routing and navigation, are the rich functionality and capability to model complex transport features such as turns, restriction and connectivity rules.

## References

1. Israelsen, T., and R. Frederiksen, 2005. The Use of GIS in Transport Modeling. In *GIS, Spatial Analysis, and Modeling*, ed. D. J. Maguire, M. Batty, and M. F. Goodchild, 265–87, Redlands, Calif.: ESRI Press.
2. Tomlinson, R., 2003. *Thinking About GIS – Geographic Information Systems Planning for Managers*. Redlands, Calif.: ESRI Press.
3. Zeiler, M., 1999. *Modeling our World – The ESRI Guide to Geodatabase Design*. Redlands, Calif.: ESRI Press.
4. Rodrigue, J.P., 2005. Network Data Models. In *Transport Geography on the Web*, <http://people.hofstra.edu/geotrans/eng/ch2en/meth2en/ch2m3en.html> , Accessed on 28.11.2005.
5. Intelligent transport systems — Geographic Data Files (GDF) — Overall data specification, <http://www.4d-technologies.com/writegdf/> , Accessed on 28.11.2005.
6. Interoperability Clearinghouse Glossary of Terms, <http://www.ichnet.org/glossary.htm> , Accessed on 28.11.2005.
7. Preliminary project of General City Plan of Plovdiv, <http://www.plovdiv.bg/Obiava/Prilozhenie1%20i%20202.htm> , Accessed on 28.11.2005.
8. UNETRANS Data Model, <http://www.ncgia.ucsb.edu/vital/unetrans/> , Accessed on 28.11.2005.

## Biography of authors

Lyubomir Filipov



Lyubomir Filipov is a Phd student in Sofia University “St. Kliment Ohridski”, Faculty of Geology and Geography, Department of Cartography and GIS. He holds a BSc in Geography and MSc in Cartography and Geographic Information Systems. His major research interests are in the field of GIS and Cartography, Spatial Modeling, Urban and Regional Development, Transportation Studies. His Phd thesis is concentrated on the development and analysis of spatial models for Urban Transportation Systems in GIS environment for the city of Plovdiv.

Boris Davidkov



Assoc. Prof., Phd. Boris Davidkov is a lecturer in Sofia University “St. Kliment Ohridski”, Faculty of Geology and Geography, Department of Cartography and GIS. He works in the field of thematic cartography and reads lectures on

several disciplines: General Cartography, Thematic Cartography, Touristy Cartography, Socio-Economic Cartography, History of Cartography and others on a bachelor and masters levels. He has a number of publications in these areas and practical experience in applied cartography (a Phd theses on Complex mapping for Pazardjik region – 1993).