Data Models for Moving Objects in Road Networks - Implementation and Experiences

Milos Vojinovic*, Zeljko Cvijetinovic**, Dragan Mihajlovic**, Momir Mitrovic**

* Geomatics Company MapSoft, Belgrade, Serbia
** University of Belgrade, Faculty of Civil Engineering, Department of Geodesy and Geoinformatics, Belgrade, Serbia

Abstract. Paper deals with the specific LBS scenario – Fleet management (FM) and more specifically with systems for Automatic vehicle location (AVL). Well designed and implemented spatial data model for moving objects is one of the most significant elements of any AVL system. In practical applications the results of the latest scientific research are seldom applied, despite the fact that this area has been developing intensively for more than 20 years. The reasons for this are analysed in the paper. Short analysis of functionality of these systems is presented considering the impact of these functionalities on the implemented data model for moving objects and more specifically their impact on spatio-temporal component of the model. The paper especially reviews the possibility of using road networks as a basis for the representation of moving objects data models and a fact that these models are rarely used in practical applications. A solution overcoming this situation is proposed. The solution assumes transition from the system that is not based on road network to the system that is based on network. There are quite few research papers dealing with OSM data models. Therefore, a significant space in this paper is dedicated to the description of these models since OSM data can be valuable for this type of applications.

Keywords. Moving objects data model, road networks, fleet management, vehicle tracking

1. Introduction

One of the most significant elements of any Automatic vehicle location (AVL) system is a well conceived data model for moving objects. Research

Published in "Proceedings of the 11th International Symposium on Location-Based Services", edited by Georg Gartner and Haosheng Huang, LBS 2014, 26–28 November 2014, Vienna, Austria.

on these models started more than two decades ago. However, results of the research are seldom fully implemented in practical applications. A basic reason for this is a fact that it is quite possible to develop and implement FM AVL that is applicable in many practical application scenarios simply by using standard RDBMS tables without spatial extensions, standard SQL language and services available from industry content providers.

This and other reasons for this situation are briefly analysed in the paper. Short analysis of AVL systems in terms of their functionality requirements is given in the paper. Implications that these requirements have on data models and spatio-temporal analysis aspect of using these models are especially analysed.

Another very important aspect for the specific FM AVL application is availability of the high quality road network data. There are quite few research papers dealing with OSM data models. Therefore, a significant space in this paper is dedicated to the description of these models. Classification of these models according to their application domain has been introduced and some typical models have been described. Quality of the data is also analysed.

Special consideration is given to the aspect of using road network as a basis for implementing data model for moving objects and the fact that these models are rarely used in practical applications. Paper reviews reasons, i.e. objective circumstances that influenced such a situation. All the circumstances were analysed, starting from current models of GPS/GPRS devices for vehicle positioning and data acquisition, through telecommunication networks status and possibilities of applying international and industrial standards in this area.

A solution overcoming this situation is proposed. The solution assumes transition from the system that is not based on road network to the system that is based on such a network.

2. FM AVL Scenario

FM AVL is one of the typical LBS scenarios. It is a mixture of management and tracking scenarios (*Figure 1*).

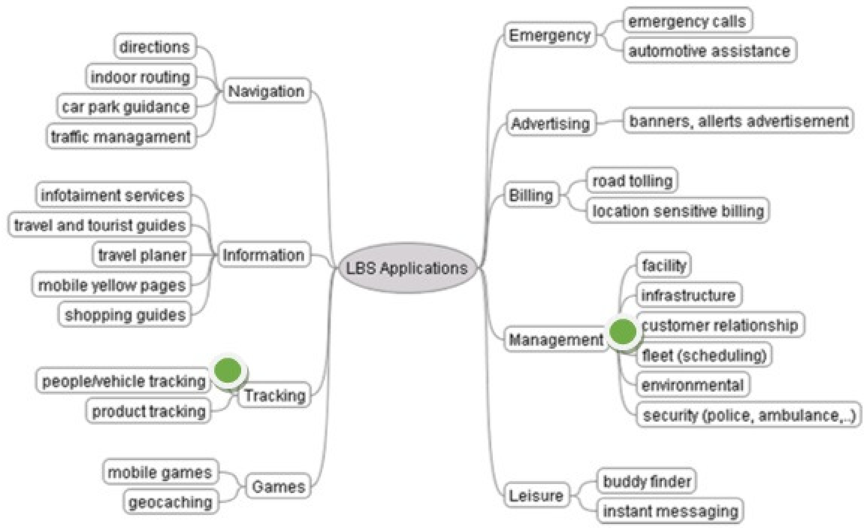


Figure 1. LBS application categories / Foundations of LBS (Steiniger et al. 2006)

Specifics of this scenario can be summarized through LBS components used by the system (*Figure 2*):

Communication network – GSM network is used predominantly; network is well developed with high stability in data transfer; data transfer rates are sufficient to enable efficient transfer of all the data (except for maybe video files in real time); very affordable prices of data transfer enabled that large data quantities can be transmitted in real time; new EU legislations in the area of telecommunications provided that state borders are rendered as irrelevant in terms of data flow through telecommunication networks, so it is possible to implement vehicle tracking in real time at very affordable pricing;

Positioning – GNSS (GPS and GLONASS positioning); very high quality of positioning is provided; in the last year large number of companies, in addition to GPS, enabled utilization of GLONASS satellites; this provided higher data accuracy and reliability but also better availability of positioning devices; prices of positioning devices are quite affordable, so it is possible to purchase devices with basic functionalities for about 50 EUR, devices with medium functionalities for about 80 EUR and devices that can be used for implementation of complicated application scenarios can be purchased for about 100 EUR;

Mobile devices - GNSS/data devices; in the last year a large number of devices supporting GLONASS positioning emerged, providing increase in quality, accuracy and reliability of positioning; manufacturers provided modules enabling data readout from internal vehicle buses (CAN, LV-CAN, K-line and others); therefore, number and quantity of telemetric data are signifficantly increased.

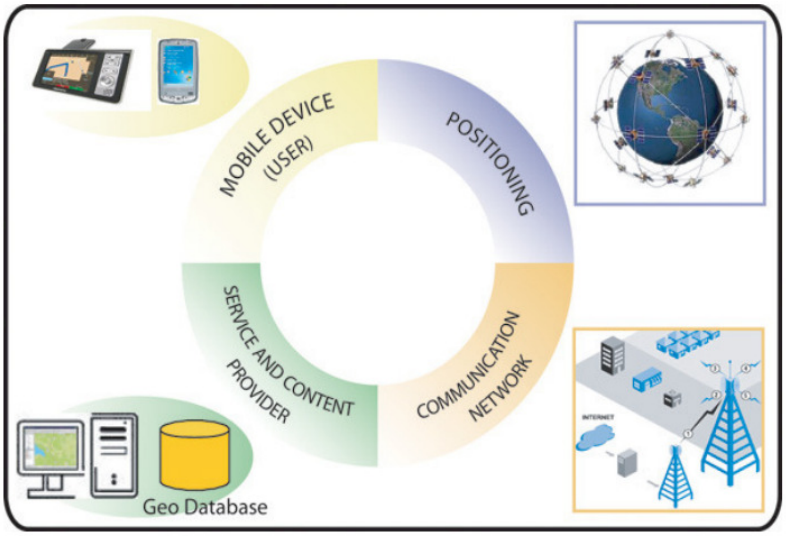


Figure 2. The basic components of LBS (Steiniger et al. 2006)

Service Provider and Content Provider – there are global data providers (Google, Microsoft), data providers offering data content highly relevant for FM scenarios (for example, Teleatlas offers navigation and traffic data) and Open Source providers (OSM data); global, regional and local leaders have been distinguished on the market (Insight 2012); customers' demands and the market shaped typical services offered by providers; therefore, almost all providers offer few standardized service packages.

Specifics of this scenario can be described through the following aspects: types of data that are collected, functionalities that are required and users' requirements regarding temporal aspects of the vehicle movement.

1. Types of data that are collected:

- Geospatial data (positions acquired by using GNSS method),
- Data based on GNSS measurements (speed, acceleration, azimuth, etc.),
- Telemetric data (fuel consumption and fuel level).

In addition to basic data, a large number of telemetric data containing additional information on vehicle is collected (fuel consumption, temperature, number of engine rotations per minute, door opening control, alarm, etc.).

Therefore, in addition to location tracking, the whole series of additional parameters is the subject of interest to users.

Besides online tracking, users also demand for reports based on collected parameters – driving style (harsh braking, harsh acceleration and harsh cornering), data on driver (driver's id, sending working orders to driver, communication with driver, etc.).

2. Functionalities that are required:

- Vehicle location display,
- Display of the vehicle track,
- Reports,
- Alarms,
- Route optimization.
- etc.

3. Users' requirements regarding temporal aspects are the following:

- Current vehicle location and status of all the parameters that are tracked,
- Vehicle locastions and status of all the parameters that are tracked in the past,
- Prediction of the vehicle location and status of all the parameters that are tracked in the future.

Official status of the collected data is very important aspect that has to be considered for these systems. This is especially related to the following data:

- **Position and time** - positional data are obtained with CEP=3m accuracy and time accuracy is expressed as RMS=1sec; this means that objects that have been tracked can be located with accuracy that is satisfying for most users; users also have demands regarding reports based on these source data;
- **Speed** - GNSS method is currently one of the best and one of the most common methods for vehicle speed measurement; speed is determined with RMS=1km/h accuracy; users also usually have demands regarding reports based on these data;

- **Mileage -** data on mileage are obtained using the following ways: from the vehicle, by calculating it from vehicle locations, and from odometer of the GNSS/GPRS device; accuracy of the mileage obtained from the vehicle is usually very unsatisfying because it depends on many external factors that are related to vehicle condition; accuracy of the mileage that is calculated from vehicle locations depends largely on the data resolution; accuracy of the mileage that is obtained from odometer is very good and it is better than 1% of the mileage, which usually satisfies users' requirements.

3. Models for Moving Objects

Beginning of the research on data models for moving objects (mobile objects is also an expression that can be used with the same meaning) is continuation of the research efforts from late eighties and early nineties of the last century (Tanse et al. 1993). Theories related to the spatio-temporal and geospatial data bases from the middle of nineties gave significant contribution to the development of these models (Sistla et al. 1997). Since 2000, international and industrial standards in GIS and LBS, i.e. OGC / ISO TC 2011 have very important influence on the development of these models (Stojanovic & Djordjevic-Kajan 2003, Macedo et al. 2008).

Characteristics that can be used to describe these models are the following: application of object oriented data models, availability of past, current or/and future object's positions and its trajectory, mobile geometry data types are supported or not, possibility of performing spatio-temporal operations and availability of the query language that has been developed for implementation of these spatio-temporal operations.

According to the basic concept these models can be regarded as:

- Location management,
- Spatio-temporal databases.

This classification actually corresponds to the classification that can be made in yet another way – according to the way that we want to track our object through time:

- Modelling of object's movement in present time and eventually in near future,
- Modelling of object's movement as it was in the past.

The first one is focused on providing answers to questions regarding actual positions of moving objects, and possibly regarding their predicted evolutions in near future.

The second approach takes into consideration a complete history of moving objects' movements (Cotelo et al. 2003, Güting et al. 2006). This is definitely more challenging requirement. A complete evolution of the moving object can be given as an attribute in a database.

According to the type of movement of moving objects the models can be classified as:

- Models based on free movement (moving objects move freely in space),
- Models based on movement in network (moving objects move on network).

In addition to models of moving objects, algorithms and languages supporting database management (queries) with spatio-temporal information have been subject of the research (Wolfson et al. 1998, Vazirgiannis & Wolfson 2001, Cotelo et al. 2001.).

3.1. Models for Moving Objects Moving Freely in Space

Even though vehicle moves continuously through predefined network infrastructure (mainly roads), this fact is usually disregarded. During the late nineties and early years of the twenty first century researches were especially focused on the problem of modelling and indexing moving objects (Wolfson et al. 1999, Saltenis et al. 2000, Porkaew et al. 2001, Güting et al. 2000.). Common characteristic of these researches is that they were based on the free movement of objects (moving objects moving freely in space).

These models are frequently used in practical applications. The reason for this is that implementation and use of the models that are based on objects movement on network are more complicated, but also due to the fact that the data required for the network model implementation are usually unavailable.

3.2. Models for Moving Objects on Network

It can be stated that research foundations in this area has been set by Güting (Güting et al. 2006.). He defined networks taking into consideration the importance of the routes (routes correspond to roads from real life and they can be represented as paths in a graph) from the following reasons:

- Roads are relevant conceptual entities from the real life; roads have names instead of crossroads or road segments between crossroads; address systems are also linked to roads or streets; model should reflect these facts;
- Using lines to describe paths, i.e. using them as means for referencing is in widespread use as a concept applied in GIS; this also indicates that

positions should be described in respect to roads and not in respect to graph edges;

- Maybe the most practical reason for this is that representation of the moving object is as small as possible; for example, if vehicle goes along the highway using a constant velocity, if its position is to be described by an edge, than a change of that description should be done on every exit/crossroad, because identification of edges is also changed; however if a position is linked to a road, than description would have to be changed only if the vehicle changes a highway.

Güting defined network as a set of routes accompanied with a set of crossroads between routes.

Using these assumptions as a foundation he built a model based on the following principles (Güting et al. 2006.):

- Definitions of the spatial network, network locations and network regions are given; model offers paths through network of graphs as a basic concept called route; route enables to make a difference between simple and dual routes, as well as to describe possible transitions of object movement (for example vehicles on crossroads);
- Networks are available through network data type; interface for relational environment is made enabling creation of network from relations, but also creation of relations from the network;
- Two data types are created gpoint and gline; gpoint type represents a position in given network gline type represents a region in the network; these data types are integrated within a system that has been defined earlier (Güting et al. 2000.), so that time-dependant constructors moving (gpoint) and moving (gline) can be applied on them; these four data types makes the core of the concept;
- Operations are also defined in order to enable interaction between gpoint and gline objects, on one side, and the network and spatial data, on the other side; some of the specific operations are defined for moving (gpoint) and moving (gline) data types, as well as for the network data type (for example, shortest path and trip).

Comparative analysis of moving objects models is given in the *Table 1*.

Source/model	Object based	Trajectory	Mobile geometries	Mobile data types	Movement on network	Spatio-temporal operations	Query language
MOST	-	Current, future	Point	No	No	No	FTL
Güting, 2000	Yes	Last	Point, region	mreal, mboolean	No	Yes	SQL extension
Brinkohoff	-	Last	Point	No	Yes	No	No
Mouza	-	Current	Point	No	Yes	No	No
Vaziriannis	-	Current, future	Point	No	No	Yes	SQL extension
Stojanović	Yes	Last, future	Point, polyline, polygon	Mobile boolean, number, string, object	Yes	Yes	SQL extension
Güting 2006	Yes	Current, future	Point, region	Mreal	Yes	Yes	SQL extension

Table 1. Models for moving objects (modified from Stojanović 2004)

4. OSM Data Model

In the last few years quality of network data has been improved significantly. Various global trends contributed to this. Maps and spatial data have become one of the most significant sources of information and both Google and Microsoft showed enormous interest for such data. Open Source community also has a great influence. Quantity of information is increased on a daily basis and there is a large number of online services offering such data.

OSM data model (Wiki OpenStreetMap 2014) is a powerful way of presenting geographic information. OSM uses topological data structure where elements (primitives) are basic components of the data content. Elements that model the physical world are the following:

- Nodes – defined points in space,
- Ways – defined linear objects and borders of areas,
- Relations – explain how elements are related to each other.

Each of these elements can have one or more tags describing the meaning of the element.

Node represents a point on the Earth surface defined by its coordinates, latitude and longitude (in WGS 84 coordinate system). Each point at least contains ID number and a pair of coordinates. Nodes are used for the definition of independent point features (for example, bench in the park, drinking-fountain, traffic light, mountain peak, etc.), but it can be also used for the definition of the shape of ways.

Way is ordered list containing 2 or more points (up to 2000) defining a polyline. They are used for the representation of linear objects such as: rivers, roads, streets, etc. They can be used to represent polygon (area) borders such as: buildings, forests, lakes, etc.

Relation is ordered structure describing link between two or more elements (nodes, ways, relations). For example, relation for turn restriction states that it is not allowed to turn from one way to another way. Relation contains a list of elements and they are calld relation's members. Each relation's member can optionally has the role in relation. For example, for turn restriction relation members can have roles *from* and *to* which describe from which way it is forbidden to turn to which way.

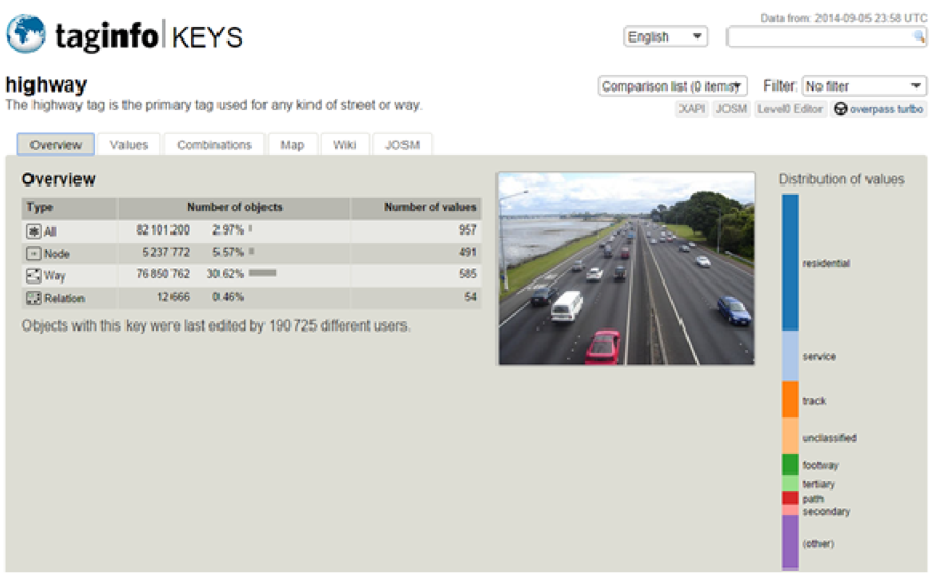


Figure 3. Frequency of using tags (OpenStreetMap Tags, Web links)

Each of elements can have tags describing the meaning of that element, i.e. they represent elements' metadata. Tags are comprised of two text fields, key and value and each of them consists of 255 characters. For example, *highway=pedestrian* denotes pedestrian path. There is no fixed list of tags and a user can create his own tags. However, there is a convention to use tags that are already defined. Also, there is a web page http://taginfo.openstreetmap.org/ (*Figure 3*) where it can be seen for any tag how often it is used and with what value, in which element, and so on.

There are also some common attributes that are stored for nodes, ways and relations. These attributes are listed in *Table 1*.

Name	Value type	Description
id	integer	Unique id of element for certain type (for example, node and way can have the same id)
user	string	Name of the user who was the last one to make a change on the element. User can change his name.
uid	integer	Unique id of the user who was the last one to make a change on the element. This id is always the same for certain user.
timestamp	W3C Date and Time Formats (YYYY-MM-DDThh:mm:ssZ)	Time of the last update of the element.
visible	Boolean ("true", "false")	Indicates whether element is deleted or not in OSM database
version	integer	Element's version. Newly created element has value of 1 and this value incrementaly increases.
changeset	integer	Set of changes that containing element update.

Table 2. Common attributes for all OSM elements (OpenStreetMap Tags 2014)

4.1. OSM Tags Important for Road Network

Tags related to the road network are only a fraction of the complete list of tags. According to application *osm2pgrouting* which is used for OSM data import into database, for import of ways representing road network only four tags are used:

- Highway,
- Cycleway,
- Tracktype,
- Junction.

These four tags can have a lot of different values. However, only key-value pairs from *Table 2* (only one tag specific for certain grup is listed) are used for road networks.

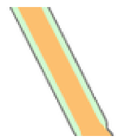
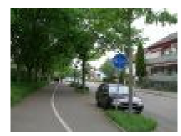
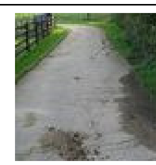
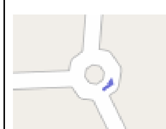
Key	Value	Element	Description	Map display	Photo
highway	motorway		A restricted access major divided highway, normally with 2 or more running lanes plus emergency hard shoulder. Equivalent to the Freeway, Autobahn, etc...		
cycleway	lane		A lane is a route that lies within the roadway.		
tracktype	grade1		Solid. Usually a paved or heavily compacted hardcore surface.		
junction	roundabout		Roundabout. This automatically implies oneway=yes, the oneway direction is defined by the sequential ordering of nodes within the Way.		

Table 3. Some of OSM tags important for road networks (Wiki OpenStreetMap 2014)

4.2. OSM Data Providers

There are many providers offering OSM data files in various data formats. OSM files for certain territories can be downloaded from internet pages of these providers.

A single OSM file for the whole planet Earth (Planet.osm file) can be downloaded from the site: http://planet.openstreetmap.org/ in .pbf and .osm.bz2 format. This site also offers a possibility to download files with daily, hourly and minute changes (.osc). Other OSM data providers are listed in *Table 4*.

WEB address	Update
ftp://ftp.spline.de/pub/openstreetmap/	Weekly
http://ftp.osuosl.org/pub/openstreetmap/	Weekly
http://ftp.snt.utwente.nl/pub/misc/openstreetmap/	Weekly
http://ftp.heanet.ie/mirrors/openstreetmap.org/	Weekly with 10 hour delay
http://ftp5.gwdg.de/pub/misc/openstreetmap/planet.openstreetmap.org/	Daily
ftp://ftp.pucpr.br/osm/mirror/	Daily
http://download.bbbike.org/osm/planet/	Weekly

Table 4. OSM data providers and their data update rate

4.3. Road Network Data Models Based on OSM Data

OSM data can be stored in PostgreSQL, i.e. PostGIS database so that they can be used for moving object modelling and routing. There are dozens of applications that have been developed to facilitate storage of PSM data into PostGIS database. Therefore, there is approximately the same number of data models used for storing OSM data within PostGIS. Some of these applications were developed exactly to meet the requirements of routing, so they build a data model that is suitable for that purpose, whereas other applications aim at displaying OSM maps, facilitating OSM data maintenance, etc.

Applications that are most frequently used are the following:

- *osm2pgrouting*,
- *osm2po*,
- *osm2pgsql*,
- *osmosis*.

Depending on what kind of data models these applications build, they can be classified into following groups:

1. Those that are appropriate for routing (*osm2pgrouting* and *osm2po*),
2. Those that are appropriate for OSM map display (*osm2pgsql*),
3. Those that are appropriate for OSM data update (*osmosis*).

4.3.1. Models suitable for routing

The routing in this paper will be discussed in the context of using *pgRouting* library (open-source library facilitating routing for the data kept in PostGIS database). However, it can be extended easily to routing using some other tool as well, so whenever in this paper some model is evaluated as suitable or not for *pgRouting* library, it certainly means that the model is also suitable or not for routing in general.

The development of *pgRouting* library also initiated the development of applications for the storage of OSM data for routing into PostGIS database. Two open-source applications that are used for this purpose are *osm2pgrouting* and *osm2po*.

OSM data are kept within a database in the following data tables:

- *types* – road types (types of segments/ways);
- *classes* – classes (subtypes) of the road;
- *nodes* – all points from the OSM file (road network points and points outside of the network);
- *relations* – relations related to the road network;
- *relation_ways* –relations between ways/segments (within road network and outside of the network);
- *ways* – segments of the road network (from node to node);
- *way_tag* – tags of road network segments;
- *ways_vertices_pgr* – road network nodes that have been created by *pgRouting* function *pgr_createtopology*.

The following diagram (*Figure 3*) depicts all tables, their attributes and their relationships.

Attribute *osm_id* of the table *ways* represents id of the segment from OSM file and attribute *gid* represents a unique identifier of the segment. The reason for having *gid* attribute is that more than one segments can have the same *osm_id*, because one OSM segment can be divided into several segments connecting network nodes. If new data set is imported again for the same area, the same OSM segments would not get the same *gid*.

Good characteristic of this model is that the table *ways* is modelled in such way that *pgRouting* functions can be called and executed for this table and attributes *length* (representing cost) and *reverse_cost* have been created to respect one-way streets (if some segment is one-way street and if that segment is oriented in one-way street direction, its cost will be equal to the street length, whereas its *reverse_cost* will be 1000000 times larger). This table also contains maximum allowed speed in both directions (attributes *maxspeed_forward* and *maxspeed_backward*) which is very important if it is required to calculate time to travel through some segment or route, as well as street names (attribute *name*).

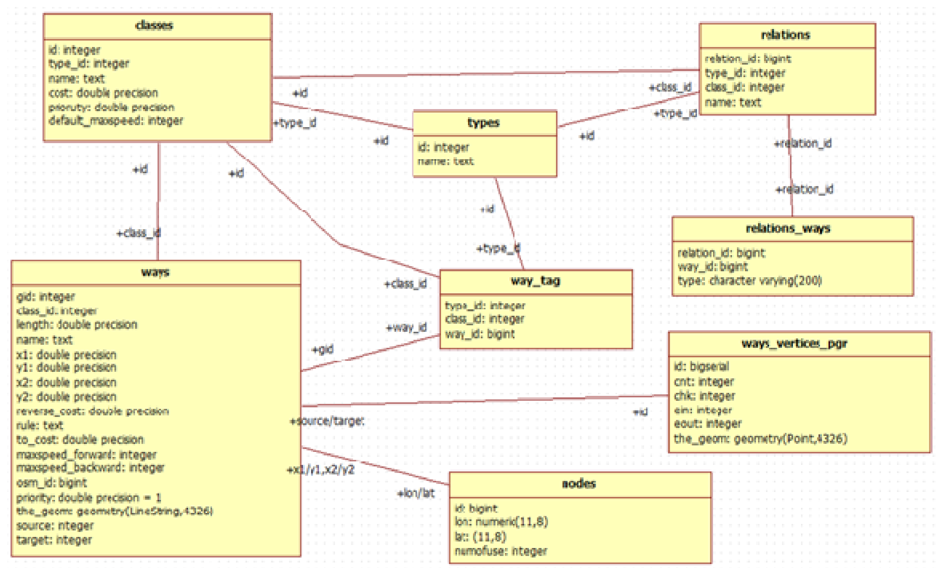


Figure 4. *pgRouting* data model

The raw OSM data contain much more data and information than it is needed for routing and that makes this tool very suitable for simple routing systems. However, there is a problem if the tool is used within routing systems where additional information such as street names, respecting of turns restrictions, etc. is needed. Currently, this model does not provide support for additional pieces of information on *nodes*, *ways* and relations that are stored as tags of OSM file. The table *way_tag* is probably designated for the segments. However it only contains type and class of the road. These pieces of information already exist in the table way.

Another big problem related to this application is that it does not support update of OSM data. If it is required to update the database containing OSM data, than new data set has to be imported again from the scratch and that will change *gid*s of the segments. However, this is not convenient if history of data changes is to be maintained.

4.3.2. Models Suitable for OSM Maps Rendering

As the title already suggests, models suitable for OSM maps rendering are developed to store OSM data in a way that would be the best for data rendering. The best known application from this group of applications is *osm2pgsql*.

The database created using this application and its model is primarily to be used for data rendering data by using Mapnik software. As such, it is not suitable for *pgRouting* functions. History of changes is also unavailable. Good side of this model is that most of the tags are stored in their own columns of the four basic tables.

Update is facilitated by the application called *osmosis*. This application reads OSC file and after that *osm2pgsql* application is used to store the data into the database.

4.3.3. Models Suitable for OSM Data Update

The best known application that builds these models is *osmosis*. In addition to update, this application provides the data model that facilitates history of OSM data changes maintenance. The model created by *osm2pgsql* application can be maintained by using the combination of applications *osmosis* and *osm2pgsql*. However, primarily function of this application is rendering of the OSM data.

Osmosis (Wiki OpenStreetMap 2014) is JAVA application for processing OSM data. This tool contains a lot of components that can be combined in order to perform more complicated operations such as: processing/reading of OSM file, database update, data sorting and indexing, etc. Some of the tasks that this application successfully performs are the following:

- creation of OSM file using the data from the database;
- OSM file data import into the database;
- creation of OSS file using history tables;
- database update;
- comparison of two OSM files and creation of OSS file;
- OSM file data sorting;
- reading of the data enclosed by the specified bounding box.

By using this application, OSM data can be imported into PostgreSQL/PostGIS database. This can be done by using one of three available data models (schema):

- Simple schema,
- Snapshot schema,
- OSM API schema.

Simple Schema, as its name suggests, is a simple schema that facilitates storage of OSM data. This is older schema type where tags for nodes, segments and relations are stored in separate tables.

It is possible to update OSM data, but this schema does not store history of data changes. The update is done by using the same code that is also used for the initial data import. The difference is that, instead of the parameter --*write-pgsimp*, a parameter --*write-pgsimp-change* is used. Also, OSS file is used, instead of using OSM file.

This model does not facilitate usage of *pgRouting* functions. Therefore, come changes are required in order to facilitate this usage.

Snapshot Schema is basically the same as a Simple Schema with additional improvement facilitating usage of *hstore* columns. This column enables that set of key/value pairs can be stored in a single field of an entity. This is very applicable for OSM data and elements' tags. This provide for memory savings.

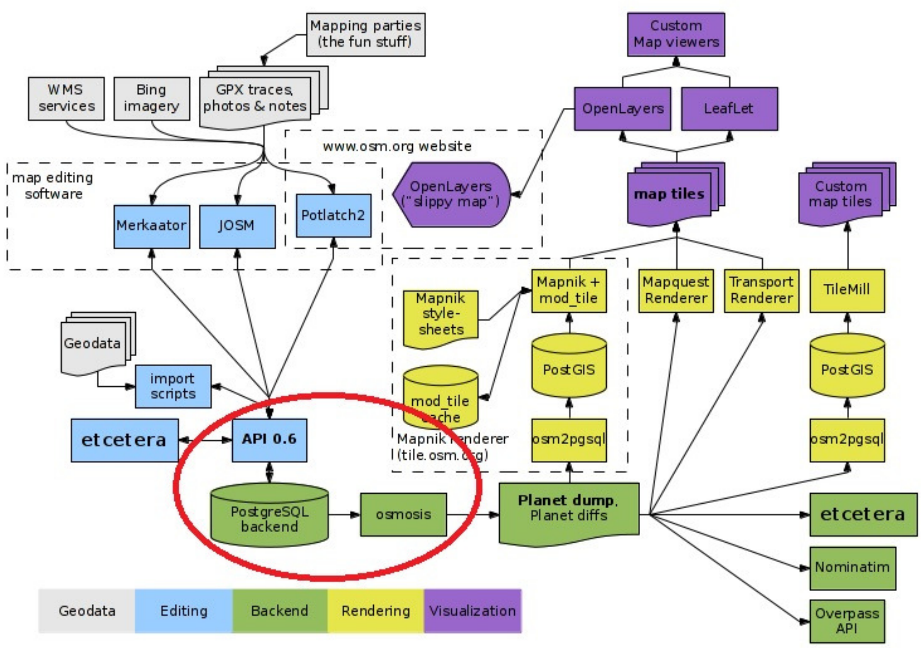


Figure 5. OSM model appropriate for data update (Wiki OpenStreetMap, 2014)

OSM API Schema represents a schema that is in the background of OSM API, i.e. it represents a complete OSM database. The purpose of this database is to answer to all OSM API requests and to render OSM maps.

OSM system model is depicted in *Figure 5*. It can be seen that this schema is used in the background and that *osmosis* is to be used for serving OSM and OSS files that are used by *osm2pgsql* application to create models for data rendering. This model contains 39 tables.

The data stored using this model can be updated in the same way as in the cases of using the previous two models. In addition to update, this model facilitates the maintenance of history of changes.

This model is also unsuitable for using *pgRouting* functions. However, this model is the most complete one when compared with all the previous models. It enables data update, history of changes maintenance. Also, all pieces of information from OSM file are stored within the database, so they are immediately available if they are needed for some reason. With minor modifications this model could be suitable for modelling movements of moving objects as well as for *pgRouting* functions and as such it would be the best solution for more complex systems for routing and moving objects movement modelling.

For moving objects movement modelling it is especially important that this model provides data versioning. This could be very important if history of vehicle movement is to be maintained.

4.4. OSM Data Quality

OpenStreetMap is open-source project enabling people all over world to collect data for their territory. Users use different data sources such as satellite images, terrain maps, GPS devices and so-called local knowledge to collect data that is precise and updated as much as possible.

In practical applications it very important that the data completely and without gaps covers the whole relevant territory and that the data is at the same time complete and consistent.

The same applies to routing applications. It is necessary that road network is complete and consistent. Otherwise, incorrect information is obtained and the routing itself loses its basic purpose – savings in time and resources. The modelling of moving objects movement is also very demanding task because insufficient accuracy of the data can undermine the whole concept and meaning of the model that is based on such data.

Since OSM data have been collected by interested people, based on their own initiative and from different data sources it is reasonable to ask a question regarding accuracy and suitability of OSM data for the modelling of moving objects and for the routing applications.

The methodology for quality assessment is perhaps best described by Oort (2006) and Hackley (2010). The methodology involves the following aspects, i.e. data quality elements: lineage, positional accuracy, attribute accuracy, logical consistency, completeness, semantic accuracy, usage, purpose and constraints, and temporal quality.

Description of these data quality elements and their relevance (grades from 1 to 5) for the modelling of moving objects and for the routing applications is given in the *Table 5*.

Aspect	Modelling	Routing
Lineage	1 - Less important aspect.	1 - Less important aspect.
Positional accuracy	5 - Very important aspect.	3 - Medium important aspect. It is good if positional accuracy is high, but topological consistency is more important than positional accuracy.
Attribute accuracy	3 - Important aspect. It is good if attributes are accurate, but it is not vital for the proposed model.	4 - Very important aspect, especially for the routing which takes into consideration various types restrictions.
Logical consistency	3 - Important aspect. It is good if attributes are logically consistent, but it is not vital for the proposed model.	5 - Very important aspect.
Completeness	5 - Very important aspect.	5 - Very important aspect.
Semantic accuracy	3 - Important aspect. It is good if semantic accuracy is high, but it is not vital for the proposed model.	4 - Very important aspect, especially for the routing which takes into consideration various types restrictions.
Usage, purpose and constraints	3 - Important aspect.	4 - Very important aspect, especially for the routing which takes into consideration various types restrictions.
Temporal quality	5 - Very important aspect.	5 - Very important aspect.

Table 5. Data quality elements and their relevance for the modelling of moving objects and for the routing applications

Haklay (2010) performed a spatial analysis of the quality of OSM street network representations in the UK. Girres and Touya (2010) performed a spatial analysis of the quality of OSM for France. Both authors use a comparison to ground-truth data obtained from the corresponding national mapping agency. Both studies have proved that the data quality is good to a great extent, but the quality is not uniform for the whole study area.

The research conducted by Neis et al. (2012) where the quality of the OSM street network was evaluated for the different regions of Germany, for the period 2007–2011 is also interesting. Neis found that, on the average, the OSM street network representation exceeds the information contained in a proprietary dataset by more than 20%.

Corcoran et al. (2013) analysed the growth of OpenStreetMap networks. They noticed that the complete road network for the USA was imported into OSM from the TIGER dataset in 2007 and 2008.

A statistical data were collected during this research showing how many new OSM elements have been added every day (*Table 6*). Statistical data have been produced using 5 updates by using OSS file (one file for each day) to update OSM data stored in API OSM data model.

	Serbia	Hamburg	Luxemburg	Slovenia
current_nodes	4255	531.6	413	3333
current_relations	0.4	2.6	0.4	2.6
current_ways	331.8	114.8	66.6	204.2

Table 6. Average number of elements added to OSM data every day for different territories

It can be noticed from the statistical data (*Table 6*) that the extent of changes (number of new nodes and ways) for the Serbia is the largest in comparison to the other areas. As far as the other areas are concerned, Slovenia has somewhat larger number of changes, whereas Hamburg (Germany) and Luxemburg already have well established road network and therefore it is reasonable that less new elements are added to OSM database on a daily basis for these territories.

The map has been published on OSM forum depicting the level of changes (increase in road network total length) during the period of 2009-2012 (*Figure 6*). A number for some region (state) on the figure, divided by 100, denote how many times is the length of the OSM road network for that region larger in 2012 in respect to corresponding length in 2009. For example, for Serbia length of the road network in 2012 is 4,33 times the length of that network in 2009. This is the biggest increase in road network length (changes in OSM data) in Europe. The lowest increase is registered for Netherland (1,21).

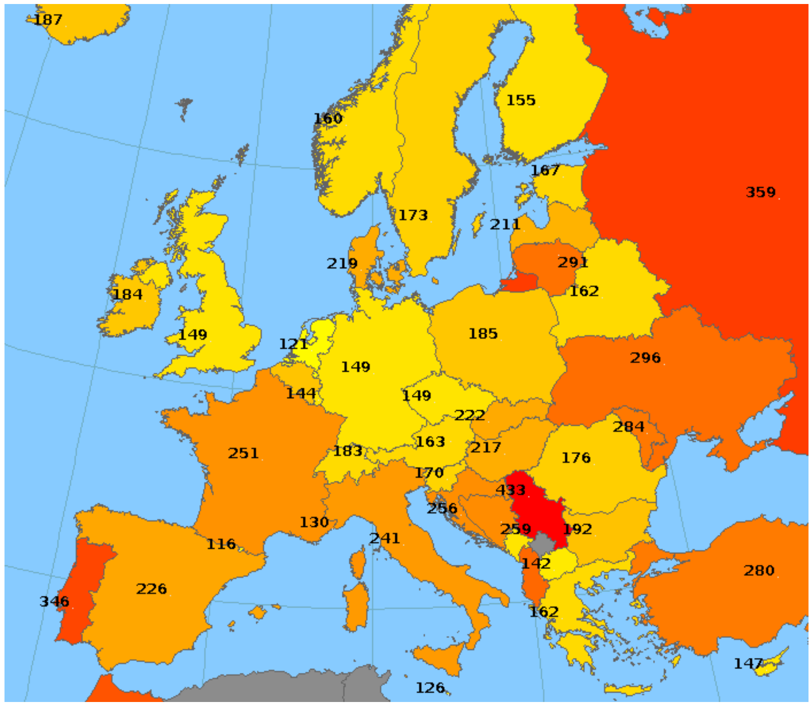


Figure 6. Changes for the road network for the period 2009-2012 (OpenStreetMap Forum, Web link)

All these studies revealed the following:

- There is a huge quantity of OSM data on a global level;
- Quality of OSM data is highly heterogeneous;
- For many application scenarios (for example, using maps as background in some GIS application or for data visualization) OSM data have immense value and they are highly useful;
- Application scenarios where OSM data, especially the data on road network, are to be used for the modelling of moving objects movement and the routing are very demanding;
- Quality of OSM data that are required for the modelling of moving objects movement and the routing is varying to a great extent for different regions in the world and in many regions the data are almost useless.

5. Models for Moving Objects Based on OSM Road Network

Even though this area has been developing for more than 15 years, models for moving objects moving on road network (MORN) are rarely used in practical applications. The reasons for this are the following:

- In addition to spatial data, a large quantity of telemetric data is collected; users require the whole range of functionalities that do not directly depends on location; therefore, models that provide very fast results are used, so system designers usually opt for models that do not take into consideration movement through road network (moving object moving freely in space - MOFS);
- Implementation of models based on road network is more complicated in comparison to implementation of models considering that object move freely in space (there is no network that constrains object movement);
- A proper network model and updated data have to be available for practical applications;
- Model implementation is directly linked to road network objects; therefore, the implementation process directly depends on the quality and up-to-dateness of the road network data; in cases where road network data model is available, but the data are not updated, the process of implementation is possible, but very difficult.

Model	Advantage	Disadvantage
Raw data	The simplest model. The easiest for the implementation. The smallest amount of information is stored. Fast queries providing data/results that can have official status (*Section 2*).	Either more spatial information is stored due to expected geometry queries than needed or spatial queries are too costly (in terms of time) because some additional on-the-fly calculations have to be made. Limitations and much lower performances for spatial queries based on road network and routing (it this functionality is provided by using external services).
Moving objects move freely in space (MOFS)	Simpler model. Easier for the implementation. Fast queries providing data/results that can have official status (*Section 2*).	Much more spatial information is stored. Limitations and much lower performances for spatial queries based on road network and routing.
Moving objects move on	Enables fast spatial queries. Enables fast and powerful queries based	The most complicated data model. The most complicated for the implementation. Requires availability of additional resources

Model	Advantage	Disadvantage
road network (MORN)	on road network data model. Enables routing.	– road network data. Much lower performances for spatial queries related to the original data providing results that are officially valid.

Table 7. Comparison of different models for moving object movement

The analysis provided in *Table 7* indicates that each of the listed models has some advantages and disadvantages. Therefore, it seems that combination of two approaches might be the best solution that would meet the users' requirements.

5.1. Requirements for the Model of Moving Objects

Factors that affect the model design are the following:

- Functional requirements (described in *Section 2* of this paper),
- Possibility of using road network model.

In order to develop and implement a model of moving object movement that is based on road network, it is required that resource that can be consumed quite fast and that is based on high quality and updated road network data is available. Additional request is the necessity that the link between the model of moving object movement and the road network model (resource) is established via data that have a permanent identification. Considering that very important aspect for FM AVL applications is also tracking of vehicle movement in the past, presence and eventually in the future, it is clear that it is of vital importance to provide maintenance of the history of changes in the database, i.e. the versioning of the road network data model has to be provided as well.

OSM data and the development of models based on these data opens possibilities that OSM data can be used as one of the resources supporting models of moving objects movement. This can be accomplished either by building a special model for this purpose or by modification of some existing data model.

The models of moving objects movement that is to be used for FM AVL scenario has to provide management with object in the past, presence and eventually in the near future (*Section 3*).

5.2. Starting Assumptions

Figure 7 depicts typical data acquisition process (registration of vehicle positions) for some FM AVL scenario. Vehicle positions are registered as blue dots. Of course, additional data on vehicle condition are also registered and linked to these positions.

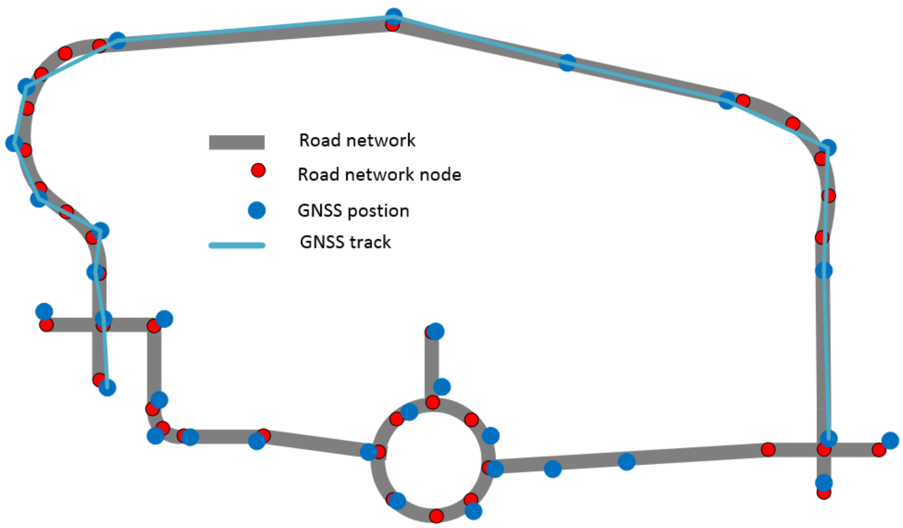


Figure 7. Registration of vehicle positions in respect to the road network in a typical FM AVL scenario

The devices used for this process have one important characteristic – they can be set to collect data by using algorithm that tracks changes in object movement (if object change its current direction of movement for an angle that is larger then defined treshold). It has already become standard that three parameters (time, distance, angle) are used to make a decision on data registration. Additional data registration can be triggered by any other event as well. For example, if value of any of the monitored parameters (*Section 2*) reaches some threshold.

In *Figure 7*, red colour indicates nodes of the road network model. These nodes are usually generated by using Douglas-Peucker algorithm.

FM AVL devices in vehicles are usually programmed to collect data according to the following principles:

- Travel distance: 200-500m;
- Travel time: 1 minute;
- Direction angle: 20-30 degrees.

Considering these facts, it can be stated that two paths depicted in *Figure 8*, i.e. the road network path (as given in the road network data model) and the vehicle movement path (as a result of data registration by GNSS/GPRS device) closely match each other.

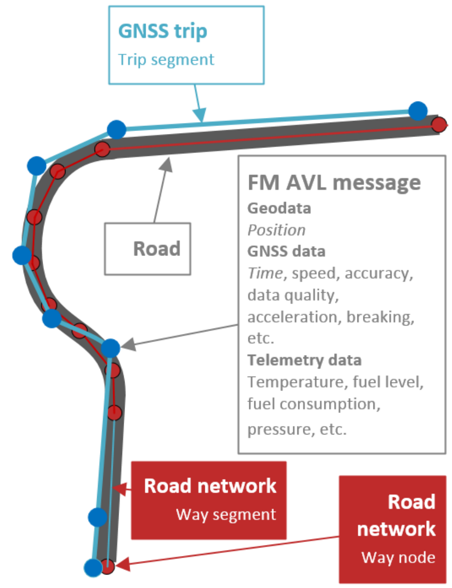


Figure 8. Data that are being registered in FM AVL system and key OSM road network data

Interesting information for this consideration is that average length for OSM road network segments in Europe is about 150m. Average lengths for OSM road network segments for some regions are given in *Table 8*.

Region	Serbia	Hamburg	Luxemburg	Slovenia
Average road network segments [m]	257.59	58.60	120.77	172.43

Table 8. Average length for OSM road network segments

If, in terms of informatics, one treats a road segment as a basic spatial unit for which information is to be linked to, than it can be concluded that OSM road network data have good spatial resolution and as such it can be used to store the information generated by FM AVL devices.

5.3. The Proposed Model for Moving Objects on Road Network (MORN)

Authors of this paper do not have any knowledge on the existence of some open-source data model that could be used in a way that all the requirements for FM AVL scenario could be satisfied. Two models (*osmosis* and *pgRouting*) have been used in the research described in this paper. These models have been modified to facilitate the modelling of moving object movement.

Figure 9 depicts FM AVL system architecture and its relationship with the system for the production of OSM data and maps.

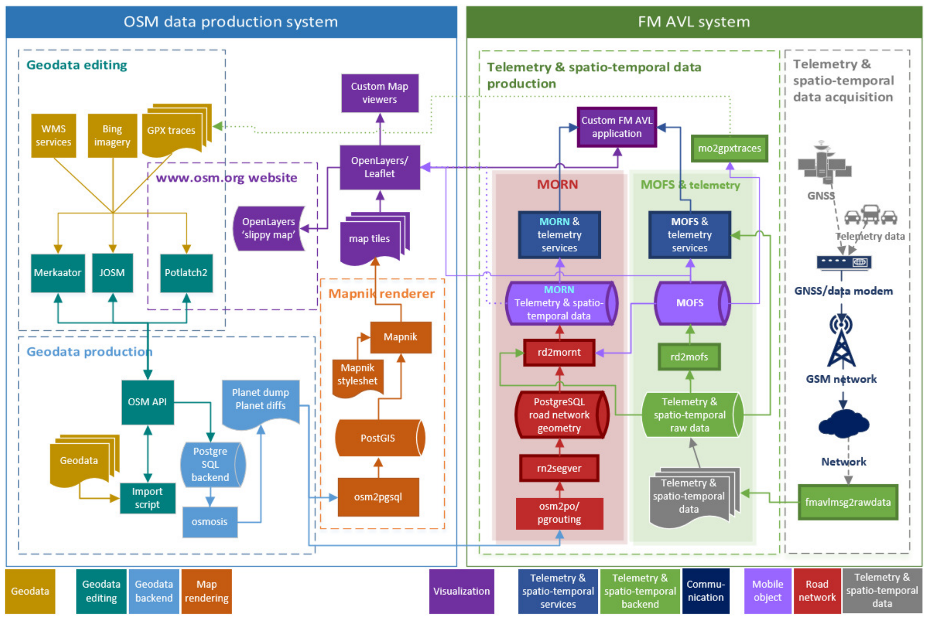


Figure 9. FM AVL system architecture and its relationship with the system for the production of OSM data and maps

Part of the FM AVL system is responsible for the acquisition of spatio-temporal and telemetric data. Process *fmavlmsg2rawdata* enables parsing of messages coming from GNSS/data devices. Usually, this component is implemented by using the TCP/IP or UDP server that translates binary messages coming through the network as streams into messages that are stored within the database.

Process *rd2mofs* translates the data that are stored within the database as a raw data into the moving objects' geometries. These objects move freely in space (MOFS), i.e. their movement is not constrained by the network. ISO/OGC geometries are used for this purpose.

Application services that are based on MOFS data (stored in the database) and on telemetric and spatio-temporal data (also stored in the database) contain all the functionalities (reports, alarms, graphics, etc.) that are consumed in custom FM AVL application.

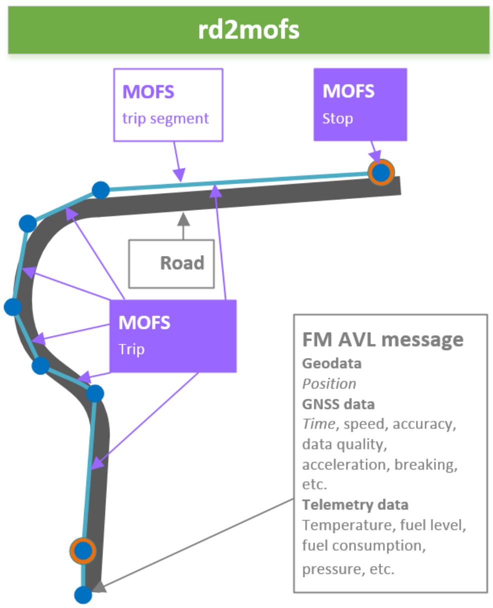


Figure 10. Process *rd2mofs* (transformation from the raw data to the moving objects' geometries (objects move freely in space - MOFS)

These components all together comprise MOFS telemetry, i.e. FM AVL system which is, in terms of the model of moving objects movement, based on telemetric data.

OSM data production system has important role during the process of building a MORN system. OSM data production system, in addition to classical maps (map tiles) and subsystems for data visualization (Open Layers and Leaflet), provides OSM data (via Planet dump data), as well as the road network geometry that is stored in the database. Very important role here plays the process *osmosis* that facilitates the data versioning.

Data model that is suitable for routing is provided by the process *pgrouting*. The process *rn2segver* is responsible for the preparation of the data in a format that is suitable for data modelling based on road network, i.e. road segments, as well as for providing the data versioning which is very important for this model.

The process *rd2mornt* reresents a component that is used to build a model of moving objects movement which is based on road network. Input pieces of information (*Figure 11*) for this process are the following: road network (segments/ways and nodes of the road network), messages from the FM AVL system and trip segments from MOFS system that is based on movement of moving objects freely in space.

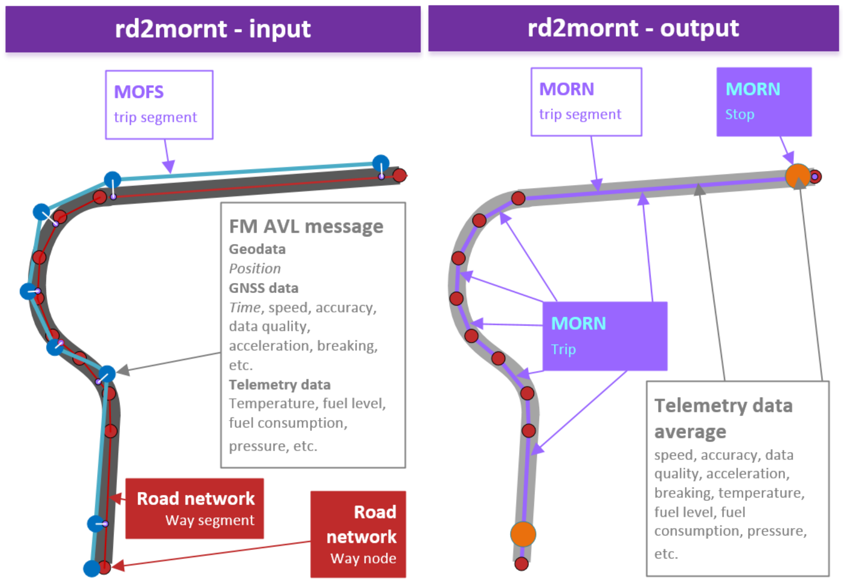


Figure 11. Process *rd2mornt* (transformation of the raw data to the moving objects road network data) – input data (left), output data (right)

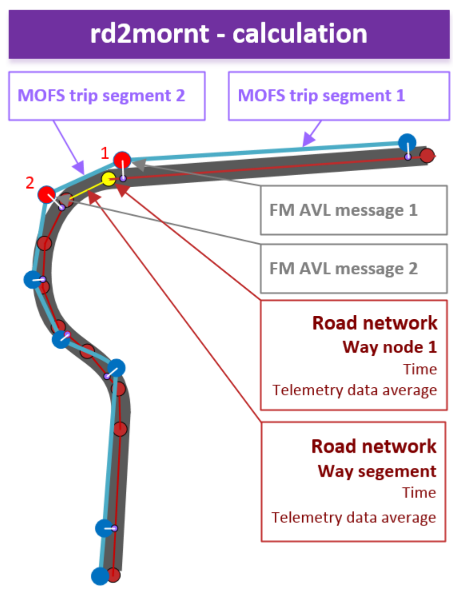


Figure 12. Process *rd2mornt* (transformation of the raw data to the moving objects model based on road network) - calculation

Algorithm for calculation is depicted in *Figure 12*. All the FM-AVL messages are sequentially processed. When message 1 and message 2 come for the processing, corresponding segments of the MOFS, i.e. trip segments 1 and 2 are assigned to these messages. These segments are taken to provide spatial context and to determine the closest corresponding segment of the Road network. After that, positions of message 1 and message 2 are projected onto OSM road network segments/ways. These projections and vehicle speed values for points 1 and 2 are used for calculation of estimated moment in time when vehicle passed through the point Way node 1. Similar principle is used for the calculation of average values of telemetric readings/parameters and all of these data are linked to the trip segment of the road network.

Figure 13 depicts the data comprising models MOFS and MORN.

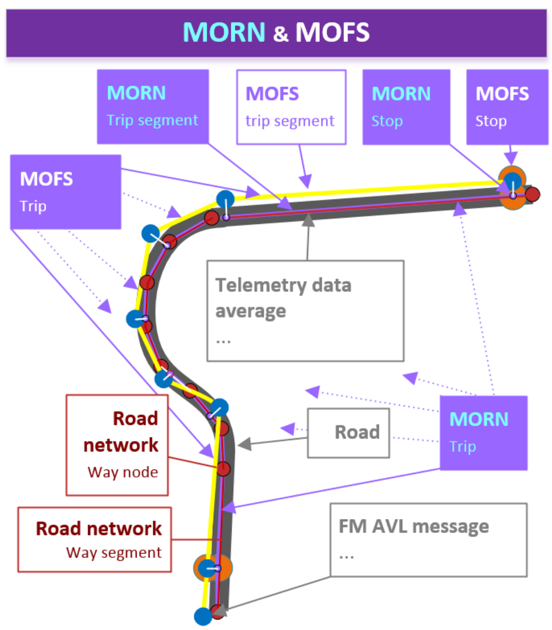


Figure 13. Basic elements of the moving objects models MORN and MOFS

Figure 9 shows the relationship between these two systems (MORN and MOFS). Planet dump files containing the latest version of OSM data which, due to the processing by *osmosis* procedures, also respect versioning are basis for building the road network geometry and topology. FM AVL system also uses Open Layers. Because of the cases where some parts of the network have insufficient data quality, an option is provided which facilitates conversion of the data from MOFS by using *mo2gpxtraces* process into GPX files suitable for processing by Geodata editing procedures.

5.4. Implementation of the MORN Model

Implementation of this model is still in the development stage. MobTrack:24 FM AVL system is used as an application framework (Vojinović et al. 2011, *Figure 14*). The system has been developed since 2006 and it has been in commercial use since 2008.

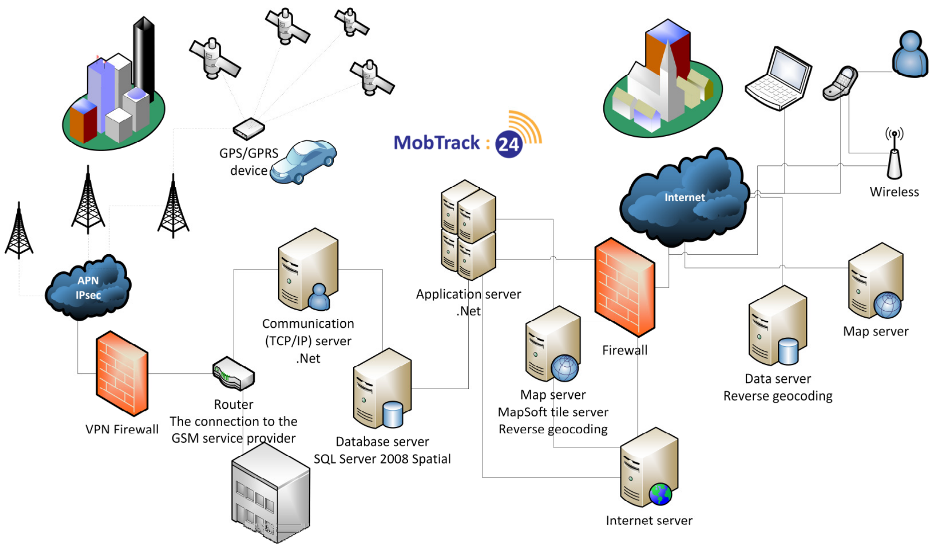


Figure 14. MobTrack:24 system architecture

System was working with raw data during the period 2008-2011. Since 2011 system uses MOFS data. Implementation of the model that is proposed within this paper has been going for several months.

OSM data for the whole Europe has been imported into the system. The procedure for data update has been implemented. The system has been implemented using the described MOFS conceptual model (*Figure 9*). The existing MOFS model, whose implementation is depicted in *Figure 15,* has been extended using concepts given in *Section 5.3*.

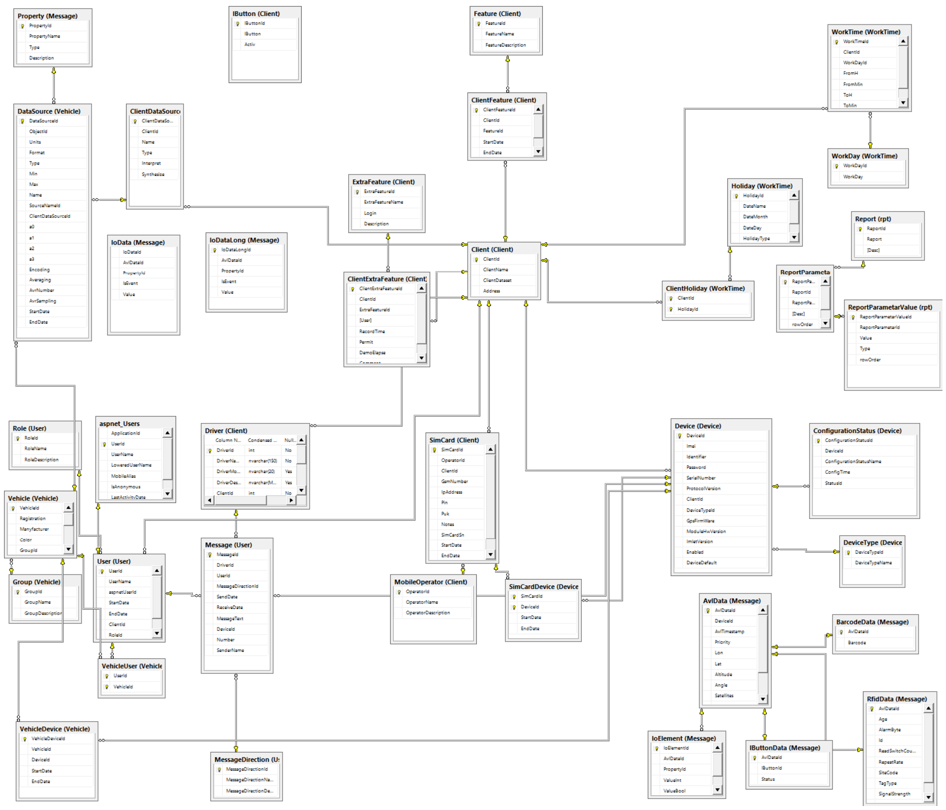


Figure 15. MOFS database design before MORN implementation

The implementation of MO Trip segment, MO Trip and MO Stop classes has been done in RDBMS. It can be seen from *Figure 16* how these classes have been implemented as database tables.

In the *Figure 16*, as well as in other figures, red colour indicates tables that implement OSM Road Network classes, grey colour indicates Raw Data tables/classes and purple colour indicates classes that belong to the moving objects model (white text denotes MOFS classes, blue text denotes MORN classes).

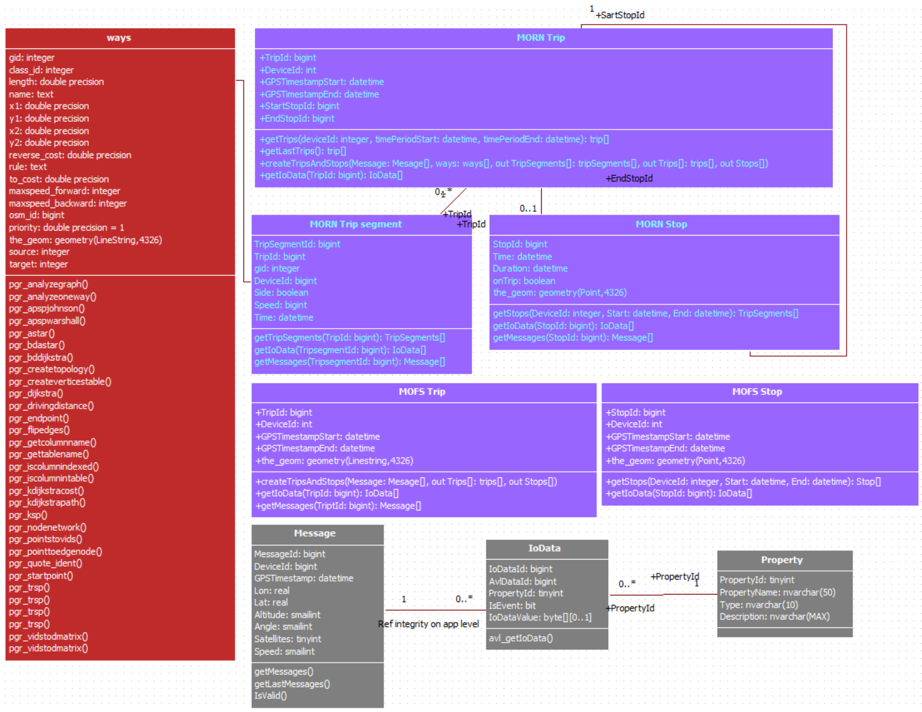


Figure 16. The MORN model obtained by extending existing MOFS model by using concepts proposed in this paper

The first results indicate the expected outcomes. One of them is that mobilisation of additional hardware resources was required in order to support work with the new data and the new data model. However, some of the functionalities are significantly improved in terms of execution time (display of vehicle movement, routing, etc.). The main problem that has been noticed is the quality of OSM data which is rather heterogeneous.

Some concrete indicators of the proposed FM AVL system capabilities and performances will be available in the next few months. What seems to be a conclusion is that, due to the insufficient quality of the data, system which is based on MORN in FM AVL scenario cannot be used as the proper solution for the model of moving objects movement, at least for the time being. Therefore, a back-up option has to be provided in case that MOFS model has to be used to answer the user's requests. However, the existing trend in the development and data quality increase indicates that systems based on the road network will have its use for FM AVL applications in the very near future.

6. Conclusions

One of the main reasons in implementation of models based on network is that data with sufficient quality were unavailable until recently. Even though the situation is not ideal nowadays, utilization of OSM data is certainly an interesting possibility.

There are various models using OSM data for different applications. Those which are suitable for moving objects movement modelling have to support object identification and to facilitate data versioning and routing. In order to provide efficient import of OSM Road Network data into the proposed models for moving objects movement, *osmosis* and *pgRouting* open-source libraries are used. Also, additional procedures have been proposed to obtain the proposed model which is suitable for modelling moving objects movement and for routing.

Based on collected statistical data related to road network geometry that are generated by FM AVL (GNSS/data) devices in standard FM AVL scenario it has been concluded that path geometries of moving objects roughly correspond to road network geometry. Density of the OSM road network and lengths of ways/segments comprising this network are such that they can be used as spatial reference for modelling and storing data generating by FM AVL devices. The whole concept of the proposed model for moving objects movement on network (MORN) is based on this assumption.

Implementation of the proposed model is done within a commercial FM AVL system. The system uses a model that is based on moving objects movement in free space (MOFS). Therefore, existing MOFS model is extended and modified according to proposed principles to implement the proposed MORN model. First experiences indicates that system which is based on MORN, due to the insufficient quality of the data, in FM AVL scenario cannot be used as the proper solution for the model of moving objects movement, at least for the time being. First test indicate that back-up option has to be provided when vehicle is located in a region without high quality road network. In that case, existing MOFS model has to be used.

Current research on OSM data quality indicates that there is a trend in data quality improvement. Therefore, moving objects movement models based on OSM road network are likely to be used in near future.

Model for tracking moving objects that is based on road network as well as the model that is not based on road network are equally significant and have to be supported for vehicle tracking application scenario due to required functionalities. Therefore, approach for testing these models will be designed and implemented using a commercial system for tracking vehicles - MobTrack:24. Procedures that will be designed and implemented within the system will enable utilization of the proposed models. These procedures will be designed to respect all the specifics and limitations of both implemented solutions. The results of these experiments might be very useful, because they might be significant in cases where the transition from models and systems that are not based on road network towards the systems and models that are based on road network is considered. It is expected that valuable practical experiences obtained from using these solutions within the MobTrack:24 FM AVL system will be obtained and analysed in terms of their performances.

The research is still in progress and it is going to be continued. The research done so far was mainly aiming at modelling moving objects movement on network. In future, the research will be also focused on routing. It is planned that two variants of AVL system for routing are to be implemented and tested in real applications. The first solution is based on using publicly available routing services such as Google and Open Street Maps and the second one is based on using *pgRouting* open-source solution for routing. Comparative analysis of the two implemented solutions will be conducted. The first solution is based on online services and that provides some advantages such as: updated spatial data, no need for the storage and maintenance of large quantity of data within proprietary system etc., but there are also some disadvantages such as limited availability, reduced system performances (in terms of system responsiveness and processing speed), price for using the services, etc. The second solution requires significant resources for the development and implementation of system for routing, investment in system installation, maintenance of the system and data etc., but, on the other side, this approach provides working autonomy and better system performances and responsiveness.

References

Cotelo, Lema. J. A. C., Forlizzi, L., Güting, R. H., Nardelli, E., & Schneider, M. (2003) “Algorithms for Moving Objects Databases”, The Computer Journal 46(6): 680-712. Networks”, The VLDB Journal, 15(2):165-190.

Corcoran, P., Mooney, P., & Bertolotto, M. (2013). Analysing the growth of OpenStreetMap networks. Spatial Statistics, 3, 21-32.

Girres, J., Touya, G., 2010. Quality assessment of the French openstreetmap dataset. Transactions in GIS 14 (4), 435–459.

Goodchild, M. F., & Li, L. (2012). Assuring the quality of volunteered geographic information. Spatial statistics, 1, 110-120.

Güting, R. H., Böhlen, M. H., Erwig, M., Jensen, C. S., Lorentzos, N. A., Schneider, M., & Vazirgiannis, M. (2000). A foundation for representing and querying moving objects. ACM Transactions on Database Systems (TODS), 25(1), 1-42.

Güting, R. H., de Almeida, V. T., & Ding, Z. (2006) "Modeling and Querying Moving Objects in Networks, The VLDB Journal — The International Journal on Very Large Data Bases archive, Volume 15 Issue 2, June 2006, Pages 165-190

Haklay, M. (2010). How good is volunteered geographical information? A comparative study of OpenStreetMap and Ordnance Survey datasets. Environment and planning. B, Planning & design, 37(4), 682.

Insight, B. (2012). M2M research series–the global wireless M2M market. Accessed July 27, 2013.

Macedo, J., Vangenot, C., Othman, W., Pelekis, N., Frentzos, E., Kuijpers, B., ... & Theodoridis, Y. (2008). Trajectory data models. In Mobility, Data Mining and Privacy (pp. 123-150). Springer Berlin Heidelberg.

Neis, P., Zielstra, D., Zipf, A., 2012. The street network evolution of crowdsourced maps: openstreetmap in Germany 2007–2011.

Porkaew, K., Lazaridis, I., & Mehrotra, S. (2001). Querying mobile objects in spatio-temporal databases. In Advances in Spatial and Temporal Databases (pp. 59-78). Springer Berlin Heidelberg.

Šaltenis, S., Jensen, C. S., Leutenegger, S. T., & Lopez, M. A. (2000). Indexing the positions of continuously moving objects (Vol. 29, No. 2, pp. 331-342). ACM.

Sistla A. P., Wolfson O., S. Chamberlain, and Dao S. (1997), Modeling and Querying Moving Objects. In:Proc. of the 13th Intl. Conf. on Data Engineering (ICDE), 1997, 422-432.

Steiniger, S., Neun, M., & Edwardes, A. (2006). Foundations of location based services. Lecture Notes on LBS, 1, 272.

Stojanović D.: Data Management for Mobile Objects in Location-Based Services, 2004, PhD thesis, Faculty of Electronic Engineering, University of Nis

Stojanovic, D., & Djordjevic-Kajan, S. (2003). Modeling and querying mobile objects in location based services. Facta Universitatis Journal.

Oort, P. V. (2006). Spatial data quality: from description to application. Unpublished Ph. D. Dissertation, Wageningen University.

OpenStreetMap Forum (2014). "Available online: http://forum.openstreetmap.org/" (accessed on 20 October 2014).

OpenStreetMap Tags (2014). "Available online: http://taginfo.openstreetmap.org/" (accessed on 20 October 2014).

Tansel, A. U., Clifford, J., Gadia, S., Jajodia, S., Segev, A., & Snodgrass, R. (1993). Temporal databases: theory, design, and implementation. Benjamin-Cummings Publishing Co., Inc.. Chicago

Vazirgiannis, M., & Wolfson, O. (2001). A spatiotemporal model and language for moving objects on road networks. In Advances in Spatial and Temporal Databases (pp. 20-35). Springer Berlin Heidelberg.

Vojinović, M., Cvijetinović, Ž., Kovačević, N., & Pušica, I. (2011). The Development of Location Based Services for Fleet Management. International Scientific Conference and

XXIV Meeting Of Serbian Surveyors Professional Practice and Education and Geodesy and Related Fields, 24-26, June 2011, Kladovo -„Djerdap“ upon Danube, Serbia.

Wiki, OpenStreetMap (2014). "Available online: http://wiki.openstreetmap.org/wiki" Data_working_group (accessed on 5 August 2014).

Wolfson, O., Xu, B., Chamberlain, S., & Jiang, L. (1998). Moving objects databases: Issues and solutions. In Scientific and Statistical Database Management, 1998. Proceedings. Tenth International Conference on (pp. 111-122). IEEE.

Wolfson, O., Sistla, A. P., Chamberlain, S., & Yesha, Y. (1999). Updating and querying databases that track mobile units. In Mobile Data Management and Applications (pp. 3-33). Springer US.