Student thesis series INES nr 230

Developing a Pedestrian Route Network Service (PRNS)

Cléber Domingos Arruda

2011

Department of Earth and Ecosystem Sciences
Physical Geography and Ecosystems Analysis
Lund University
Sölvegatan 12
S-223 62 Lund
Sweden

Developing a Pedestrian Route Network Service (PRNS)

Att utveckla en navigeringstjänst för gående (PRNS)

Cléber Domingos Arruda 2011.

Master degree thesis in the Division of Physical Geography and Ecosystems Analysis, Department of Earth and Ecosystem, Lund University, 30 ECTS

Supervisors
Lars Harrie (Ph.D)
Division of Physical Geography and Ecosystems Analysis Department of Earth and Ecosystem Sciences Lund University
Britta Duve Hansen,
Technical Surveyor, Department of City Planning

Abstract

Route network service is becoming increasingly popular. However, although there are significant amount of route network services there are still limitations especially to pedestrian network services. Pedestrians daily make decision about their navigation choices. Developing a pedestrian route network service (PRNS) involves several factors.

During this study the analysis of several routing network services have demonstrated that the geographical data is one of the most important factors in order to develop an own PRNS. Considering the idea of estimation two different datasets for the PRNS were evaluated OpenStreetMap (OSM) and Swedish national road database (NVDB). The use of the OSM dataset for developing the PRNS was made after the comparison between both the dataset. OSM has shown more advantages in terms of completeness of route for pedestrian navigation than NVDB. The OSM dataset was created and stored in the PostGIS database.

The implementation of own pedestrian network service is intended to facilitate the developments of new PRNS and analysis and comparison of others existing PRNS. The calculation and collection of the routes to be displayed for the user are performed by extended tools within PostGIS such as pgRouting and PostgreSQL respectively. The dataset’s network topology is related to the distance and determination of route choice by the pedestrian. Thus, Geographical Information System (GIS) is also one fundamental factor used in this study to evaluate and create results. The application was implemented in the city of Lund.

One of the limitations developing the PRNS is the lack of documentation for new functions which are released by pgRouting developers. Although OSM provides an essential network for developing the PRNS, some closed residential areas, parks, and open areas are not include on the network limiting the PRNS application.

In conclusion the PRNS is a useful application in order to assist pedestrians on their wayfinding in the city of Lund. It is also intended to help further development of new PRNS such as mobile PRNS applications. However, the PRNS must be improved and the dataset network requires updating and expansion for successful operations of the PRNS applications.

Keywords: Geography, physical geography, pedestrian route network service (PRNS), OpenStreetMap (OSM), Swedish national road database (NVDB), geographic information system (GIS), PostGIS, pgRouting, and dataset.

Sammanfattning

Navigeringstjänst blir allt mer populära, men även om det finns en betydande mängd tjänster, finns det fortfarande begränsningar, speciellt för fotgängare. Fotgängare gör dagligen flera val om hur de ska hitta den bästa vägen i sin närmiljö. Att utveckla en navigeringstjänst för gående (PRNS) involverar flera viktiga faktorer.

I denna studie analyseras flera befintliga nättjänster och studien visar att geografiska data är en av de viktigaste faktorerna för att utveckla egna PRNS. Två olika datamängder för PRNS utvärderades: OpenStreetMap (OSM) och svenska nationella vägdatabasen (NVDB). Efter en första utvärdering av de två datamängderna valdes OSM som visade sig ha flera fördelar för en fotgängares navigering.

Implementeringen av en egen navigeringstjänst för gående är avsedd att underlätta utvecklingen av nya PRNS och analysering och jämförelsen av andra befintliga PRNS. pqRouting och PostgreSQL beräknar och samlar in de rutter som ska användas, med hjälp av datamängden som finns i databasen PostGIS. Topologin för nätverket i datamängden är relaterat till avstånden och fotgängaren bestämmer vägvalet. Geografiska informationssystem (GIS) är också en grundläggande faktor som används i den här studien för att analysera resultatet. PRNS implementerades i staden Lund och syftet är att underlätta för fotgängare att navigera i staden. En av begränsningarna vid utvecklandet av PRNS är bristen på dokumentation av vissa nya funktionaliteter i pgRouting. Även om OSM tillhandahåller ett grundläggande nätverk av vägar för utvecklingen av PRNS, så saknas information om vissa bostadsområden, parker och allmänna utrymmen, vilket begränsar PRNS.

Sammanfattningsvis så är PRNS ett användbart program för att hjälpa fotgängare att välja väg i staden Lund. Vidare utveckling av PRNS kan exempelvis vara en mobilapplikation. Då måste dock PRNS förbättras och datamängden kräver uppdatering och utveckling för att bli framgångsrik.

Nyckelord: Geografi, naturgeografi, navigeringstjänst för gående (PRNS), OpenStreetMap (OSM), svenska nationella vägdatabasen (NVDB), geografiska informationssystem (GIS), PostGIS, pgRouting, dataset.

Acknowledgements:

The contribution of many people has inspired, motivated, and made possible this thesis accomplishment. I would like to extend my most sincere thanks for each one of you here. I would like to extend my greatest gratitude to the University of Belize, Lund University, EMCW LOT 10, for giving me this opportunity for further my studies and making me a much broader person.

I am very thankful to my advisors and supervisors Lars Harrie from the department of Earth and Ecosystem analysis – Lund University, and Britta Duve Hansen from the municipality of Lund who patiently guided me through this thesis research. I would like to thank my co-worker Eduard Mikayelyan who in hard and good times was always a support. I would express my greatest and warm thanks to all my friends from the Geomatics program who shared their experiences, knowledge, common reflections and culture during these two years.

I also would like to express my warmest acknowledgement to my wife Sol Faride, my daughter Sol Angela, and my son Lucas Santiago who companied, encouraged, and believed in me while we struggled together to achieve the completion of this thesis. Although being so far away from me I would like to send many thanks to all those who are and somehow made a significant contribution to my life, my grandmother, mom, aunties, brothers and sister. Thanks my family because you were indeed my greatest inspiration and I love you all.

Finally, but no least I would like to thank God for all the strength and spirit I got to finalize this thesis. Thanks all those who I have forgotten to mention here but deep inside in my heart I am very thankful that somehow you helped me to make this dream comes true.

Table of Contents

Abbreviations

ADT - Abstract Data Type
AJAX - Asynchronous JavaScript and XML
AND - Automotive Navigation Data
API - Application Programming Interface
CAD - Computer Aided Design
CC-BY-SA - Creative Commons Attribution-ShareAlike 2.0 licence
DB - Database
DBMS - Database Management System
EPSG - European Petroleum Survey Group
ESRI - Ergonomics and Safety Research Institute (UK)
ETL - Extract Transform and Load (spatial data)
FGDC - Federal Geographic Data Committee ()
FME - Feature Manipulation Engine
HTML - Hypertext Markup Language
HTTP - Hypertext Transfer Protocol
ISO - International Organization for Standardization
GeoJSON - Geospatial data interchange format based on JavaScript Object Notation
GIS - Geographic Information System
GPL - General Public License
GPS - Global Positioning System
GPX - GPS eXchange Format
JDBC - Java Database Connectivity
JOSM - Java OpenStreetMap
LBA - Location Based Applications
LBS - Location Based Service
LGPL - Lesser General Public License
NVDB - Nationella Vägdatabasen
ODBC - Open Database Connectivity
ODMG - Object Database Management Group
OGC - Open Geospatial Consortium
OLS - OpenGIS Location Service
OODBS - Object-Oriented Database System
ORDBMS - Object Relational DBMS
OS - Operation System
OSM - OpenStreetMap
PDA - Personal Digital Assistant
PHP - Hypertext Preprocessor
POI - Point of Interest
PRNS - Pedestrian Route Network Service
RDBMS - Relational DBMS
SALAR - Swedish Association of Local Authorities and Regions
SOA - Service Oriented Architecture
SQL - Structured Query Language
UI - User Interface
WGET - World Wide Web gets
WKB - Well Known Binary
WKT - Well Known Text
WMS - Web Map Service
XML - Extensible Markup Language

Part I: Background and theory

1. Introduction

1.1 Background

Pedestrian route network service (PRNS) is used to assist pedestrians for their wayfinding. The main focus is on pedestrians who navigate in the municipality and are searching for safe, short, and convenient routes to navigate from one point to another. To do this, the pedestrians should plan. The idea is to eventually develop an application that would help the user to identify the best possible route utilizing their knowledge of the municipality and the PRNS application.

The routes of the PRNS have to offer a wide range of qualities (e.g. convenience, attractiveness, and safety) in order to fulfill the requirements of different types of pedestrians. Besides those route qualities there are three other important qualities that should be taken into consideration based on the previous qualities. There are advantages and disadvantages for pedestrians regarding their navigation in contrast to drivers. For example, pedestrians are always "free" to choose the best and convenient way for their navigation while drivers have the same freedom but they have to follow signs and roads regulations (e.g. "motorcycles only"), speed limit, etc. Pedestrians are exposed to the environment surrounding them (e.g. climate, noise, pollution, etc.) [29, pp. 109-118]. The variety of attributes that influence a pedestrian's choice of route is extensive. It depends on the reason and the preference of each pedestrian in addition to other qualities and requirements of the pedestrian's needs such as a safer or shorter route, etc. Planning a route network requires an analysis of the data used (e.g. authority and free or commercial data), routing service, database management use, and other aspects.

Today, most navigation systems are not designed specifically for pedestrians. Instead, these navigation systems are usually focused on the navigation of motor vehicles. Some examples are the personal digital assistant (PDA's) and navigation systems found in smart phones. Pedestrian navigation implies crossroads, parks, streets, passes between buildings, sidewalks in the community, etc. Often these details are not presented in standard navigation systems. Furthermore, one should consider that there are several different groups of pedestrians within a municipality. Therefore, it would be interesting to develop a PRNS to assist differing types of pedestrians to navigate within the municipality.

This project is intended to develop a PRNS to help pedestrians in a wayfinding approach. PRNS can help the pedestrian identify the complex features change in a developing municipality. Using the PRNS, pedestrians can find the best route for their navigation in the municipality. The infrastructure in a city frequently changes; to identify those changes is crucial to pedestrians especially in complex urban areas. This is interesting for the different type of pedestrians who navigate the city daily or simply look for a more convenient way to walk. For example, students, workers, and tourists are some of these pedestrians who would have different intentions for using a particular route to navigate. In the case of students and workers, the safer and shorter the route is, the better it is. While for tourists, an attractive, well-structured, and well-marked path is preferable. Thus, it is also important to develop a PRNS application to provide accessibility to reliable information about the routes and the municipality for any type of pedestrian.

This project aims at assisting and providing pedestrians with fast and convenient information about route in a municipality. This project was started because of the lack of specific route network information for pedestrians in the municipality of Lund. The PRNS can enable service for pedestrians in the municipality just like the existing routes, such as route networks for motor vehicles and bicycles. The interest in developing navigation tools for pedestrians is increasing in the context of computing technology and location-based service (LBS), especially with navigation tools for pedestrians who have visual, perception, or other types of impairment [26, pp. 331-338]. Thus, providing users with services such as the PRNS motivates and enhances this interest toward the development and creation of successful navigation tools for pedestrians.

The thesis proposal is to develop a PRNS using a pedestrian navigation network in the municipality of Lund. This implies combining the available and accessible data and tools such as open source or commercial route network licenses to create an application allowing users to generate routes and maps for navigation purposes within the municipality of Lund.

The pedestrian route network service is not as common as motor vehicles or bicycle route network services in relatively small municipalities. Additionally, those type of route network services are usually integrated in mobile hand units, which are not specific for pedestrians. The use of a route network service varies from region, to region usually depending on the demand of the municipality.

For example, Stockholm is a complex and busy metropolis; route planning for a navigation service is being implemented there by e-Adept, providing accessibility solutions for pedestrians, especially those with disabilities [4]. However, it is done with the use of mobile telephone and PDAs with integrated telephony capabilities. There are also other projects like

HaptiMap that focus more on providing services to navigators with visual impairment [14]. To develop the PRNS the use of the wayfinding concept that involve, the pedestrian's orientation, route selection, and recognition of the path, among other aspects, are essential. Thus, every element of the creation of a PRNS has to be analyzed and adapted.

1.2 Objective

The general objective of this thesis is to study pedestrian route network services (PRNS). The specific objectives are:

a. Evaluation of geographic dataset for pedestrian route network service (PRNS).
b. Implementation of an own PRNS.
c. Comparison of the own PRNS with other available PRNS services.

1.3 Methods

This section introduces the general method of developing a pedestrian route network service (PRNS). Some essential aspects for developing a PRNS are the preparation, analysis, comparison and implementation of the data and own PRNS (cf. Figure 1 - Part II). These aspects among others are used to determine the use of the right dataset for the implementation of PRNS. Part II of this thesis describes two case studies. Section 9 describes the method and preparation of this thesis. It focuses the workflow, how to convert and project, access and manipulate, and view the datasets. The methods used to perform the analysis or evaluation and comparisons of the data are described in Case study I in Chapter 10 which concentrates on how to download, compare geometric and topological data, and fix the dataset before the creation of PRNS. Subsequently, one of the methods used to reach the results obtained from the evaluation of the data and own PRNS implementation is described in Chapter 11. It describes the importance of application, data, programming language, technical solution, completeness and analysis of PRNS. Chapter 12 is the Case study II where methods of comparison between own PRNS and other existing PRNS are performed in order to obtain results and draw conclusions from this study.

The development of the pedestrian route network service (PRNS) involves several steps, which ensure the completion of the PRNS. First, in order to identify the needs of the PRNS, regular consultations on materials, documentation, and resources performed under supervision and advice set off the creation of PRNS. Following existing and suggested ideas of route network service became crucial to achieve this study. Second, a list of activities such as the writing of theoretical part for a better understand of the problem, the collection of data, documentation, resources, and information related to the topic, follow the objectives stated in section 1.2, and consistent survey and use of program systems are some of the methods used to develop the PRNS (cf. Figure 1 – Part I).

1.4 Disposition

The theoretical section of the thesis in Part I start by describing the geographic data as well as storage methods and examples of geographic data. Among other things, it explains the complexity of using geometric data and its storage in a spatial database. Furthermore, Part I also describes the data source OpenStreetMap and the Swedish national road database (sw. NVDB) and network algorithms.

One of the main parts of this thesis is the evaluation of the network data. The evaluation of the data includes the comparison and selection of the best dataset, the preparation and/or tailoring of the dataset, and estimation of its quality. Those fundamental needs of this evaluation depend on each dataset's characteristics and purposes. Then the evaluation and comparison between existing route network services and this thesis' own PRNS.

The implementation of the pedestrian route network service (PRNS) is intended for users such as students, parents, tourists, etc., who wish to find the most convenient way to navigate based on the shortest path in the municipality of Lund. Most of the route network service developed today is utilized for motor vehicle navigation. However, the PRNS is projected to be extended and developed for not only car, but other future purposes such as mobile applications.

To conclude this thesis, the analysis, result, discussion, and conclusions are the last sections summarizing the entire study of the PRNS for the municipality of Lund.

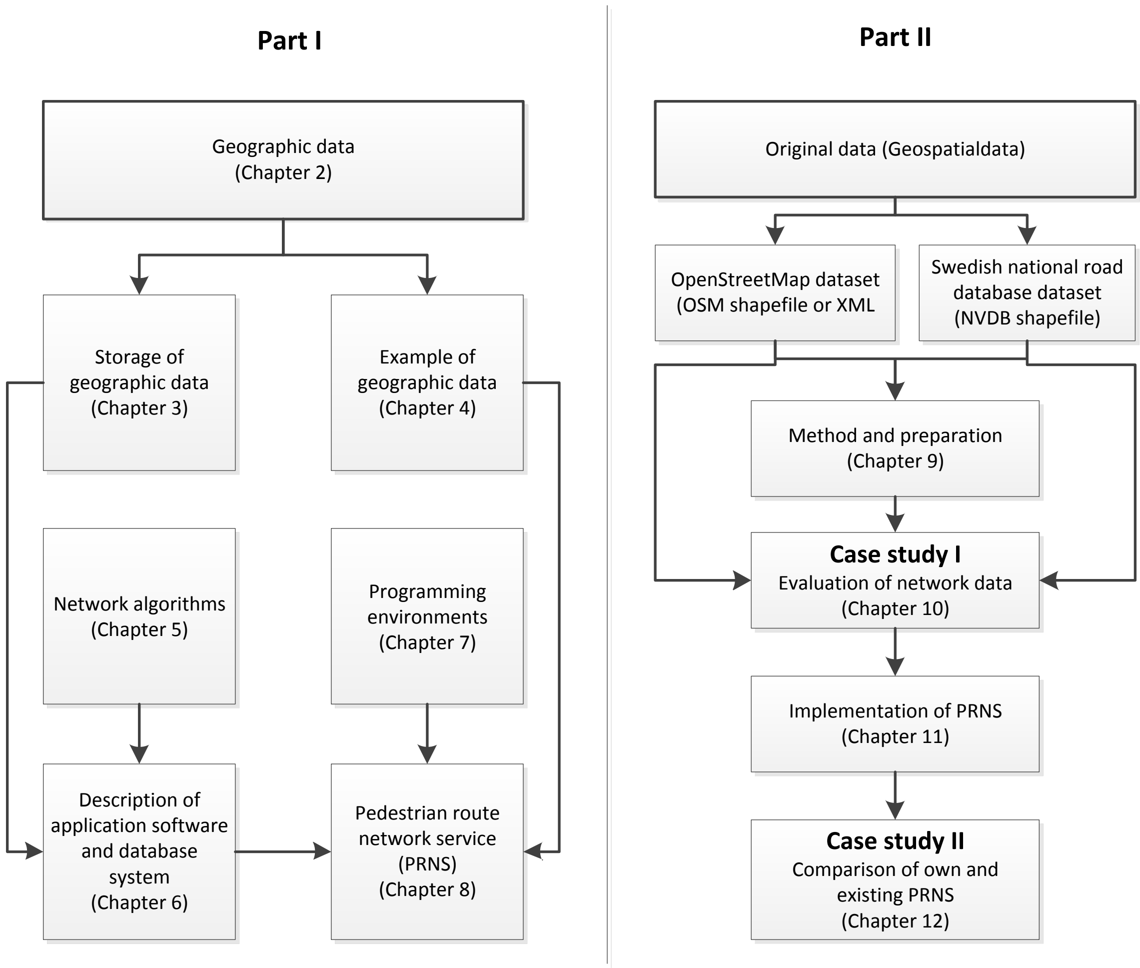


Figure 1 - Show the structure and the workflow of the methods used in this thesis.

2. Geographic data

2.1 Definition

Geographic data can be described as spatial data that can be built up from atomic elements or facts about the complex geographic world [25, pp. 68-70]. Geographic data link place, time, and attributes. Geographic data are more than maps; they describe objects and relations in space in a certain scale. Geographic data are usually used in a wide range of applications such as route network services for vehicle and pedestrian navigation. There are several methods of data collection; this process is considered one the most exhaustive and often most expensive task in terms of Geographical Information Systems (GIS). Geographic data comes from a variety of sources and generally the data about a particular feature is unique, correct, and representative of physical reality (e.g. earth).

Attributes identify a place or an entity differentiating it from another attribute (cf. Figure 2.1). “The range of attributes in geographic information is vast” [25, p. 69]. Time and location are basically tied with the property or attribute of a particular entity [25, pp. 69-83]. Thus, all geographic phenomenons have spatial and temporal components [52, pp. 1-34]. The world is very complex and it is difficult to know exactly the amount of atomic elements needed to make a complete representation of geographic data. The combination of these attributes, time and places can make it even more complex if one analyzes it in detail.

There are two types of geographic data: vector and raster. They can be classified in primary and secondary geographic data depending on their origin. The data can be obtained in either digital or analog format. Primary geographic data source are obtained with direct measurement of objects (cf. Table 2.1) in digital format which is used for GIS project. For example, raster SPOT and IKONOS Earth satellite images, and vector building survey are considered primary geographic data. Secondary geographic data source are derived from earlier studies or from other systems (cf. Table 2.1) in digital and analog datasets. For example, raster scanned color aerial photographs, and paper maps that can be scanned and vectorized. One has to convert secondary geographic data in the desired format in order to use it in a GIS project. Table 2.1 shows some of the classifications of the primary and secondary geographic data between vector and raster format as well as their origin.

Table 2.1 Based on [25, pp. 199-201] the table shows the geographic data classification for data collection purpose with example of each.

Classifications	**Raster**	**Vector**
Primary (Direct measurement)	Digital satellite remote-sensing image	GPS measurement
	Aerial photographs	Survey measurement
Secondary (Collected from existing source)	Scanned maps or photographs	Topographic maps
	Digital elevation models from topographic map contours	Toponymy (placename) database

Geographic data consist of geometric data and attributes. Geometric data are of several types, for example points, linestring, networks or polygons. In order to create routing network service, you need a network with a correct topology.

The vector data type such as point, line and polygon represents different objects in the real world. Points consist of as a single coordinate pairs [52, p. 85] and can represent buildings, soil pits, wells, etc usually small geographic features (cf. Figure 2.1b). Another type of geometric data is a line (polyline). Lines have a series of ordered coordinate pairs that can also be called polylines (cf. Figure 2a). Lines represent real-world features such as roads, streams, railroads, etc. Polygons have lines segments that are closed to form a certain polygon area. A polygon is usually used to represent large areas such as a country, city, soil areas, etc. Notice that a polyline forms the boundary of the polygons. In another words, the end-point of an *edge* of the polyline is denoted as the *vertex* of the polygon (cf. Figure 2.1c). However, an array of cells does the same feature representations in raster data (cf. Figure 2.1d) [25, pp. 181-184].

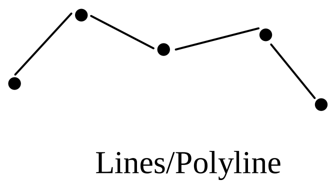


Figure 2a - Lines/Polylines

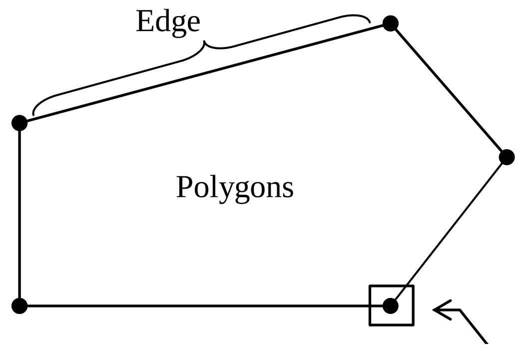


Figure 2.1c - Polygon

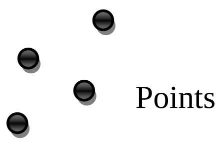


Figure 2b - Points

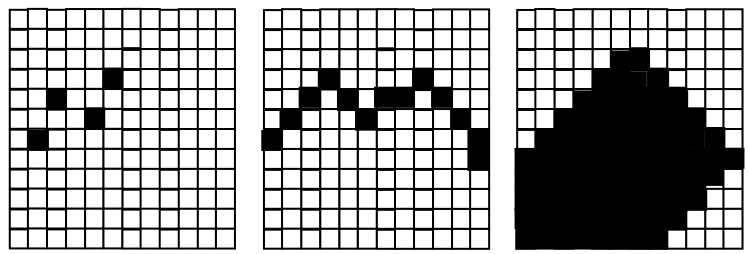


Figure 2.1d - Raster representation of points, lines, and polygons

Attribute data is used to describe non-spatial properties of features and attribute data can be of several types [25, p. 69]. There are some common used types of attributes used to identify or differentiate one entity from another such as the name of a particular place, the name of a road, length, speed of that road, etc. A typical example would be a type called *Road* which can represent all the roads in a region such as Route 22 or Dalbyvägen as members or instances of that type. Instances are identified within a class of entities by the attribute data that also helps to distinguish an instance from other member of that same class. Attributes can be described in the form of text, numbers, or geometry; however, there are other types of data that can compose the attributes data of one entity [8]. In that manner, a table can be generated to represent the real world with information about a road entity (cf. Table 2.2).

Table 2.2 - Attribute table of Roads.

Id	Road_name	Speed	Type	Location	Length
1	Stortorget	50	Motor	Lund	92.4657
2	Dalbyvägen	60	Urban	Lund	270.569
3	E22	80	Motor	Lund	654.630

2.2 Metadata

Metadata provide information about the dataset and have some information such as the owner, coordinate system, file type, and projection (cf. Table 2.3). The metadata record represents the *who, what, when, where, why,* and *how* of the resource [9]. It provides essential information for further processes. Metadata is much like a library's catalog [25, pp. 245-247]; data search and discovery organizes all the relevant information for an easier retrieval of the data. This, should specify when the data were created, what the name of the applications used to create the data, where the data location or source are from, who created the data (e.g. author's phone number, address, and e-mail, etc.), and where it can be used. Thus, metadata basically describes the contents in a dataset allowing easier access to and interpretation the data.

Table 2.3 - Example of metadata.

Metadata
Owner
Coordinate System
File Type
Projection …

There are standards of metadata and one of them is the ISO 19139. Geographic data are shared and distributed among private and public organization (e.g. Swedish national road database, and OSM respectively). In order to maintain the standard the ISO 19139 was created to provide a common XML specification for describing, validating and exchanging geographic metadata. There are some common task the user follow for managing the ISO 19139 compliant metadata (cf. Table 2.4). The ArcGIS metadata (cf. table 2.4) can be generated by the use of the ArcCatalog metadata editor. It follow the Federal Geographic Data Committee (FGDC) which is an interagency committee that promotes the coordinated development, use, sharing, and dissemination of geospatial data on a national basis to generate Content Standard for Digital Geospatial Metadata (CSDGM) which is the United States of America federal metadata standards. The ISO 19115 is one of the standards followed by ArcGIS metadata. It defines the schema required for describe geographic information and services. The ISO 19115 provides information about the identification, the extent, the quality, the spatial and temporal schema, spatial reference, and distribution of digital geographic data [9].

Table 2.4 - Examples of metadata based on ArcGIS 10 help: showing elements of metadata.

ArcGIS Metadata	**Metadata format: ISO 19139**
Resource Identification	**Roads**
Citation	**ISO 19139 metadata content**
Title: Pedestrians Route Network	• Resource Identification Information
Creation date: 2010	• Spatial Representation Information
Publication date: 2010	• Reference System Information
Resource Identification Information	• Data Quality Information
Citation	• Distribution Information
Title: Pedestrians Route Network	• Metadata Information
Creation date: 2010	
Publication date: 2010	
Presentation format: mapDigital	

2.3 Route network data

Route networks can be defined as the representation of paths, streets, and roads essentially used for navigation purposes. A route or road network should describe its information in a way that informs us about its fundamental aspects in the real world. They are usually represented by lines or links that run through defined points (nodes) (cf. Figure 2.2). For example, in the Swedish national road database (sw. NVDB), it is documented that a reference line's position is stored in three dimensions. There are three coordinates used to represent the roads: an x-coordinate, a y-coordinate, and a z-coordinate [46, pp. 5-6]. The z-coordinates correspond to the height above the sea-level, thereby, allowing the road to fit (e.g. in a map) with other features such as lakes, buildings, and parks.

Below Figure 2.2 based on [52, pp. 1-34] shows a vector network which has connection between nodes representing the roads features.

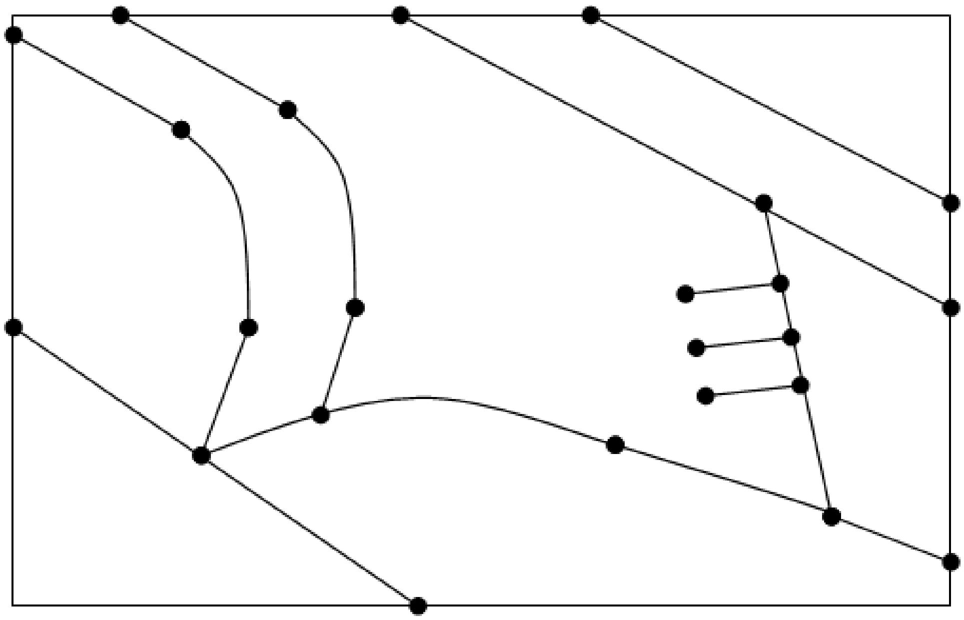


Figure 2.2 – A sample of a route network. (Source: adapted from [52, p.17])

2.4 Pedestrians' network

In order to create a pedestrian's network you have to consider paths, sidewalks, ramps or stairways, bridges, and crosswalks where navigation is done by different groups of pedestrians. For example, students, disabled or handicap pedestrians using wheelchair, and parents navigating with babies using carriage wagons [17, pp. 1-128]. You should also consider open spaces such as parks, market places, or "large open squares" frequently used by pedestrians. Usually, pedestrians navigate through parks instead of follow a particular route in a route network (cf. Figure 2.3). These optional pathways are often not stored in route networks like bicycle and motor vehicles route. However, it is relevant to store the pedestrian's route network data in order to help pedestrians to find, for example, the shortest and most convenient navigation path within the city.

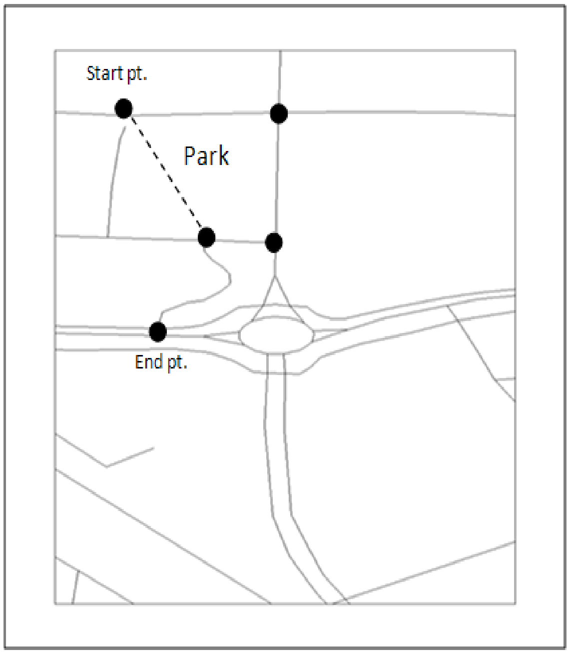


Figure 2.3 – Sample of pathway in an open space – (e.g. parks, market place, large square, etc)

___________	route network (e.g. roads, streets, etc)
-------------	optional route (e.g. pedestrian walkway, pathway)
●	vertex

2.5 Quality of geographic data

In order to understand quality one has to take into consideration several aspects of data representation. Geographic data are directly related to their quality [23, pp. 93-96]. The representation of the real world is very complex, so it is impossible to make a perfect representation of the world; thus, uncertainties will always be present and quality labeling is essential.

According to Kresse and Fadaie [23, pp. 93-96] data quality is distinguished by several terms. However, there are two main categories that describe the quality information of the data. The first category is the data "quality elements", which contain *quantitative* quality information. In contrast, the second category is the "data quality overview elements", which contain the *non-quantitative quality information.* Data quality elements and data quality overview elements are divided into groups (see below).

There are several quality parameters: data accuracy, precision, lineage, bias, completeness, consistency, currency, format, granularity, reliability, and timeliness are some of the aspects of data quality reference. Some aspects of data quality reference from the US Federal Geographic Data Committee's are: attribute accuracy, positional accuracy, logical consistency, completeness, and lineage. But there are others components of quality which can be found on the *Federal Geographic Data Committee (FGDC's) Web pages (*www.fgdc.gov*).* Data quality has several aspects, but two major aspects are the accuracy and its precision [52, pp. 336-339]. Besides accuracy and the precision of the data, information about a dataset must be kept up-to-date.

The concept of data quality is crucial to geographic data. The ISO basic standards definition of quality is ISO 9000 which has the same quality characteristics of ISO 19113 [51, pp. 1-3]. The ISO 19113 denotes that "*inherent characteristic of a product, processes or system related to a requirement*". To establish the principles for describing geographic quality, the characteristics of any product and the quality of data make the geographic data valuable.

Today, the contribution, sharing, and development of geographic data are vast and increasing. It is due to the amount of applications and users of geographic information systems (GIS) that demands different levels of quality. Data are often used in other applications besides the one for which it was intended by the producer. Therefore, the quality of information is one of the vital aspects in order for one to determine the geographic data to be used in a particular application. The following are some quantitative elements of spatial data quality [23, pp. 93-96]:

- Completeness inform about the presence and absence of relevant data in the database to be disregarded during analysis.
- Logical consistency concerns the degree of adherence to logical rules of data structure, attribution, and relationships.
- Positional accuracy is the accuracy of the position of the features.
- Temporal accuracy describes the accuracy of the temporal (e.g. data and time) attributes and temporal relationship of the features.
- Thematic accuracy refers to the correctness of quantitative and non-quantitative attributes, and correctness of features classification and relationships. In another words, features can be classified incorrectly while being analyzed. For example, in a route network a polyline can be classified as a highway when it is a street in reality.

  The previous paragraph describes the quantitative elements of spatial data quality. The following three subgroups defined by ISO 19113 in this section are the data quality overview elements which include non-quantitative quality information and it composed of [23, pp. 93-96]:

- Purpose describes the reason for creation and the intention of use of a particular dataset. The purpose has relevant information motivating these aspects.
- Usage describes the application of the dataset used.
- Lineage simply delineates the history of the dataset with relevant information regarding the dataset (e.g. the creation process, its life cycle, etc).

3. Storage of geographic data

3.1 File system and databases

Digital data, such as geographic data, can be stored in file systems or databases. A file system is a method of storing and organizing computer files. Operating systems (OS) have file systems that organize all created folders and files, such as MS-Word document (e.g. doc), images (e.g. jpeg and png), and other types of files. The file system stores, organizes, and allows the creation of new files and folders on the hard disk. The folders and files are added in a "file tree" (cf. Figure 3.1).

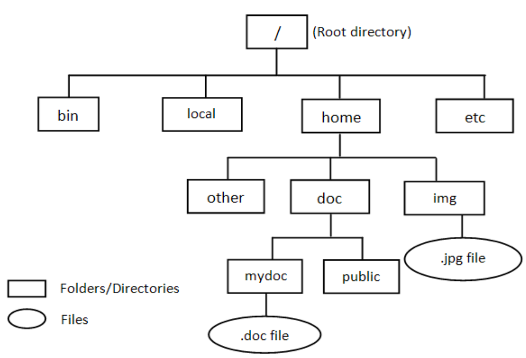


***Figure 3.1 – File System based on** [10].*

In a database a database management system (DBMS) is used for the data storage and retrieval (cf. Figure 3.2). A database management system (DBMS) is a software application which creates processes facilitating the organization, storage and access of data. A DBMS is required when one works with complex databases, especially when there are many users using and/or accessing the databases. The DBMS keeps the integrity and the longevity of the databases.

The DBMS has several capabilities that are of great interest for GIS users. Below there is a list of these capabilities [25, pp. 218-220].

- The DBMS data model mechanism is capable of representing objects digitally in a computer system (e.g. text, numbers, floating-point, integer and dates). In order to manage geographic object type the DBMS can be extended.
- The data are usually in a standard data type (e.g., characters, numbers, and date). If the data is not in a standard data format it can be converted into the standard required by the standard loader. It is done by the use of converter software applications. Then the data can be loaded after the conversion.
- The DBMS indexes are used to make the searching of data much faster.
- The DBMS query language supports the queries. The most common language is SQL (Structured Query Language).
- The DBMS security provides control and access to the database. There are options for restricted access and manipulation of the database for creating, updating, reading, writing and deleting permission within the database, accomplished through the DBMS security.
- The DBMS-controlled update manages and coordinates the multiple-user access and transactions in the database.
- The DBMS backup and recovery protects the data in the database from system failure and undesired updates in the database that can occur unexpectedly.
- The DBMS database administration tools basically maintain and manage all the actions cited above, such as maintaining the indexes and backing up and recovering the database.
- The DBMS application programming interface (APIs) allows user interaction.

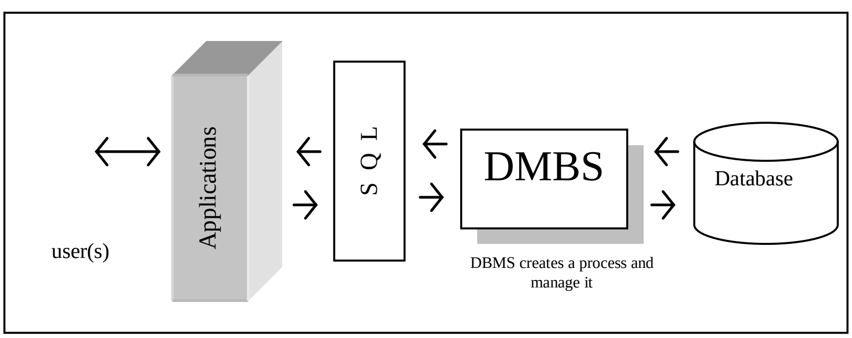


Figure 3.2 – Simple sample of a data base management system (DBMS).

Large amount of data is generally managed by DBMS. However, one should understand the concept of secondary storage which helps to determine the best way of storing and retrieving data. Data are constantly being transferred from primary to secondary memory and data transfer should be minimized to avoid the decrease of its access time.

The GIS databases are based on DBMS nowadays. These capabilities are essential and capture the GIS users' attention, thereby motivating the GIS vendor to enhance the software by including the DBMS in it or by providing an interface that supports the DBMS. There are different types of DBMS for GIS users, but the three main types are: relational (RDBMS), object oriented (ODBMS) and object-relational (ORDBMS).

3.2 Relational database

A relational database is one of the most used database systems. One of the reasons why relational database is so popular is due to its simplicity and capability of relational and associational attributes. A database management system (DBMS) is managed by the relational DBMS or RDBMS whereas the database relationships are treated in the form of a table. Organization of data is one of the essential strongholds of a relational database. The organization of the data facilitates the collection and matching of common attributes characteristics (e.g. in a route network group by type of road) found in the dataset. These processes are usually performed using the relational database management system (RDBMS) and the data is stored in tables allowing a consistent and well-organized method of structure. Information is organized in tables and is defined as relation tuples or row are all distinct from one another, while the columns keep the attributes that are stored in a particular order in which must maintain the relation of the data and each attribute name.

The order that the rows and columns are is irrelevant but data has to be organized in an order where the information is not duplicated. Usually it is done by having multiple tables separating the information and normalizing the tables. Consequently, the data can be easier retrieved from the database using a structure query language (SQL) [43]. However, one still have to deal with the limitation that relational database presents such as the storage of more complex data (e.g. geometric data) described in section 3.4 – 3.5.

3.2.1 Structured query language (SQL)

Structured Query Language (SQL) is the communication language of the database mostly used by the relational databases models. SQL provides the possibility for the user to create, maintain and search for data. It involves the storages of data into and retrieves of data from tables in the database (cf. Table 3.1).

Customer Table:

Table 3.1 – Sample of Relational Database table.

CustID	FName	LName	Address	City	Zipcode
1	Mikael	Enkland	Kämnärsvägen 11J lgh 457	Lund	22640
2	Silva	Perez	Stora Tvärgatam 7P	Lund	22646
3	Robert	Johansson	Stora Tvärgatam 4A	Lund	22646
4	Maria	Nunes	Korsgatan 67C	Malmö	22260
5	Ola	Hall	Klostergården 56F	Lund	22122
: :	: :	: :	: :	: :	: :

One type of SQL command is request. The user is the one who determine the specification of the query in order to get the desired results from his or her request. For example, "*retrieve or find all records that have a zip code 22646*". A simple SQL for this statement could be as follow:

```
SELECT      Customer.FName, Customer.LName, Customer.Address
FROM        Customer
WHERE       zipcode = '22646'
```

The DBMS fetches all the information from the tables by using an index method which usually is sorted in an alphabetical order either by a unique numerical field such as the social security number or field(s) which distinguish each record on the tables.

Recently the most common way of data storage has been to use a relational database. Today, people are still storing most of their data in relational databases. In some ways relational databases are beneficial, such as for storing simple data like text and numeric data.

3.3 Hybrid data structure

A hybrid GIS structure is comprised of spatial data files and non-spatial data files (cf. Figure 3.3a and 3.3b). The spatial data or geographic and attribute data are linked by a common key to the non-spatial data which are attributes of the feature. In the hybrid data structure, the maintenance of databases integrity, security, and reliability is complex [52, pp. 260-261]. However, there are advantages to using the hybrid data structure. One of these advantages is the ability to make use of a set of generalized dataset (e.g. raster data) in order to perform a faster-searching or browsing and sampling of the databases [35, pp. 11-12]. Spatial data and non-spatial data are managed in modules independently. Furthermore, while using a hybrid data structure, users can perform an analysis of the application without noticing the type of data (e.g. raster or vector data) that are processed.

In this way, according to [37] "nearly a standard are the well-known ESRI shape files that consist of at least a .dbf-File, a .shp-File and a .shx-File". Usually this type of file format has a topology. If not, the data must be manipulated and transformed by the use of designated software and tools.

Figure 3.3a Geometric shape (e.g. polyline). Shows an example of a spatial data file (e.g. geometric aspect of the roads).

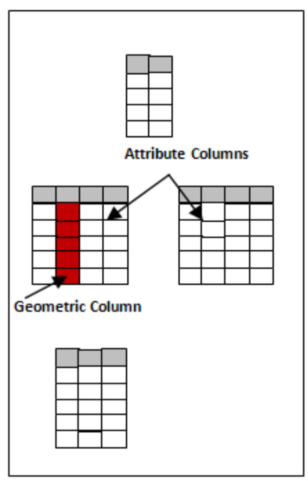


Figure 3.3b Feature and attribute tables. Shows an example of a non-spatial data file (e.g. name, type of road, max speed, etc.)

3.4 Object-oriented database

The object oriented database "era" started in the 1980s. Object oriented database system (OODBS) follows more or less closely, a standard proposed by the Object Database Management Group (ODMG) in 1993. This implies that OODBS not often depends on a unique model. The spatial and geographic objects are represented by objects in a homogeneous way in OODBS [40, pp. 100-107]. OODBS is basically another approach as an alternative to relational model paradigm which allows user to define data types with some characteristics. Some of the characteristics of the OODBS are object identities, object types, object classes, object methods or operations, object encapsulations and object inheritances.

Object identity denotes the unique identity of each object by using an identifier (e.g. *id, oid*) independent of their attributes values. The structure of one object and the operations that can be performed on it is related to the type of object. For example, one object can be of integer, string and or more complexes type such as geometric type. All the objects that have the same type are grouped into classes. The objects within a specific class have the same behavior and structure which have a set of methods (or operations) used for the objects of that class. A *class* implements abstract data type (ADTs) which is described in section 3.5. Thus, object encapsulation simply hides this implementation of ADTs and its structure from the outside world (e.g. users). Inheritance allows classes, and objects to share common properties. It also allows the definition of subclasses based on the existing classes. The term superclass or parent class is an upper class from which other classes (e.g. subclasses or child classes) are derived. The subclasses inheriting the properties of the superclass (cf. Figure 3.4a). The operations and methods of a class are inherited by the subclasses when it is created. Inheritance induces the idea of polymorphism which let the instances of a subclass be replaced by an instance of the superclass.

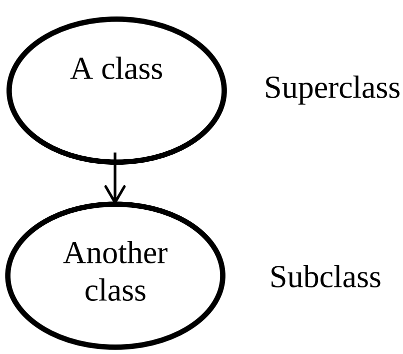


Figure 3.4a – Shows the flow direction of subclass inheritance of properties from superclass.

3.5 Extended relational database (object–relational database)

An extended relational database is used for handling spatial data, since a standard RDBMS does not contain geometric data types.

An abstract data type (ADT) is a data type that provides an extension to the operations of those which are consistent. It extends the structure query language (SQL) type system. One way to extend the capability of a relational database is to use abstract data types (ADT) [15, pp. 13-21]. These additional ADTs allow the control and management of the non-existing type in a standard database system for example, geometric and topological queries to the database. ADTs permit the user to be familiar with the detailed associated functions. Today ADTs are implemented in languages such as C language. In case of the geometric objects it is implemented and follows the standards of simple feature access to be used in a standard relational database. ADT is also used to define the columns type for tables in a database

3.6 Spatial database

One example of an extended relational database is a spatial database. Spatial data needs to be modeled somehow differently due to its complexity. Spatial databases are databases that allow the storage and retrieval of data related to objects in space, such as points, lines, and polygons. The spatial database, unlike traditional databases, is capable of process spatial data types, usually called geometry or feature [43]. Usually a spatial database is comprised of images and vector layers (e.g. land parcels, transportation, ecological regions, soils, etc) [42, pp.1-21]. Unlike traditional data types such as number and text, spatial data is much more complex. In this case the ADTs are geometric object types or point, linestring and polygon. For example, the ADT could be a combination of the type polygon and an association function such as the adjacent function. Polygons usually "*share the same boundary",* especially in a parcel. Therefore, the use of the adjacent function allows one to verify whether this condition is true.

The common geometric ADTs are points, lines and polygons. The ADTs are used for spatial data and they can also be treated as any other data type such as text, number, integer, etc (cf. Example 3.1). Similarly, ADTs and spatial operations or functions can be used to retrieve information by using "*extended*" SQL queries [15, pp. 15-17]. Two examples are: the point in a polygon test used to verify the existence of a feature object is in a specific place (e.g. a house or building in a neighborhood), and polygon area used to (e.g. measure the area of a park).

There are two levels that one should be aware of in terms of the ADTs. One is the logical level which most users are familiar with dealing basically with SQL-level. Level two is the physical level which deals with the x-coordinates and y-coordinates stored for points. This level is hidden (or encapsulated) from users since it is implemented by the creator of the database and it is not necessary for users to work with it.

```
CREATE TABLE roads (
        INTEGER      roadID       NOT NULL    PRIMARY KEY
        TEXT         name
        TEXT         road_type
        POLYLINE     geometry     NOT NULL
            :            :
            :            :
        )
```

Example 3.1 – Shows the use of polyline ADTs for spatial data in the road table.

There are operations defined on the ADT. Two examples of spatial operations on spatial ADT are:

1. Distance – determines the shortest distance between any two points in two geometries (cf. Figure 3.5a).
2. Intersection – compares two geometries returning a geometry that contains the points with the same dimensions in the geometries being compared (cf. Figure 3.5b).

In order to compute a distance between points and lines one should estimate and study the length of the line segment and the distance between the point and the two end point of the line segments [15, pp. 37-45].

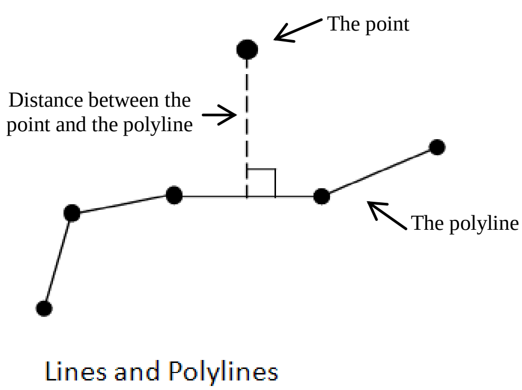


Figure 3.5a - Shows distance between a point and a line segment.

The Figure 3.5b show the operation on two sets denoted as set A and set B resulting in an intersection between them and points which belongs to both sets denoted as set C. Sometimes simple operation like this can help one to solve several issues in the geographical information system GIS query [15, pp. 1-12].

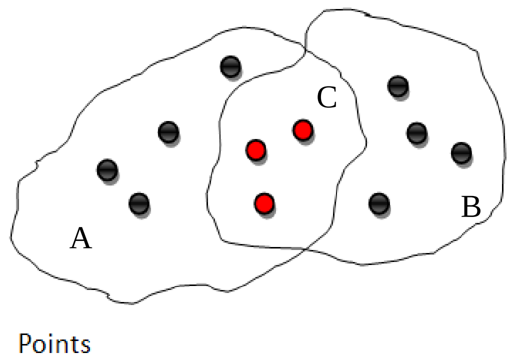


Figure 3.5b – Shows points and intersections

3.7 Database solutions for geographic data

Figure 3.6 shows four database solutions for GIS programs. Each type of these database solutions uses a different approach. Figure 3.6 i) shows the use of an "in house database system", which is the most commonly and used method today to store geometric data, while the attributes are stored in a standard relational database. In this manner, for the management of the geospatial data the use of hybrid, described in section 3.3, is essential. Figure 3.6ii shows the use a program on the top of the relational database, which adds the functionality of handling geometric in a form of an extended SQL. Figure 3.6iii shows the use of an interface in a form of an extended SQL and the geometry functionality has a direct interaction with the database. Figure 3.6iv shows the use of the "in house query language" approach and it also uses a spatial object-oriented database.

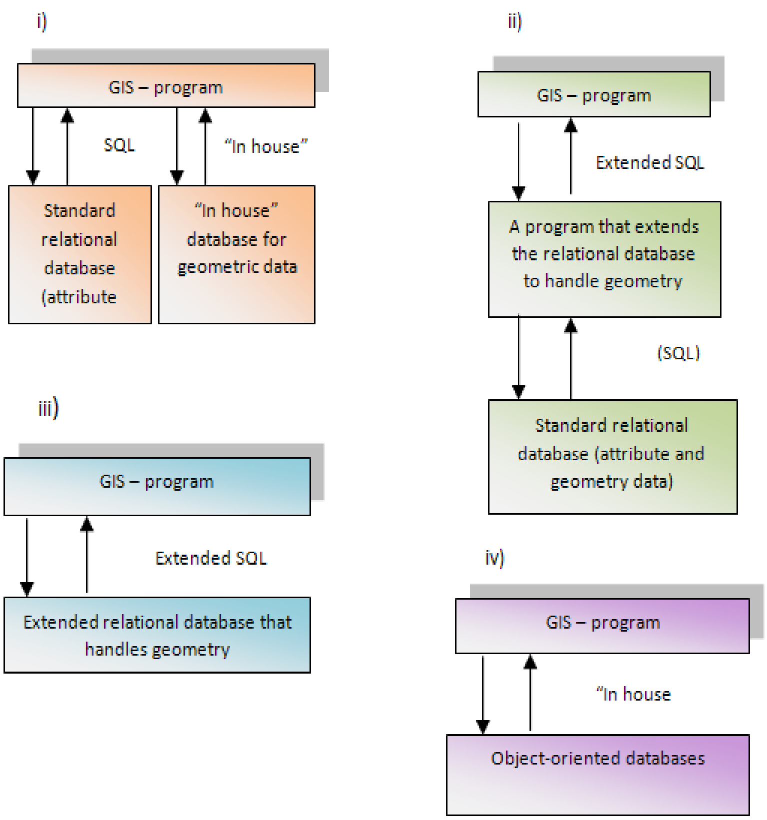


Figure 3.6 – Show four different solutions for a GIS program implementing and handling spatial data [15, pp. 14-15].

NOTE: *"In house" means non-standard format which also means it is proprietary.*

4. Example of geographic datasets

4.1 Types and usage of geographic data

Three commons types of geographic data are authority, commercial and volunteered. Authority data is important for decisions by the governments and authorities dealing with legal processes. Licensing geographic data implies evolvement of government's law, policies and institutions. The data are usually not accessible by the public. However, people need to have access to geographic data to become educate about the data and its legal procedure of usage in detail to avoid improper use of it. Democracy is important in a society but access to government information sometimes is a subject equal protection [24, pp. 19-30].

Commercial business collects sensitive geographic data in order to sustain their objectives. One of these objectives is to avoid the misuse of the geographic data. Commercial business generates a wide variety of products, information and services such custom data processing services in order to serve the private firms, public and the government. Some of the services provided by such commercial business are the Web-based services that support a variety of applications including in-vehicle navigation, location-based services, Web mapping, and asset tracking [24, pp. 31-38].

Volunteered geographic data are of the interest of many since it would be beneficial to not only to the public domain but also to government and private sectors. Several contributors of open geographic data source such as OpenStreetMap provide geographic data to anyone to use. However, some of these geographic data are under a license restriction accountable to government and public benefit [24, pp. 157-176].

What mainly differentiate these three types of geographic data examples are the accessibility by users. For example, previously, people had to pay for a license to use maps and geographic data. Today, projects such OSM makes it possible for anyone to create, use and develop projects (maps) while using their creativity, various application tools, and others components without those restrictions. There are also similarities between these geographic data types such as the usage, purpose, standards, etc.

4.2 OpenStreetMap (OSM)

OpenStreetMap (OSM) is an open source network of contributors for creation, development, and sharing of geographic data within that network and to the public. By using the geographic data from OSM one should agree upon under the terms and condition of an open source license. Thus, users can freely create, edit, modify, and share a variety of map data with one another under that open source license agreement. It is denoted as “Creative Commons Attribution-ShareAlike 2.0 licence” (CC-BY-SA). CC-BY-SA facilitates the sharing and remixing of geographic data by attributing the author or the licensor and share alike distributing the resulting works under the same or similar license. OSM offers this possibility for us however; there is still substantial work and development to be done. One can access, engage on, and contribute to enhance OSM geographic data by accessing its project website where several tasks are listed for geographic data developers.

The core elements of OSM project are its OSM’s website, technical infrastructure, contributors and developers. They are constantly creating, maintaining, and improving the OSM’s infrastructure. Consequently, it allows creation of maps, the interaction and increasing of people in this community of developers. There are demands for many aspects for OSM, however, the increase of tools development, access to data, use of application domains, software platforms and hardware devices is also a motivation to this community to continue to grow.

There are several contributors (people from different locations in the world) involved on projects such as OSM. One of the contributors, AND (Automotives Navigation Data), has donated the entire street map of Netherland to OSM project as an example. Netherland is considered one of the primarily countries having a complete street-level mapping. United States of America (U.S.A.) has contributed with comprehensive streets and highway coverage for the entire country [34]. OSM has demonstrated how this project has been increasingly successful, given that, “in 2006 Yahoo has given the OSM the right to use its satellite imagery web service to trace roads and other features” [34].

4.2.1 OpenStreetMap data model

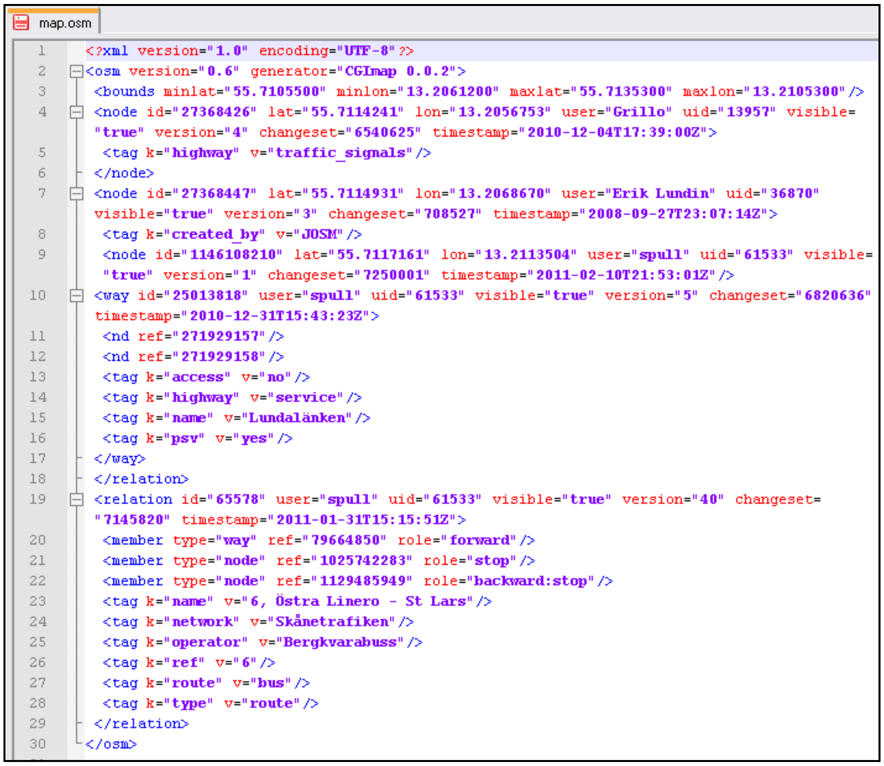
The basic geometric elements of the OSM data model are nodes, ways and relations and they have an arbitrary number of properties. However, in order to have a detailed description of those elements, annotations named tags may be added to these elements. For example, the highway tag is used for highways (e.g. roads and paths). These tags and features have to follow a certain set of recommendations in order to create, interpret, and display a base map (cf. Figure 4.1). Contributors can edit the OSM map features using the elements and input tags through the edit interface of the OSM site. These basic elements have equivalents in the internal OSM database as *data primitives* [34].

Figure 4.1 – Map of Municipality of Lund from OSM.

Data from OSM are frequently changing since there are many people contributing with new geographic data. So, one should keep in mind that a set of geographic data often is updated (cf. Figure 4.3).

OpenStreetMap (OSM) data are used for various purposes by different types of users. As an open source, OSM has contributors from all over the world creating a massive network of users. The OSM network allows one to access free geographic data from any part of the world. Using OSM ones can create maps, extract information from the dataset to enhance ones business directory or create own application such as own pedestrian route network service, combine OSM with another type of data, edit and create new products from it, etc. OSM data since the beginning of OSM project has been update frequently. In fact, there are many contributors such as GEOFABRIK in Germany which update OSM data daily.

Figure 4.2 shows a simple example of OSM file looks like. Some of the content of this file are the elements of the OSM dataset (e.g. nodes, ways and relations). As an example, nodes hold information like id, latitude, longitude, user, uid, if it visible or not, version, and timestamp. The tags are additional feature with information about the elements.



```
map.osm
<?xml version="1.0" encoding="UTF-8"?>
<osm version="0.6" generator="CGImap 0.0.2">
 <bounds minlat="55.7105500" minlon="13.2061200" maxlat="55.7135300" maxlon="13.2105300"/>
 <node id="27368426" lat="55.7114241" lon="13.2056753" user="Grillo" uid="13957" visible=
 "true" version="4" changeset="6540625" timestamp="2010-12-04T17:39:00Z">
  <tag k="highway" v="traffic_signals"/>
 </node>
 <node id="27368447" lat="55.7114931" lon="13.2068670" user="Erik Lundin" uid="36870"
 visible="true" version="3" changeset="708527" timestamp="2008-09-27T23:07:14Z">
  <tag k="created_by" v="JOSM"/>
  <node id="1146108210" lat="55.7117161" lon="13.2113504" user="spull" uid="61533" visible=
  "true" version="1" changeset="7250001" timestamp="2011-02-10T21:53:01Z"/>
 <way id="25013818" user="spull" uid="61533" visible="true" version="5" changeset="6820636"
 timestamp="2010-12-31T15:43:23Z">
  <nd ref="271929157"/>
  <nd ref="271929158"/>
  <tag k="access" v="no"/>
  <tag k="highway" v="service"/>
  <tag k="name" v="Lundalänken"/>
  <tag k="psv" v="yes"/>
 </way>
 </relation>
 <relation id="65578" user="spull" uid="61533" visible="true" version="40" changeset=
 "7145820" timestamp="2011-01-31T15:15:51Z">
  <member type="way" ref="79664850" role="forward"/>
  <member type="node" ref="1025742283" role="stop"/>
  <member type="node" ref="1129485949" role="backward:stop"/>
  <tag k="name" v="6, Östra Linero - St Lars"/>
  <tag k="network" v="Skånetrafiken"/>
  <tag k="operator" v="Bergkvarabuss"/>
  <tag k="ref" v="6"/>
  <tag k="route" v="bus"/>
  <tag k="type" v="route"/>
 </relation>
</osm>
```

Figure 4.2 – Shows how is an OSM file structure.

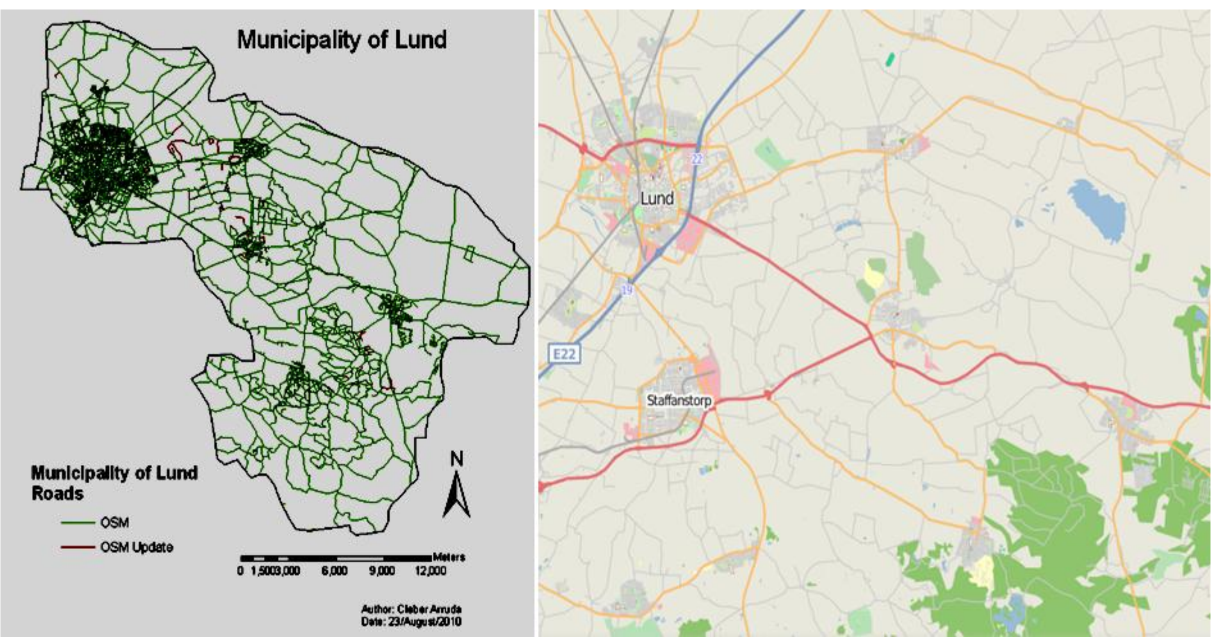


Figure 4.3 – OSM data municipality of Lund Map.

4.2.2 Editing OpenStreetMap data

Furthermore, users can manipulate data through features such as Java OpenStreetMap (JOSM), which is similar to the GIS application software and application programming interface (API). JOSM is a very useful and dynamic tool. It allows users to link OSM features to photo and audio notes, supports data conflict resolutions, and perhaps most importantly, it can be extended, using several independently developed plug-ins such as the Web mapping service (WMS) background imagery, Yahoo aerial imagery live recording of external GPS data, and data and tagging scheme validation tool [34].

4.3 Swedish national road database (sw. NVDB)

Swedish national road database (sw. NVDB) is a database that contains up-to-date information about roads and bicycle routes in Sweden. NVDB is used for applications such as navigation, transportation planning, traffic security, etc. The NVDB has information about the road network which describes not only geometrically but also topologically all the roads and some of bicycle routes in Sweden [47-48, pp. 5-6]. The roads are classified and represented as lines that connect by coordinated points. NVDB also contains information about the properties and use of the roads such as the width, speed limit, the road's name, and so on.

In order to have all this information about the Swedish national road database various contributors, such as the Swedish Transport Administration (Trafikverket), Swedish municipalities and county councils (SALAR), forestry companies (Skogsindustrierna) and the Swedish mapping, cadastral and land registration authority (Lantmäteriet) cooperate. The Swedish Transport Administration (Trafikverket) is responsible for the traffic on Swedish's roads, railways, sea (waterways) and flight (airways). Trafikverket is an essential contributor providing important information like traffic regulation, serving as one of the many important information services that Trafikverket provides to the continuation of its development and contribution to NVDB. The Swedish municipalities are within an association called the Swedish Association of Local Authorities and Regions (SALAR). SALAR is an additional contributor to NVDB, which is comprised of 290 (two hundred and ninety) municipalities and 20 (twenty) county which include the regions of Gotland, Halland, Västra Götaland and Skåne. They come together to provide and support services to each region and local municipality and to the Swedish government [45]. Swedish Forest Industries Federation (Skogsindustrierna) is another contributor to NVDB; it has several associated Swedish and European members companies dealing with the issues of the proper use of wood products. Finally, Lantmäteriet is the Swedish mapping, cadastral, and land registration authority providing basic geographic and land division information.

5. Network algorithms

5.1 Shortest path algorithms

The shortest path is one of the most important operations in a route network. Usually it is used for estimating the distance travelled by car, foot, bicycle, etc in order to determine the shortest route. Shortest path algorithms are usually based on the Euclidean distance and sometimes in terms of time travel [52, pp. 214 - 217]. When a pedestrian determines a route to navigate, there is an array of elements to consider, such as road type (e.g. motor road, sidewalk, pedestrian path) with its specific distance value or travel time (cf. Figure 7.1). Note that this distance can also be called “weight”.

This study concentrates on the development of a pedestrian route network service to provide a convenient, short and safe route for pedestrians. The essential component of these aspects within a network is the computation of the shortest path, using shortest path algorithms. There are several shortest path algorithms such as the A* (A star) algorithm, the shooting star algorithm and Dijkstra’s algorithm. Dijkstra’s algorithm is the most well-known shortest path algorithm; other more advanced algorithms are often an extension of this algorithm.

Classification and measurement of the shortest path must be executed based on a cost or weight ratio that is determined by the travel time or by the distance between two points [6, pp.41-47]. When it comes to structure routing networks applications, a vector map with a representation of the routes in a graph-based is also an essential factor (cf. Figure 7.1).

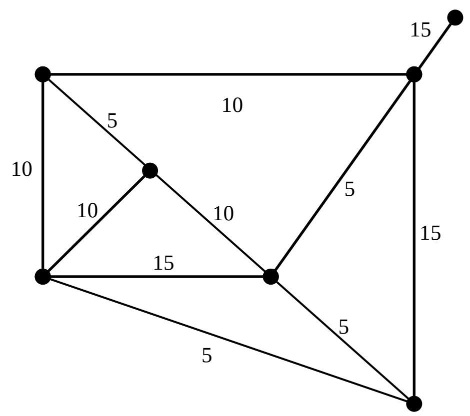


Figure 7.1 – Route Network with distance cost or weight.

5.2 Dijkstra's algorithm

Dijkstra's algorithm is one the most popular shortest path algorithms. Dijkstra's algorithm estimates the shortest path from a node visiting all the other nodes in the network. Thus, Dijkstra's algorithm is categorized as a single source shortest path [52, pp. 214-217]. In order to be functional Dijkstra's algorithm must use a network with no negative distance [15, pp. 9-10]. In case the distances are negative then the shortest path is unable to be found.

At the beginning of the iteration Dijkstra set all the target weight to infinity except the start node and all adjacent nodes to the start node. Then it starts to visit all the nodes starting by the one which has the lowest weight of the adjacent and unvisited nodes. The iteration continues and the nodes are updated based on the comparison of the weight of adjacent and unvisited nodes. This process goes until all nodes are visited. Dijkstra algorithm keeps the lowest weight of the vertex being processed from the source in an array. The shortest weight of the source to itself is zero.

Besides the rules the Dijkstra's algorithm requires a source and target ID, the id attribute and a cost column are needed in order to estimate the shortest path. Furthermore, in case of the specification of a reverse cost, Dijkstra's algorithms requires an extra column, which will also has a cost (reverse cost).

5.3 A-star (A*) algorithm

A* (pronounced "Ay star ") is another shortest path algorithm which can be viewed as an extension to Dijkstra's algorithm. A-star (A*) algorithm uses a heuristic approach to find the shortest path from a determined starting node (or the source) to a destination (end node or the target). During its interactions, the A* algorithm looks for the closest nodes from the starting point to the end node. The A* algorithm uses the Euclidean distance when estimating the distance between two points. Thus, it is also convenient for geospatial information. There are also problems that the A* algorithm presents: it is the "worse-case" complexity (or the $O(n^2)$). On the other hand, A* is better than Dijkstra's algorithms in the average-case time complexity [14, pp. 214-217].

One of the approaches used by A* is a suitable evaluation function in order to compute goal-directed shortest path. The algorithm sometimes requires that all possible start locations to all possible destination exist to compute the shortest path. This approach is called all-pair shortest path. The algorithm conserves an open list that tracks the nodes that must be visited and a closed list that tracks the visited nodes.

Each algorithm has its own characteristics and usage. For example, A* is usually used on in the industry of transportation. The reason is because it demonstrates a decrease in computational time and flexibility compared with other algorithms.

5.4 Shooting star algorithm

Another shortest path algorithm is called "Shooting Star". Unlike Dijkstra's algorithm and A* algorithm, the shooting star algorithm uses a link to link approach to perform the routing search instead of a vertex. Problems like "parallel links" can be address and solved by the use of the shooting star algorithm. However, it has similarities to the other algorithms, with regarding to search through links close to the target, and its use of heuristic method like A*.

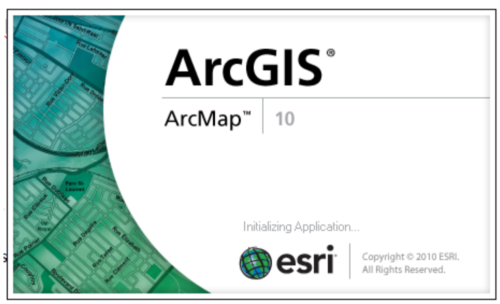
The shooting star algorithm works based on the link-to-link passage cost. This is one of the attribute of the adjacent edges structure or the cost of the line graph, which represents any limitations or penalties. Based on this the algorithm can use the same iteration concept of A-star [12]. Shooting star is a new function and has new extensions of it. For example, shooting_star_sp and shooting_star_smart which are available now on the pgRouting and github website. The shooting_start_smart is the newest function released Feb. 2011 and there is not any official documentation about it.

6. Description of application software and database system

6.1 Introduction

In this study we use several application software and tools in order to organize, view, and manipulate the data acquired. The application software and tools described in this chapter are used and considered essential for this type of study.

6.2 ArcGIS



ArcMap is an ESRI's GIS software which has tools and functionalities that provide capabilities such as data collection, data import, editing, restructuring, transformation, display, query, analysis, routing, geostatistical and spatial analysis among others [25, pp. 167-172].

After the collection of the OSM and NVDB data, this software is used to view the data and perform changes in a coordinate system, edit, and make an analysis of the data (cf. Figure 5.2).

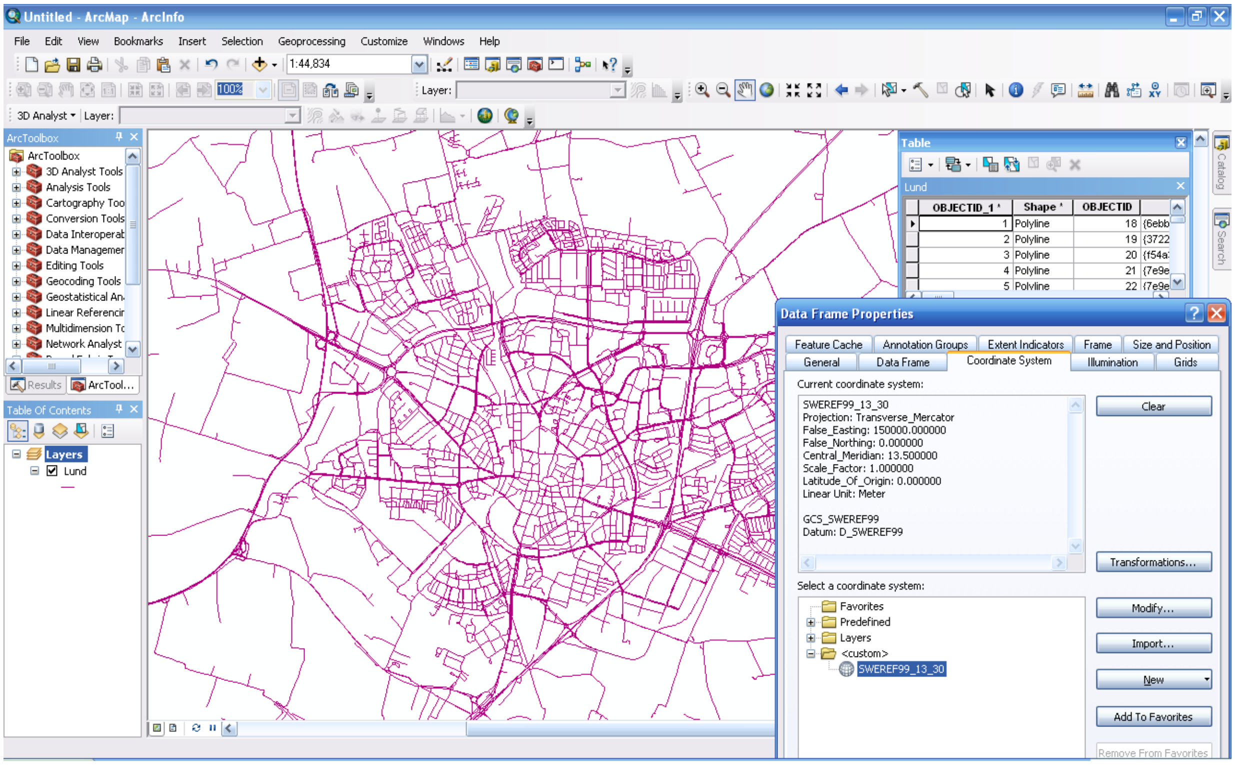
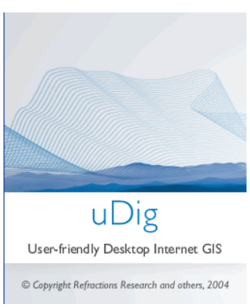


Figure 5.2 Screenshot of ESRI ArcMap. The windows on the right in front of ArcMap shows the table with NVDB's pertaining data and its attributes; the windows in front of the table shows the data frame properties where the coordinate systems can be changed.

6.3 uDig



uDig is an open source program under the license Lesser General Public License (LGPL). It provides support in desktop GIS, data access, editing, and viewing of data. In this project analysis, uDig is used to show the routing results derived from the pgrouting (cf. Figure 10.10). uDig can be downloaded from http://udig.refractions.net/download/

6.4 FME – Feature Manipulation Engine

Feature Manipulation Engine (FME) is a GIS program able to extract, transform and load geographic data (spatial ETL). It performs these operations and a variety of data interoperability, challenges which one often encounters while managing proprietary and evolving data formats, adapts to new schemas and lack of standards and difficulties accessing, restructures, integrates, and distributes data [11]. Geographic Information Systems (GIS) have been used to display, manage, and analyze geographic data.

Usually spatial data comes from different sources. It implicates a variety of formats that need to be translated, transformed, and integrated in order to be distributed and used. FME is a flexible and powerful platform which offers those functionalities. The use of FME in this study is to tailor the data. This is vital for creating a PRNS since topology manipulation, projection, and viewing of the data (cf. Figure 5.4) can be accomplished.

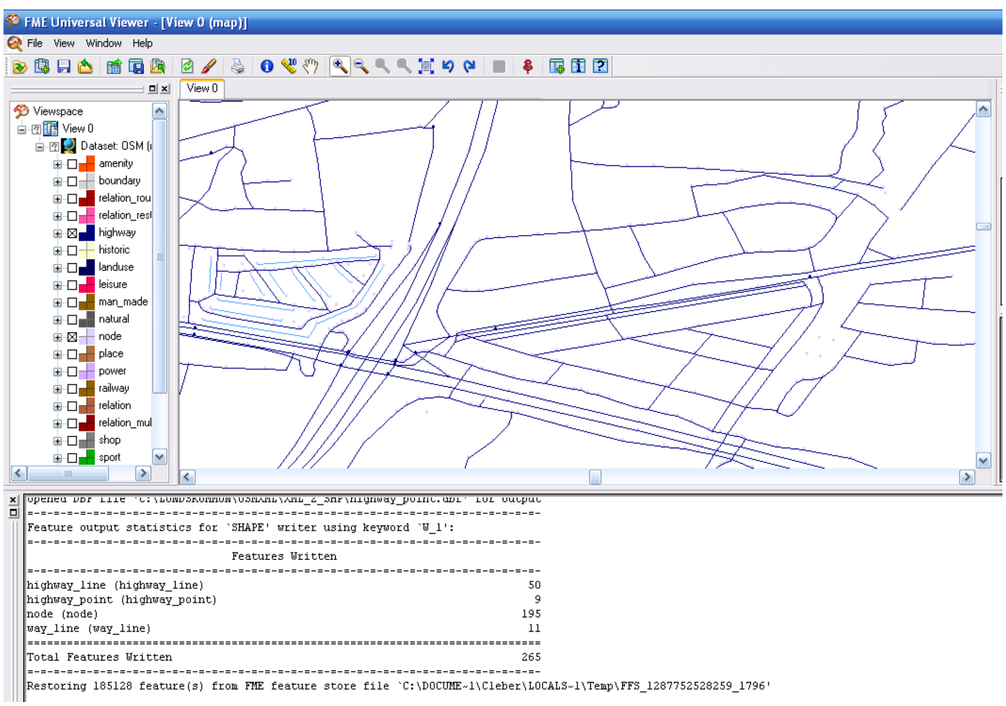


Figure 5.4 Sample of FME Universal Viewer showing the loaded and building process of OSM data.

6.5 PostgreSQL

PostgreSQL or just Postgres is an example of open source software. Postgres is object-relational database management system (ORDBMS). It supports most operating system (e.g. Linux, UNIX, Windows Mac OS X etc) but has its limitations when it comes to table size, row size, field size, and number of columns (depending on the column type). However, this software is well known and one of the most used when it comes to object relational databases management system. It has many technical features developed and provided by many open source developers, having a wide potential to continue its expansion and improvement.

- PostgreSQL– an open source object-relational database management system downloadable from http://www.postgresql.org/download/

6.5.1 PostGIS

One example of a spatial database is PostGIS. It is an open source spatial database based on PostgreSQL. PostGIS follows the Simple Features for SQL specification from the Open Geospatial Consortium [32]. PostGIS allows the manipulation of geometry columns providing spatial information of geometries such as points, linestrings, polygons, multipoints, multilinestrings, multipolygons and geometrycollections. By the use of AddGeometryColumn function, one of the several functions of PostGIS, these geometry columns can be created within the database. PostGIS uses spatial operators for determining geospatial measurements such as areas, distances, lengths and perimeters. These spatial operators also allows one to determine geospatial set of operations, like union, difference, symmetric difference and buffers. In terms of access PostGIS has its R-tree spatial indexes for high speed spatial querying performance. These indexes are supported to provide high performance query plans for spatial and non-spatial queries [39] [21].

PostGIS is basically used for the storage and manipulation of spatial objects like any other object in the database. Pedestrian routes navigation is associated with geometric locations. The information gathered from the route network is stored into the PostGIS database, which allows the manipulation of the spatial data. For example, the storage and retrieval of data. PostGIS works as a spatial database management system along with pgAdmin database administration as front-end clients (section 5.6).

Usually there are several tables storing information in a PostGIS database, where the data tables are separated having information with its own name feature. PostGIS also supports the use of SQL queries statements which is facilitated by the use of binary spatial relations.

- PostGIS – a spatial database extension to PostgreSQL downloadable from http://postgis.refractions.net/download/

6.5.2 pgAdmin

pgAdmin is an open source administration and development platform for PostgreSQL. pgAdmin runs in various type of platform commercial and non-commercial platform such as Linux, Mac OSX, FreeBSD, Solaris and Windows. pgAdmin is also friendly graphical user interface (UI) which allows one to write simple or complex SQL queries in order to manipulate data in the database. Consequently, user can create and manage tables, develop complex database and queries to store data to and retrieve data from the database. pgAdmin is an application that supports all the PostgreSQL features and PostGIS and in this way it make easy the administration of the database[36].

6.5.3 pgRouting

pgRouting which is an extension of PostGIS, adds routing functionality to the PostGIS/PostgreSQL. pgRouting is a further development of pgDijkstra (by Camptocamp SA). pgRouting was extended by Orkney Inc., and is currently developed and maintained by Georepublic under the General Public License version 2 (GPLv2). pgRouting works with geodata of type linestring or multilinestring (e.g. shape files or OSM data) [37, 21, pp. 3-4].

pgRouting provides functions for [21, pp. 3-4]:

- shortest path Dijkstra: routing algorithm without heuristics.
- shortest path A-star(A*): routing for large datasets (with heuristics).
- shortest path shooting-star: routing with turn restrictions for real road networks (with heuristics).

- traveling salesperson problem (TSP) – based on ordering points and using straight line (euclidean) distance between nodes.
- driving distance calculation (isolines).

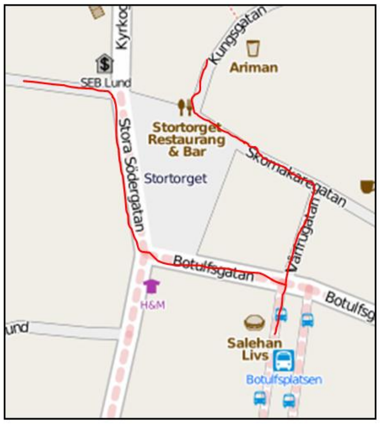


Figure 12.7 – Typical trajectory to reach bus terminal Botulfsplatsen (map: Stortorget in Lund view from OSM).

Besides these functions, pgRouting offers several advantages with regards to the database routing approach [37, 21, pp. 3-4]. They are as follows:

- Accessible by multiple clients such as PC's or mobile devices through java database connectivity (JDBC), open database connectivity (ODBC), or by directly using Procedural Language or PostgreSQL Structured Query Language.
- Uses PostGIS for its geographic data format, which in turn uses Open Geospatial Consortium (OGC) data format such as the well known text (WKT) and well known binary (WKB).
- Accessibility to open source software (e.g. qGIS and uDig) in order to modify the data and attributes.
- While routing, pre-calculation is not required; the data is changed directly.
- The "cost or weight" parameter can be dynamically calculated through SQL. SQL can retrieve the value to be estimated from different tables or fields.

This study owes much to pgRouting.org project, which made a tremendous contribution to the field of developing routing network services by providing documentation and resource for creation of routing network services.

pgRouting requires several parameters in order to perform the routing process. Thus, the selection of attributes is very important in order to perform the pgRouting functionality, especially if one intends to use different approaches (e.g. different algorithms such as Dijkstra, A*(Star) and Shooting Star), because each algorithm also uses different parameters from the dataset to find the shortest path, as described in section 5.1.

6.6 OpenLayers

OpenLayers is an open source software used to display maps in most common of today's web browser such as firefox, google chrome, internet explorer, etc. It is a JavaScript library which has not service side dependencies. OpenLayers develops, maintains, and shares several examples of functions and code snippets that user can use on their application. These functions and code snippets are available on OpenLayers website http://docs.openlayers.org. Today, several web map service such as Bing and Google maps implements the APIs for the web-based geographic application. OpenLayers uses such similar implementation by using JavaScript API.

7. Programming environments

Programming languages are created and classified in various ways. These languages, especially high level languages, need to be translated (or compiled) into machine language in order to be interpreted and executed by a computer. Programming languages provide means of computing devices communication and are set in a complex context of interaction [44, pp 1-12].

In this project we have only used the program languages JavaScript and PHP described below in section 7.1 and section 7.2. In order to improve the accessibility of web-based GIS applications and the representation of maps, JavaScript and PHP are essential factor in this project.

7.1 JavaScript

JavaScript is an interpreted language. The meaning of interpreted in this case is that JavaScript does not need to be compiled into machine code like other languages such as C/C++ language. JavaScript language can be directly interpreted by a JavaScript interpreter incorporated in a web browser. Therefore, JavaScript is a language used to make a dynamic interaction between applications and other languages, such as HTML. JavaScript is also a much easier type of language to write than e.g. C/C++, but it becomes complex when it comes to correctness of the code.

Using JavaScript, one can create an external document of the .js type or extension, or simply integrate it in another document like an HTML or PHP document. JavaScript is object-based, allowing interaction with the object model, also facilitating events handling several activities to happen. JavaScript permits the use of basic data types and variable declaration. Furthermore, its functions enable the representation of graphic elements such as lines, ellipses, polygons etc.

The use of JavaScript among other elements enables the employment of raster and vector representations, using maps and their geographic objects [1, pp. 490-494].

7.2 PHP

PHP originally meant *personal home page*. However, it now stands for *Hypertext Preprocessor* a recursive acronym. PHP is a HTML embedded scripting language mostly used for web development. One of the interesting characteristics of PHP is that it can be used to access database such as MySQL and PostGIS. Another example of use of PHP is the combination of it with AJAX in order to get and exchange information with a particular server [38]. PHP is used in this study to access the PostGIS database and post the contents retrieved from the database to the web map server.

8. Pedestrian route network service (PRNS)

Today there is a variety of services providing assistance for pedestrians through navigation systems, interactive maps online, and other navigation technology. These services provide a fast way for pedestrians to find their route and retrieve maps either online or with a mobile device [22, pp. 395-400]. However, those services also have their weakness and often lack of maintenance. The use and combination of the traditional map with other technological facilities are still essential and have their advantages [30, pp. 1853-1861].

Besides open source data, there are several routing or online routing services for different purposes. For example, the OpenRouteService (section 12.2) and CloudeMade (section 12.3). However, one very interesting program used to create and manipulate routing service is pgRouting (section 11.4) where users can integrate their own data (e.g. downloaded from OSM) and provide routing functionality while using the PostGIS database or PostgreSQL.

There are several factors that influence the choice of a particular route by a pedestrian. The distance is one of these factors, related to network geometry. Perhaps some pedestrians would choose a different route based on the geometry or topology presented from point A(departure) to point B(destination); this route is comprised of nodes and edges in a route network (cf. Figure 2.2), based on the shortest distance. In contrast, a driver would not be satisfied with the same path. Since pedestrians often prefer the shortest path to navigate, a pedestrian route network system is not satisfactory for a driver. Furthermore, the route choice made by the pedestrian is, of course, also determined by the desire of a specific path, the aspect of the surrounding environment, and the feeling of safety and security. Naturally, pedestrians choose their way by instinct and knowledge of the route. However, they are not aware of so many factors that are fundamental in their spatio-temporal activities [28, pp. 1-14].

The reason for using only roads, streets, and bicycle paths (shared with pedestrians) for the development of the PRNS is that there is a lack or limitation of pedestrian route network data from providers such as OSM project, Swedish national road database (NVDB) and the municipality of Lund. The data from these providers are navigation networks that have its attributes and descriptions, which help to determine the specific route for the PRNS. For instance, a road characterized as a highway is part of the network, but it has other attributes, such as speed, which is of no interest for a PRNS. There are different ways to determine a limitation or preference for different types of roads. Weight cost is another approach to take into consideration in this project. Using data from e.g. OSM, two tables can be generated (cf. Tables 8.1 and 8.2) and used to estimate the weight cost of each type of road in the network [21, pp. 23-26].

Table 8.1 - Type of Road

type_id	type_name
1	Highway
2	Bicycleway
3	Road
4	Junction
5	Walkway

Table 8.2 – Road Classes

class_id	type_id	class_name	cost
301	2	Lane	
302	3	Living_street	
303	3	Motor_way	
304	1	Bridleway	
207	5	Pedestrian	

The cost column can be updated by the use of SQL queries or advance routing queries with a designated values for each class_id into the cost column. These can be done by the use of the function UPDATE query. For a fast and better performance, an INDEX query can be created. The use of various tools such as FME, PostGIS, and pgRouting, among others, is also used to help developing the PRNS.

The completeness of the routing process is done by using the data from the providers mentioned above. The data differ from one another; however, the values, limitations, and attributes of them can be analyzed and compared after a routing process. Based on the data from each source, there can be similar or different results in terms of search, speed, accuracy of data, correct and best path, etc. The analysis and results based on the data is described in detail in Part II, Chapters 9 and 10.

Part II: Case studies

9. Method and preparation

9.1 Introduction

The method and preparation chapter describes the study area which the pedestrian route network service (PRNS) is implemented. The chapter is divided into sections which describe the workflow of the methodology used to develop the PRNS. It includes the downloading, converting, manipulating and viewing of the data.

The case studies performed in this project were discussed at several meetings with employees at the municipality of Lund. The first case study is to evaluate and compare the data. The second case study is to compare route network services.

9.2 Define the study area (municipality of Lund)

The case study for a pedestrian route network service is performed in the municipality of Lund (cf. Figure 9.1). Lund has 24.99 km^2 with approximately 109 000 inhabitants. Lund is a student city with a large population of university and high school students, as well daycare children navigating daily to schools. Lund located approximately at latitude 55° 42' and longitude 13° 12'

9.3 Coordinate system projection and conversion of dataset

Another crucial aspect of this study is the study of the coordinate system also denoted as spatial reference system. Coordinate system is usually used to assign coordinate to a location, and allows one to interpret a set of coordinates as a representation positions on the Earth (real world space).

The municipality of Lund is located in the Scania area (zone). The default reference system for the OSM dataset acquired from the OSM website is WGS84 and it has to be converted and projected from its original reference system to SWEREF 99 then to SWEREF 99 13.30. The reason for this is that Sweden is divided into several reference system zones (using different central meridians in Transverse Mercator projection) and Lund is on the SWEREF 99 13.30 zones [7].

Another step of the data preparation is the conversion of the shape file, obtained from the previous conversion and manipulation of the dataset to Structured Query Language (.SQL) files. These SQL files are used to add data into the PostGIS database by the execution or running of an SQL command. This command dumps or deposits the data converted from the shapes files into SQL as attributes into the database with its pertaining information, as shown in Table 10.1 and 10.2.

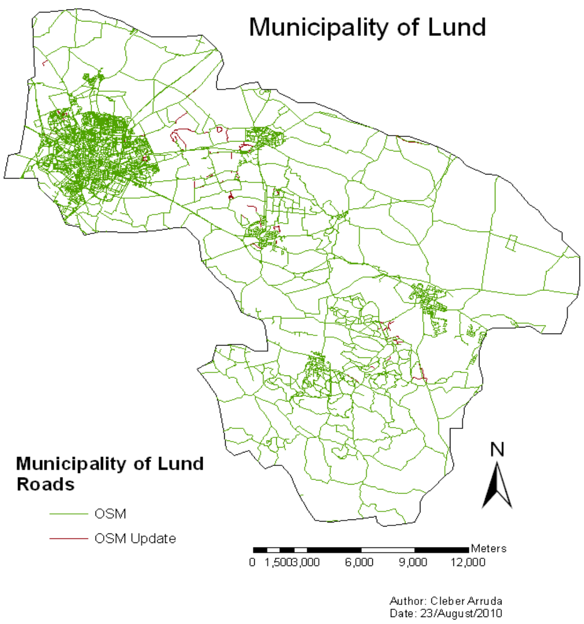


Figure 9.1 – Map of Municipality of Lund based on data from both tables in Figure 10.2.

9.4 Workflow of the methodology

There are several steps that one must perform including e.g. downloading and installing software applications and tools described in Chapter 6, preparing the data and validating them before the implementation of a routing process with pgRouting. The following steps described in the workflow methodology (cf. Figure 9.2) are:

- The first step is to download or obtain the datasets described in section 10.3.1 and 10.3.2.

- The second steps are conversion and projection of the dataset (section 9.3).
- The third steps accessing and manipulation.
- The fourth and last step is viewing or visualizing the data and maps.

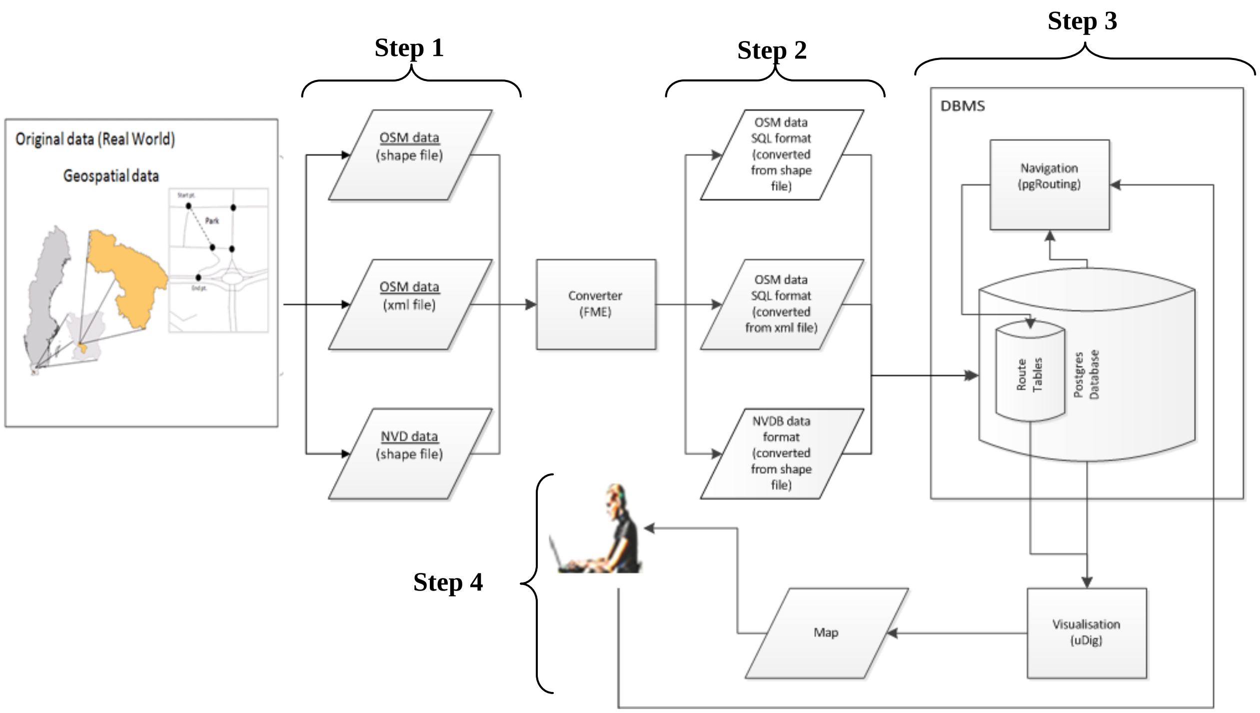


Figure 9.2 shows the workflow of the methodology.

9.4.1 Downloading the dataset

In step 1 the dataset from NVDB and OSM were downloaded or obtained from either their providers or website (sections 10.2.1 and 10.2.2). Then they were tailored (cf. Figure 9.2) by the use of ArcGIS and other software applications in order to get the correct topology for the road network. The OSM has several elements however, as described in section 4.1.1 the ones of interest are nodes, ways, and relations which can also be downloaded individually from other elements. The features of interest for this analysis is the roads, bicycle and pedestrian paths, and others elements on the network used by pedestrians navigation. Tags are additional feature which complements each element by providing information about it (section 4.1.1). In other words, all links (routes) in the network which can be used by pedestrians for navigation must be downloaded with their pertaining attributes to be used for this study and the development of the PRNS.

9.4.2 Accessing and manipulating the data using PostGIS DBMS

In step 3 (Figure 9.2) the PostGIS database management systems (DBMS) allows the interaction of the users with the database allowing the manipulation and retrieve of information from the database. One has to use the Structured Query Language (SQL) statements and the PostGIS built in function in order to accomplish it.

9.4.3 Viewing the data and maps

In step 4 (Figure 9.2) we can view and perform some analysis of the dataset and the data retrieved from the database. The viewing of the road network maps is not possible to perform in PostGIS. However, with the help and use of some applications software such as uDIG described in section 6.3 or ArcGIS described in section 6.2 the visualization of the maps are possible.

10. Case study I: Evaluation of network data

A main issue to create a pedestrian route network service is to use an appropriate road dataset. This case study dedicates to the evaluation of the datasets OpenStreetMap (OSM) and the Swedish National Road Database (NVDB) with respect to pedestrian navigation. Especially the comparison is made in the perspective of creating an own pedestrian route network service (PRNS) using pgRouting (cf. Chapter 4). The dataset analysis is done in collaboration with Eduard Mikayelyan (a Geomatics student at University of Lund).

10.1 Pedestrian network

Most of the route networks within a city are shared between motor vehicles, bicycles, and pedestrians. Identifying pedestrian network in a 2D map representation is complex if one does not use specific tools and applications software. The OpenStreetMap (OSM) dataset, roads are of several types. Figure 10.1a shows three different types, cycleway (exclusively for bicycles), footway (designated to pedestrians including path such as walking track and gravel path), and pedestrian (exclusively for pedestrians in shopping plazas areas). Tags always can be added to these elements in order to expand its classification. For example, for pedestrian the tag area can be added by setting area=yes and for footway the tag bicycle can be added by setting bicycle=yes. Notice that the two first maps of the municipality of Lund are from different area. However, the map on the left and the map below which is also from OSM are from the same area.

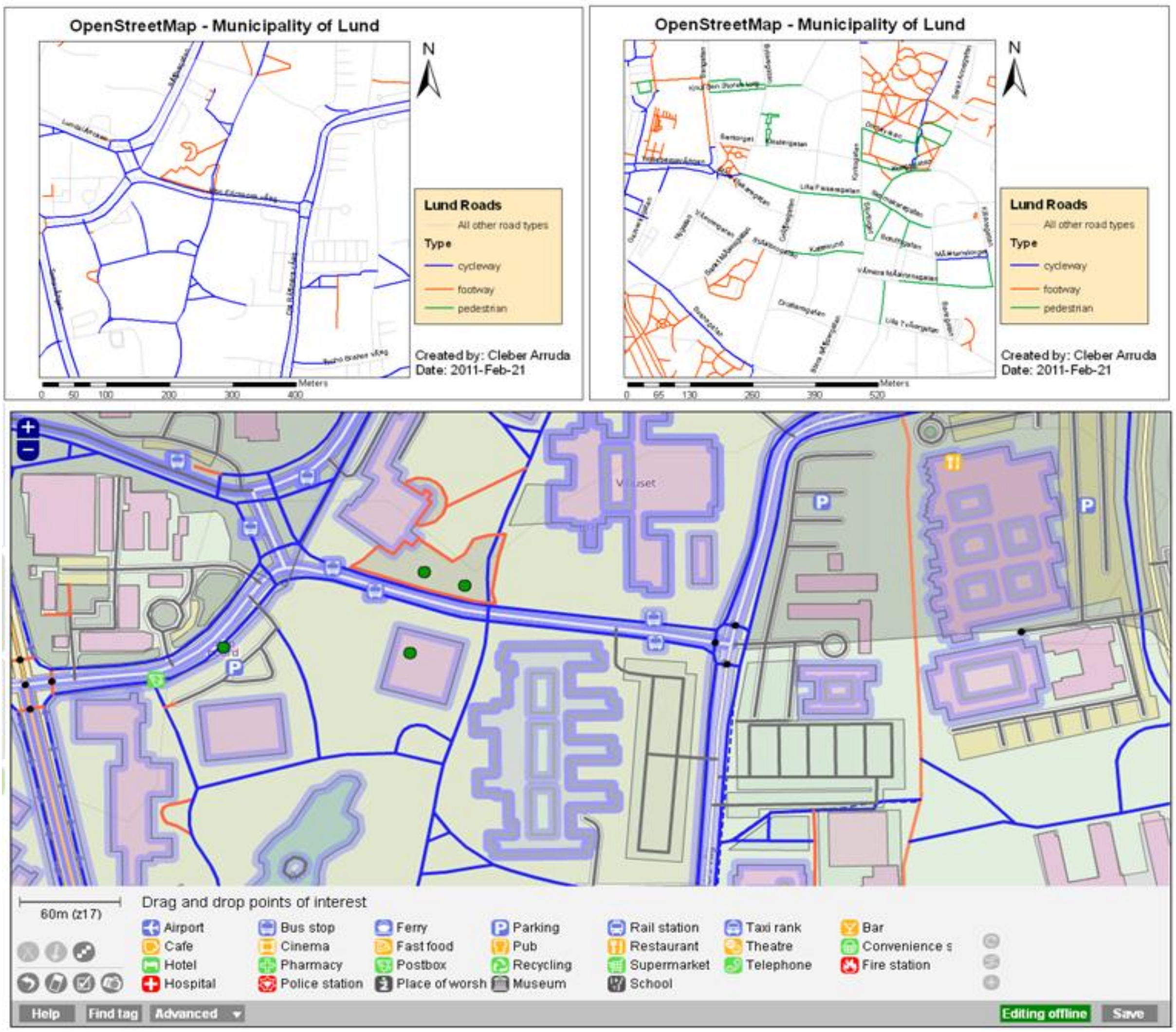


Figure 10.1a – Map from OpenStreetMap dataset showing roads network of municipality of Lund.

10.2 Creating geographic databases

In this study we use geographic data from OpenStreetMap (OSM) (section 4.2) and Swedish national road database (sw. NVDB) (section 4.2). One example of a geographic database table is illustrated d by Figure 10.2. This geographic database table contains attributes which differentiate it from a traditional relational database. For example, the field Shape where the geometric type is determined. However, other relevant fields such as the field type where the type of the road is classified and the field max speed where the speed of the roads are specified are crucial for the implementation of the pedestrian network service (PRNS).

Attributes of LKroads_Clip

FID	Shape *	osm_id	name	ref	type	oneway	bridge	maxspeed
0	Polyline	4448972			motorway	1	1	0
1	Polyline	4453888		E22	motorway	1	0	0
2	Polyline	4453934		E22	motorway	1	0	0
3	Polyline	4454192	VÃ¤varegatan		residential	0	0	0
4	Polyline	4454193			residential	0	0	0
5	Polyline	4454194			service	0	0	0
6	Polyline	4454195			footway	0	0	0
7	Polyline	4454196			footway	0	1	0
8	Polyline	4454197			cycleway	0	0	0
9	Polyline	4454198			footway	0	0	0
10	Polyline	4454199			footway	0	0	0
11	Polyline	4454201			cycleway	0	0	0
12	Polyline	4454202			steps	0	0	0
13	Polyline	4454204			tertiary	0	0	0
14	Polyline	4454205	Bangatan		tertiary	0	0	0
15	Polyline	4454207	Bantorget		residential	0	0	0
16	Polyline	4454209	TrollebergsvÃ¤gen		tertiary	0	0	0
17	Polyline	4454210	Kung Oscars vÃ¤g		secondary	0	0	50
18	Polyline	4454211	GetingevÃ¤gen		secondary	0	0	0
19	Polyline	4454214	TornavÃ¤gen		secondary	0	0	50
20	Polyline	4454222	KlostergÃ¥rdsvÃ¤gen		tertiary	0	0	50

Figure 10.2 – Sample of geographic database table from OSM dataset.

10.2.1 Downloading OpenStreetMap (OSM) dataset

There are several ways to download data from the OpenStreetMap (OSM) website or its associate contributors. The dataset of OSM is available on its website. Users can obtain the dataset in several data formats such as .osm files, .shp files and .pbf (protobuf binary format). The data is usually in the format of eXtensible Markup Language (XML). Data can be downloaded partially (e.g. in a specific region in the world, a city, a country etc.). However, one can download the whole world as a dataset. The OSM data is available for commercial and public organizations.

Small regions or areas can be obtained from the OSM website using an embedded export tool which facilitates user interaction with the map. The user can select the desired area by drawing a square box and export or download it as a map.osm file (cf. Figure 10.3). However, the download area is limited to a specific size. It should contain a certain amount of node about 50000 nodes otherwise one has to use or request a smaller area, or use planet.osm site in order to download larger areas. World Wide Web gets (WGET) is another approach in order download data. The user uses the bounding box (bbox) specifying the minimum latitude and longitude as well the maximum latitude and longitude of the desired area. For example, a wget bbox request statement would look like this:

wget -O map.osm http://xapi.openstreetmap.org/api/0.6/map?bbox=11.54,48.14,11.543,48.145

It can be either executed in an ms-dos command or in a web browser address bar such as Mozilla Firefox. The target location for this study is the Scania area in Sweden specifically the municipality of Lund (cf. Figure 10.3).

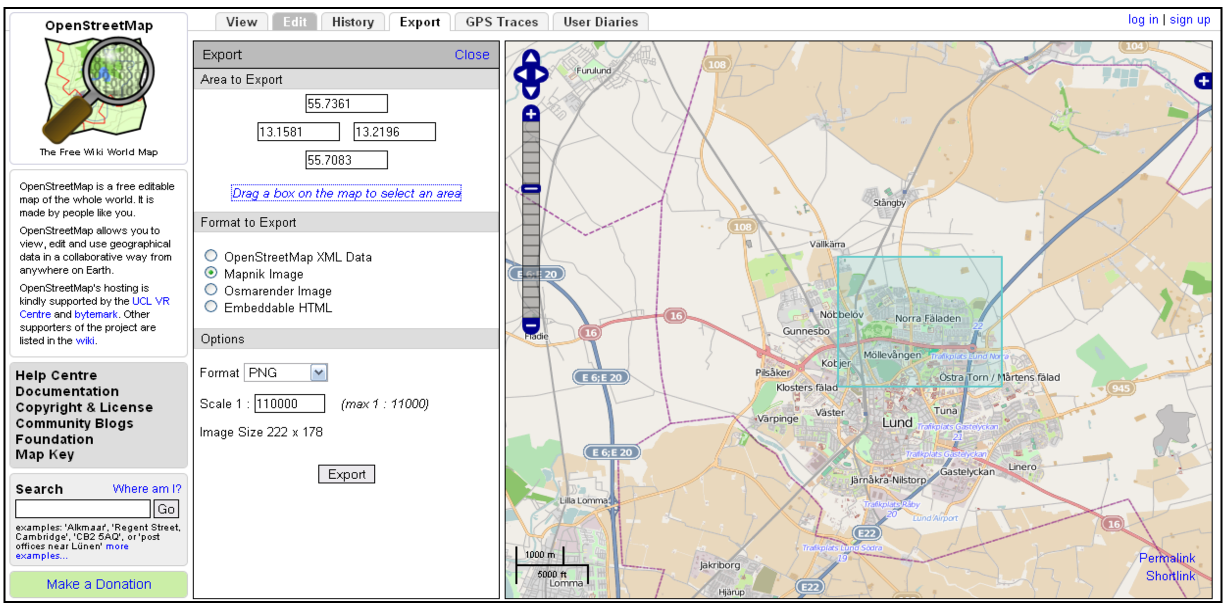


Figure 10.3 – Shows OSM export tools using a square box to download part of municipality of Lund area.

Another method used to download OSM dataset is by visiting sites of contributors and providers of OSM such as GEOFABRIK, CloudeMade, and Maquest where you can access OSM and download it in various format. They are some of the most visited sites and used to download data for this thesis. One of the reasons that GEOFABRIK is popular when it comes to download OSM dataset, it is because it update the data daily. The datasets are available at their website (e.g. GEOFABRIK – http://download.geofabrik.de. The datasets can be downloaded as XML OSM data, bzip2 compressed, protobuf binary format, ESRI shape file (EPSG: 4326) in zipped format [13]. The dataset can be download from these contributors and provider free of cost.

10.2.2 Downloading Swedish national road database (NVDB) dataset

NVDB data are also available for commercial and public organizations. NVDB data can be viewed for free of charge through NVDB on the internet. However, it required user name and password to access and view the data. In order to download the NVDB data one has to agree and sign in a contract for use of Lastkajen. Lastkajen is an application program that allows a user to access road data from NVDB. The data are in four different formats as follow:

- .nvdb – an internal data format by NVDB that requires the use of "Slussen", which is basically a platform application that allows one to use the NVDB dataset format.
- .xml – format based on the SS637004, a Swedish standard, which requires import functionality based on the standard's xml schema.
- .shp – shape file geospatial vector data (digital map) format used for GIS software. It is developed and regulated by Environmental Systems Research Institute (ESRI). It supports point, line, and area elements [53, pp. 174-178].
- .mdb – personal geodatabase that allows one to deal with complex relationships, also in ESRI environment.

The NVDB data used in this study were provided by the municipality Lund (one of NVDB's contributors).

10.3 Comparison of attribute data

During the analysis one of the important aspects of the datasets is the attribute tables (cf. Tables 10.1 and 10.2 below). It contains the attributes and information to compare OSM and NVDB for the PRNS. Looking at the table one can study the importance of each attribute and eliminate or include some attributes in order to fulfill the requirements to develop a better application.

10.3.1 OpenStreetMap (OSM)

In this case the attribute “TYPE” from Table 10.1 has some unclassified roads but one of the unclassified roads, “FID” 1192 has a “MAXSPEED” of 30. Notice that the attribute “MAXSPEED” has a value of 0 (zero) for the attribute “TYPE” classified as motorway with REF “E22”. So, it make complex the analysis and determination of the best dataset base on its quality parameters. The estimation of quality parameters of the data depends on the purpose and use of the dataset.

Table 10.1 The table below contains some of the attributes of OSM dataset

FID	Shape	OSM_ID	NAME	REF	TYPE	ONEWAY	BRIDGE	MAXSPEED
1192	Polyline	2387337	Sankt Lars vÃ¤g		unclassified	0	1	30
1193	Polyline	2387338		E22	motorway	1	0	0
1194	Polyline	2391705	BrunnshÃ¶gsvÃ¤gen		unclassified	0	0	0
1195	Polyline	2391705			cycleway	0	0	0
1196	Polyline	2399928	Sankt Petri kyrkogata		residential	1	0	0
1197	Polyline	2401871			footway	0	0	0
1198	Polyline	2401871			cycleway	0	0	0
1199	Polyline	2401879	SÃ¶dra Esplanaden		tertiary	0	0	0
...	...	...	...	...	...	...	...	...

10.3.2 Swedish national road database (NVDB)

NVDB has a more complex table in terms of attributes and it is simple in terms of roads type (cf. Table 10.2). NVDB focus only in two main road types: bicycle and vehicle network. There is also a difference between the geometric column SHAPE * type between OSM and NVDB described in section 10.4.

Table 10.2 The table below contains some of the attributes of NVDB dataset

OBJECTID *	SHAPE *	NETOBJECT	FROM_DATE	TO_DATE	SEQ_NO	Nättyp	Företeelsetillkomst	ROUTE_ID *	FROM_MEASURE	TO_MEASURE	SHAPE_Length
1105	Polyline Z	{93d5f809-2f	20090204	99991231	6	cykelnät	10849:20703	10849:16256	0	1	36.393447
1106	Polyline Z	{93d5f809-2f	20090204	99991231	7	cykelnät	10849:20703	10849:16262	0	1	26.931075
1107	Polyline Z	{93d5f809-2f	20090204	99991231	8	cykelnät	10849:20703	10849:16850	0	1	25.344967
1108	Polyline Z	{93d5f809-2f	20090204	99991231	9	cykelnät	10849:20703	10849:16916	0	1	214.633506
1109	Polyline Z	{93d5f809-2f	20090204	99991231	10	cykelnät	10849:20703	10849:16922	0	1	306.376168
1110	Polyline Z	{93d5f809-2f	20090204	99991231	11	cykelnät	10849:20703	10849:17756	0	1	255.943421
1111	Polyline Z	{93d5f809-2f	20090204	99991231	12	cykelnät	10849:20703	10849:17762	0	1	154.629815
1112	Polyline Z	{3e39231c-6	20081205	20090115	0	bilnät	10850:3231	10850:3216	0	1	27.45495
1113	Polyline Z	{3e39231c-6	20081205	20090115	1	bilnät	10850:3231	10850:3220	0	1	180.040644
1114	Polyline Z	{3e39231c-6	20081205	20090115	2	bilnät	10850:3231	10850:3226	0	1	98.408922
1115	Polyline Z	{83435ae0-6	20090115	99991231	0	bilnät	10850:3231	10850:3216	0	0.33418	9.174995
1116	Polyline Z	{83435ae0-6	20090115	99991231	1	bilnät	10850:3231	10850:3216	0.33418	1	18.279956
1117	Polyline Z	{83435ae0-6	20090115	99991231	2	bilnät	10850:3231	10850:3220	0	1	180.040644

10.4 Comparison of geometric data

Geometric data analysis can refer to geometric aspect of image, pattern, and shape analysis. In another word, it is the study of shapes and spaces. Thus, comparison between geometric data can be performed in one or all of these aspects. In this study the focus is more into the shape analysis.

10.4.1 OpenStreetMap (OSM)

The basic geometric data type of OSM is nodes (point), ways (link), and relations (section 4.2). The OSM nodes are points which are used to represent points of interest (POI) in the real world such as bank, gas station and schools. Nodes also connect the ways and can be considered part of the ways geometric data type as the start or end point. OSM ways are lines or links that represent roads railroads, rivers, etc. The ways and nodes combined are the two fundamentals geometric data types to form the links of the OSM network. The ways in the network usually are not defined with a direction. The use of tags would denote a direction for a specific way (e.g. tag oneway=yes). Relation is just used to model the relationship of the geographical objects. Relation is not used in this study however it has is an important aspect of OSM dataset. OSM geometric data type has its own characteristic but there are methods used to expand it usage by the use of tags (section 4.2) [16, pp. 1645-1650].

10.4.2 Swedish national road database (NVDB)

According to NVDB road network model, it is fundamental to describe the roads in reality in a proper and fundamental level. As well-as OSM, NVDB has its collection of road network links represented by lines and points. NVDB denote the lines which represents their roads network as “reference lines” which is a depiction of the roads’ positions. These reference lines are created with connected straight lines and a group of coordinate points as illustrated in Figure 7.1. On difference between OSM and NVDB is that OSM lines are stored as 2D dimension (cf. Table 10.1) and NVDB reference lines are stored in 3D dimensions (cf. Table 10.2).

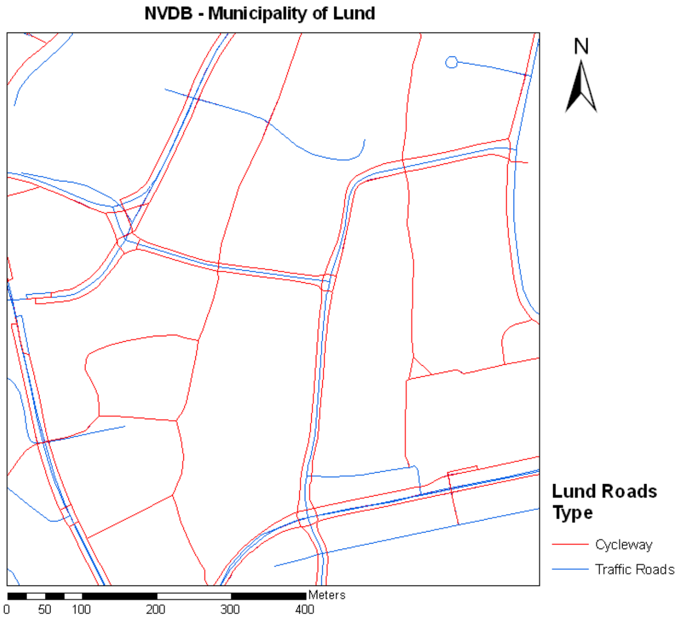


Figure 10.1b – NVDB roads dataset shown the same area as Figure 10.1a.

10.5 Comparison of topological data

Geographic data are representation of objects in the real world and its relation in space, linking places, time, and attributes as described in section 2.1. The datasets used in this project are road networks represented by geometric data lines or polylines and points (cf. Figures 10.5 and 10.6). It is important to maintain the topological aspect of the data and its topological relationship with the objects in the algorithms in the network analysis [2, pp. 359-364]. NVDB's topological base elements are nodes, reference links, and ports.

NVDB concept of the road network model is that it consists partly of the road network's geometry and partly of its topology [45, pp. 6-37]. It is because sometimes in the network reference links can be entirely or just part of it used to represent a road segment in the network. Reference links are geometrically represented as a reference line. Intersections as well as a roads' end points are called nodes and are represented by a defined point [45, pp. 6-37].

Figure 10.5 Map of Lund from OpenStreetMap. The square is the area displayed below on Figure 10.6

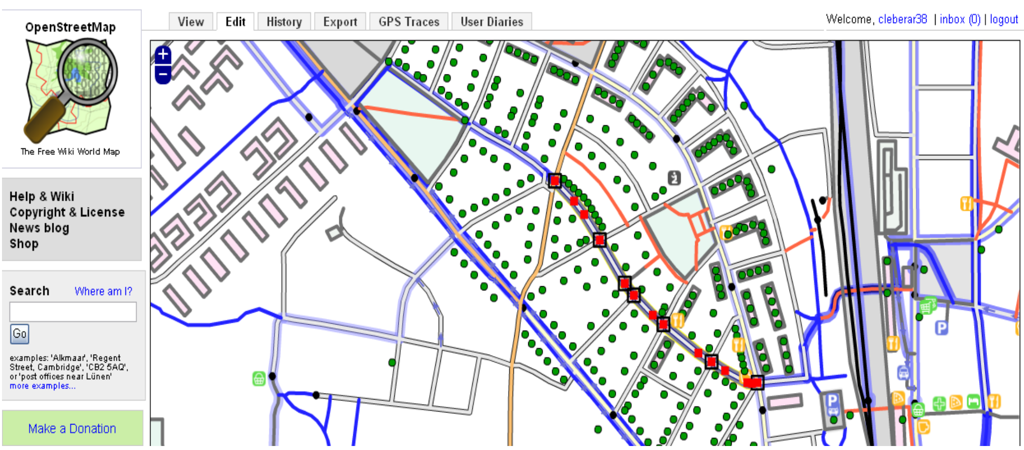


Figure 10.6 shows how data can be edited in OpenStreetMap and displays different geometric features that represent (e.g. roads, intersections, railroads, building etc).

Nodes

■

Examples of roads type

- footway
- highway
- motor vehicle
- pedestrian
- railroad

The shortest path algorithms use the network to estimate their paths. A problem in the route network was found during the pgRouting execution using the OpenStreetMap dataset. This is shown by intersections without nodes as seen in the graph-like figure, Figure 10.7.

10.5.1. Nodes

Objects in a network are described as simple graphic elements such as nodes. Nodes are points in terms of geometric elements which can represent points of interest (POI) and small geographic features on the real world. For instance, building, soil pit, wells, traffic signs, etc. Nodes can be interconnected by ways to define intersections, merges, and change in direction and usually it is denoted as the start and end point of the way. It is important to mention that nodes in the network have a specific identifier (ID) which helps the search efficiency [2, pp. 359-364]. However, there exists certain problem or connectivity error in a network which I denoted as the “missing node” causing difficulties in the search process within the network. One the reason for this problem is the data contributed by users who do not provides, and considers some pathways (e.g. parks, footways, tunnels as part of the network) (cf. Figure 10.7-10.8).

a. No nodes at intersections

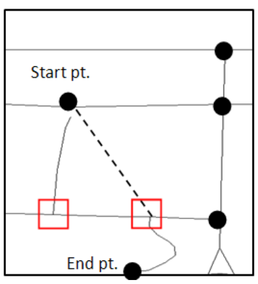


Figure 10.7 - The squares show that there are no nodes at intersections

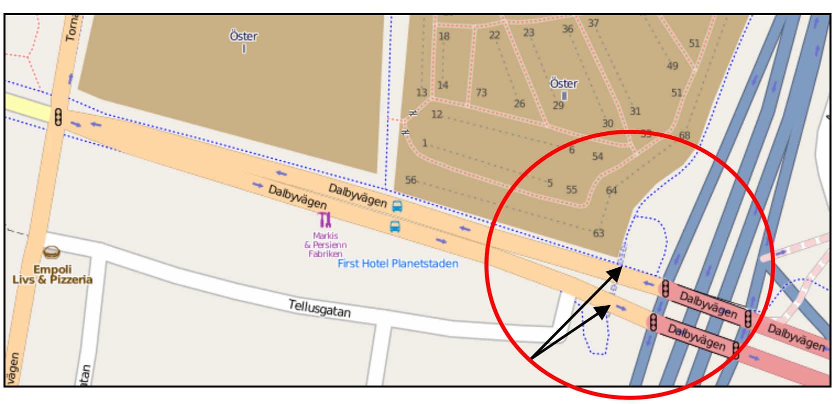


Figure 10.8 – Show the dotted line is the cycle and pedestrian navigation route under the highway bridge pointed by two black arrows.

b. Solution

There are solutions to the missing nodes at intersections problem. However, one should choose carefully one that is suitable and gives an optimal result to the problem. Some considerations were taken into account regarding this issue. ArcGIS allows you to create functions and you can generate a function which can splitting the network links on intersections if needed. However, since the area analyzed in Figure 10.8 looks like a "spaghetti bowl[1]" it becomes more complex. In order for you to create a new topology you should specify at which intersection the split should occur. The other approach is the creation of nodes by using FME, which allows one to build a topological network with edges and nodes. However, it could produce more problems because sometime there are intersections that there are no needs of nodes (e.g. cycle path under a bridge) Figure 10.8.

10.5.2 Parallel links

Ways are lines in term of geometric elements which can represent roads, river, railways, etc in the real world. It is one of the basic graphic elements of a road network such as NVDB and OSM. These lines are also denoted as links and are crucial elements in a network topology uniquely identified by the starts and ends node. Links are roads segments interconnected by nodes at intersections, junctions of two or more links (cf. Figure 7.1).

a. Parallel link issue.

A parallel link issue can be described as the link that would by forced diverted traffic, while path search navigation is performed, if the link in consideration were closed. In another word, even if the link in consideration is one of the shortest links on path, it would not be taken. Note that a parallel link may or may not be connect with the link in consideration (cf. Figure 10.9 B-C). A parallel link is one type of topological aspect of a route network that has two different paths with the same nodes as source and destination. When analyzing the dataset from OpenSteetMap, a parallel link problem was found (cf. Figure 10.9 B). As shown in Figure 10.9 B on the right side Dijkstra's algorithm routes the path, choosing the longest way.

[1] Term used to describe a road network where highways are interconnected sometimes with other categories of route such as pedestrian route cycle, etc.

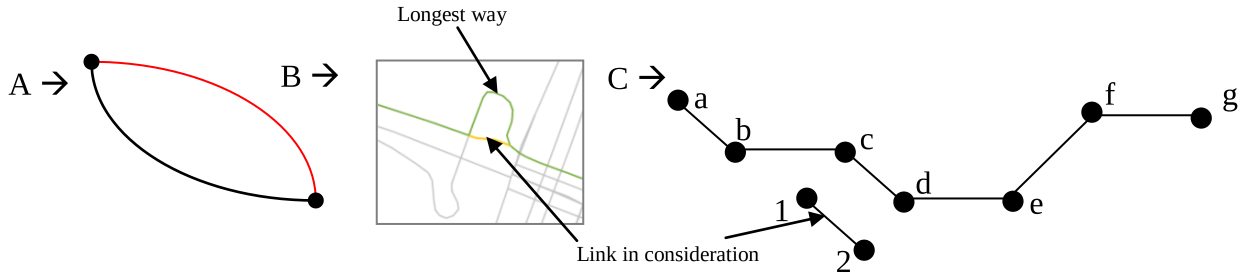


Figure 10.9 shows parallel link problems.

b. Solution

In order to solve this issue, one approach was used the introduction of the shooting star algorithm, since this algorithm can solve problems such as the "parallel links".

The geometric and topologic aspects of a route network are fundamental when developing a routing network service. The comparison between OSM and NVDB is also based on these characteristics to develop a own pedestrian network service (PRNS)

10.6 Tailoring the road datasets for creating a pedestrian route network service (PRNS)

There are some steps to add a network topology to the route network and assign to each link a unique identification. It was accomplished by the use of the assign_vertex_id function, in PostgreSQL which requires the following parameters: the table, tolerance, geometry column and the unique identification (ID) for each link on the network. As mention above, each shortest path algorithm uses different parameters. Dijkstra algorithm uses the source, target, and length attributes. A*(A-star) uses more attributes such as the latitude and longitude for start and end point for each link. As any new method implementation shooting star algorithm introduce new attributes while preparing the network table. They are the "rule" (list of edges IDs and turning restriction) and the "to_cost" (restricted passage) columns [21, pp. 20-21]. The shooting star algorithm, however, uses one additional attribute, which can be described as the reverse_cost. The sections below describe how one can manipulate the datasets for creating an own PRNS.

10.6.1 OpenStreetMap (OSM)

The Dijkstra shortest path algorithm was used on the tailored OSM dataset. The first problem that was presented was the floating tolerance parameter from the assign_vertex_id. However, this problem was solved by just projecting the dataset reference system from SWEREF 99 to SWEREF 99 13.30, which is explained in section 9.3.

The second problem was an inaccurate network topology. There were some intersections without nodes, implicating the routing process to return no line segments from the routing and shortest path process. Each intersection should have a node, except, for example, when the intersection happens to be on a bridge or tunnel. In Figure 10.10 there is an example where two edges intersect and there is no node present on it.

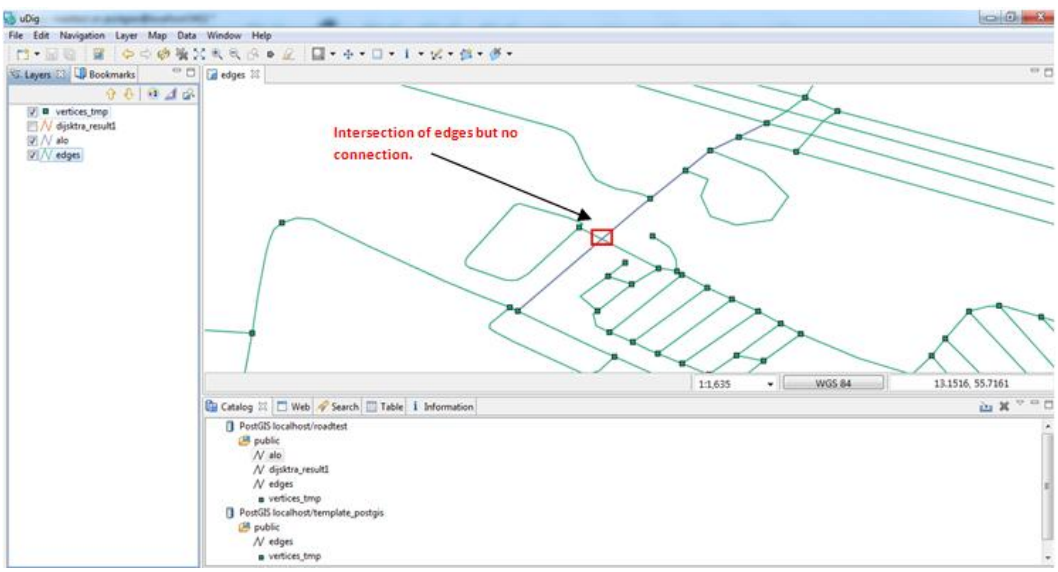


Figure 10.10 Screenshot of OSM dataset in uDig. The square shows an intersection between edges but no existence of node after the routing process using the OSM dataset.

The solution for this problem was the use of ArcGIS software to create the right topology for the route network. Based on the original nodes, new nodes were created. By adjusting and overlaying the new set of nodes, the correct route network topology was created (cf. Figures 10.11a-e). Figure 10.11a shows the network which contains all the links and nodes at each intersection plus the vertexes which are additional point that are not of our interest. Figure 10.11b shows the network with the nodes of interest on each intersection without the vertexes. However, there are intersections that should not be split and these nodes should be removed in order to avoid creation of incorrect topology.

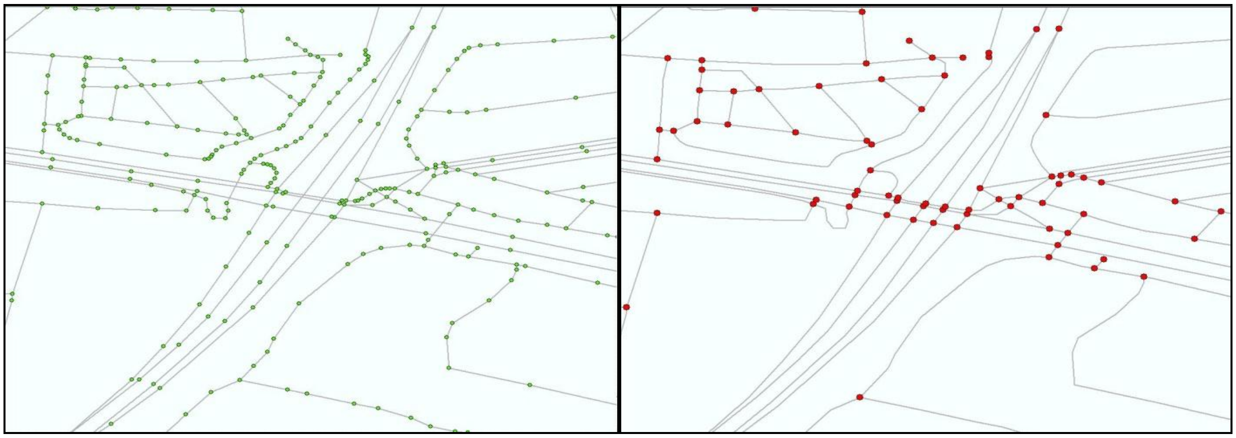


Figure 10.11a (left) – Screenshot of ArcGIS showing the original OSM nodes, including all the vertices on the line segments.

Figure 10.11b (right) – Screenshot of ArcGIS showing the intersections and nodes on the line segments on the OSM dataset.

Figure 10.11c shows the overlay of the Figure 10.11a and Figure 10.11b above which are used to eliminate the nodes and vertex that should not be used to split the lines on the network. Then after of the overlay and the use of the split function Figure 10.11d shows the correct topology that can be used for the creation of a route network service.

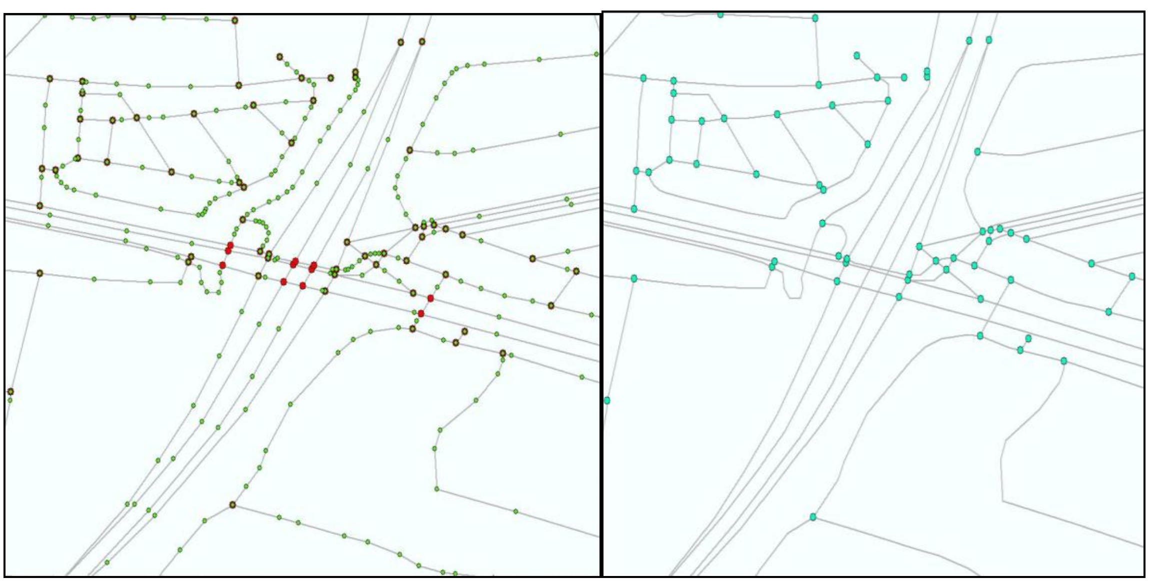


Figure 10.11c (left) – Screenshot of ArcGIS showing the overlaying and adjusting of nodes at the intersections of the OSM dataset.

Figure 10.11d (right) – Screenshot of ArcGIS showing some nodes' intersections on OSM dataset.

After the analysis and use of split method to correct the topology on the OSM network, another area of the municipality were used to verify if the topology was correctly created (cf. Figure 10.11e. A comparison between Figure 10.11e (OSM image) and Figure 10.13 (NVDB image) show the difference between both dataset used for this study. There are several roads on the network missing on NVDB image in Figure 10.13 since it does not focus on provide information on pedestrian network (section 4.2).

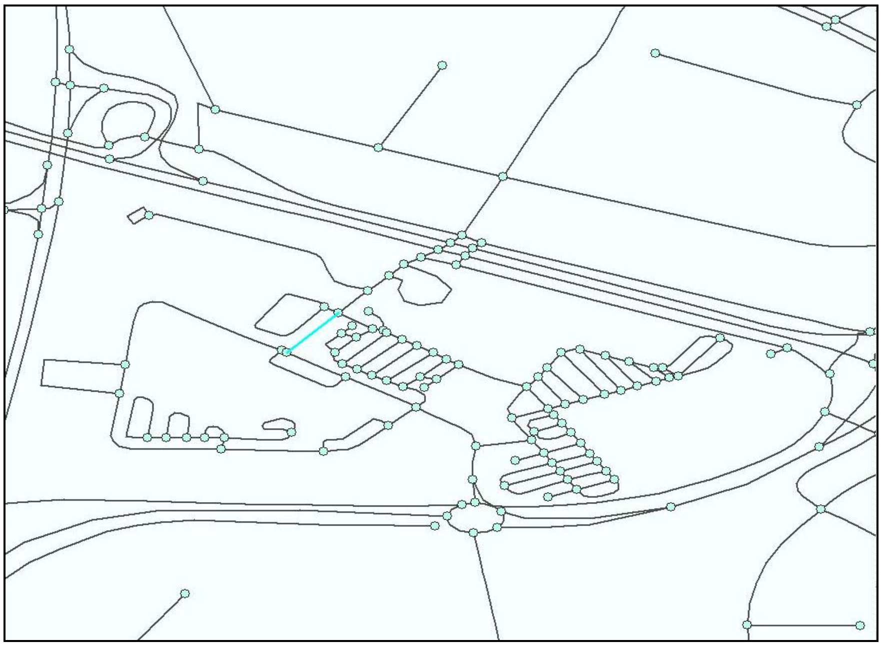


Figure 10.11e Screenshot of ArcGIS showing the correct topology and nodes at each intersection on the OSM network dataset.

10.6.2 Swedish national road database (NVDB)

The NVDB dataset network is a designated to motor vehicles and some cycle path (section 4.3). In order to fix or tailor NVDB dataset to be ready to create a pedestrian network service one have to perform substantially amount of work. It would be necessary to update and reclassify all roads that are used by motor vehicles and cycles that can be used by the pedestrians. After one would have to add new geometries and create new network topology among several other steps. Therefore, the idea of fix the data was discarded the NVDB dataset was utilized in this thesis for comparison and analysis.

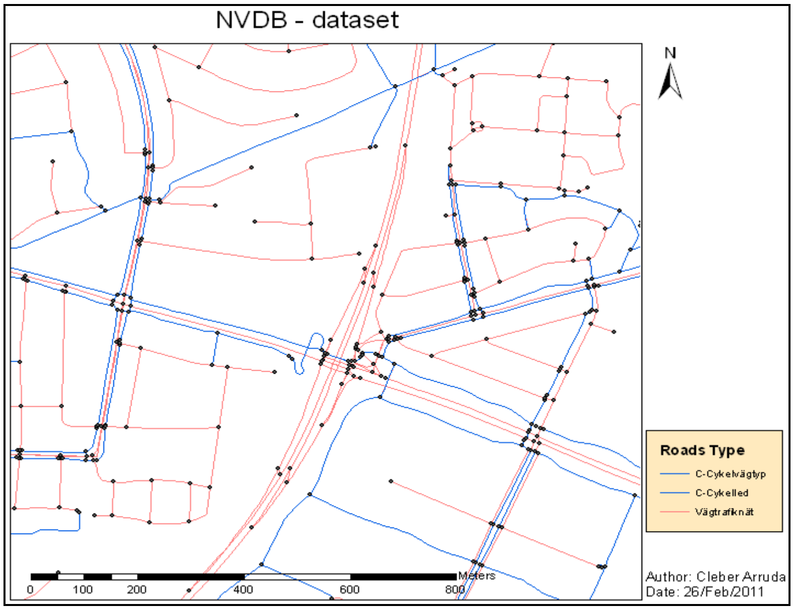


Figure 10.12– Shows a map of NVDB bicycle and motor vehicle network southeast area.

Visually one can also see the missing routes in Figure 10.13 in comparison to Figure 10.11e. The bicycle path is denoted by the routes which are assumed to be shared with pedestrian. The routes are the motor vehicles routes which include highways, streets, etc.

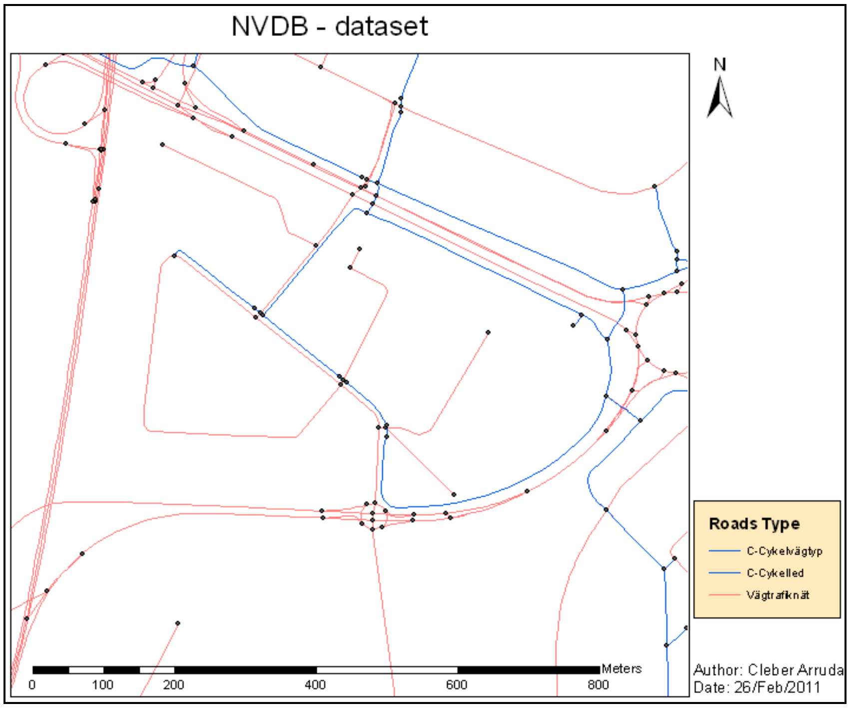


Figure 10.13 – Shows a map of NVDB bicycle and motor vehicle network northwest area.

10.7 Evaluation

During the analysis between OSM and NVDB datasets we found that both datasets present its advantages and disadvantages in term of data quality (cf. Table 10.3). The attributes from each dataset analyzed are used for its own and different purpose. For example, NVDB dataset is intended for motor vehicle and some cycle path (cf. section 4.2) having its limitation for users who wants to use it for developing an own PRNS. NVDB has a lack of specific roads for pedestrian on its road network. In contrast, OSM dataset have suitable attributes in order to create an own PRNS with different roads types and its properties. However, OSM is an open source project where data are shared and contributed among many users (cf. section 4.2). The OSM is update frequently as described in section 11.7. Section 2.5 describes data quality allowing one to realize that its analyzes are complex.

Table 10.3 – Comparison of quality between OSM and NVDB and the check mark (✓) indicates best category.

	OSM	NVDB
Guaranteed quality		✓
Positional accuracy		✓
Completeness of pedestrian network data	✓	
Updated frequently	✓	
Available for public (open source)	✓	

11. Implementation of the pedestrian route network service (PRNS)

11.1 Application of PRNS

The PRNS implementation is integrated within a spatial database PostGIS (cf. section 5.5). The dataset used is from OpenStreetMap (OSM). The dataset is added into the PostGIS database using the Structure Query Language (SQL). There are also some others functions and features that will be also stored in the database in order to maintain its integrity. One of those functions is the new “shooting star smart” function released in 2011-Feb [5]. Those functions and features are provided by the pgRouting contributors and can be added using a SQL files which can be downloaded from pgRouting or PostGIS web site (Chapter 6).

Information such as the type of road, speed limit, reference, and osm_id, are maintained and stored in the database as shown in Figure 10.2. Sometimes a weight or cost value is given to each type of road and the convenient path is determined based on that weight or cost. This information is used by the algorithms to determine the shortest path on the network.

11.2 Data

The aim of this study is to implement a route network service after the comparison of two datasets from different sources. Consequently, use the appropriate dataset in an own routing network service (e.g. pedestrian route network service (PRNS)) using pgRouting. The datasets used for these comparisons are the Swedish national road database (sw. NVDB) dataset and the OpenStreetMap (OSM) dataset. At the end of this study, a decision should be made to determine the best dataset, between OSM and NVDB datasets, that suites for the development of the PRNS, based on the analysis of these datasets.

It is important to take into consideration various aspects to determine the quality of the OSM and NVDB datasets. The NVDB dataset has its limitation of roads for pedestrian on it is network (section 10.7). However, according to NVDB’s documentation NVDB shall have a guaranteed and well-defined level of data quality based on its quality model. The qualitative requirement depends on the NVDB’s number of users and their needs. OSM dataset is more appropriate to be used in the development of own PRNS. The OSM dataset has more pedestrian roads types on its network than NVDB. According to OSM documentation on its website, the quality assurance is performed based on rules and data analysis, or using the quality assurance tools that provides a means of manually reporting, or some combination of the two [34].

The datasets are also compared in terms of its network data model. It basically reveals the different geometric and topological aspects of the data (section 10.4 and section 10.5)

11.3 Workflow of routing service

Structured Query Language (SQL) described in section 3.2. The SQL queries are the ones which retrieve the information from the database (PostGIS) using the PostgreSQL features and tools. It is done through another programming language called PHP. PHP file, as showing in Figure 11.8, gets the result from the queries and returns it to be post on the web map service which is executed by the HTML file. The Hypertext Markup Language (HTML) file uses the OpenLayers as showing in Figure 11.9, a free open source JavaScript, to display and manipulate the maps on the web map service. Figure 11.3 shows the workflow of the routing service in detail.

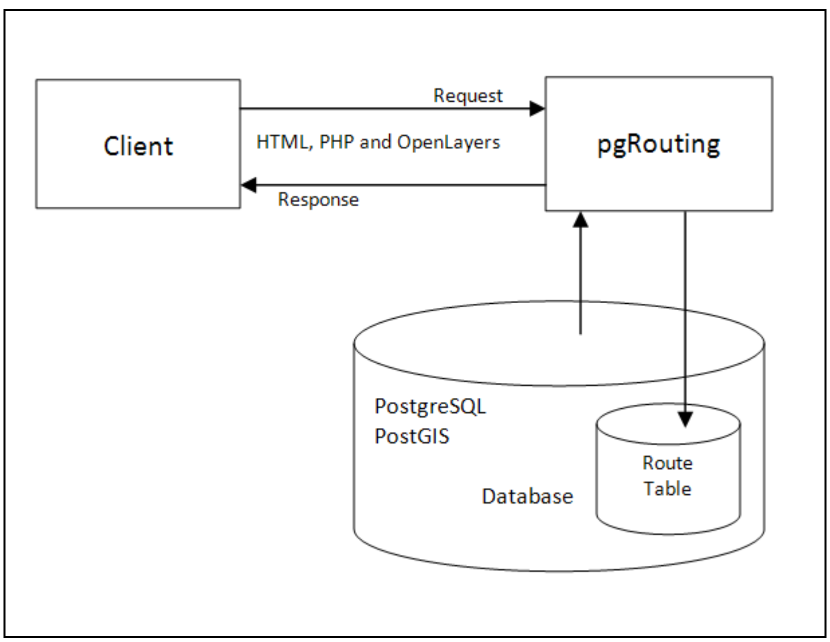
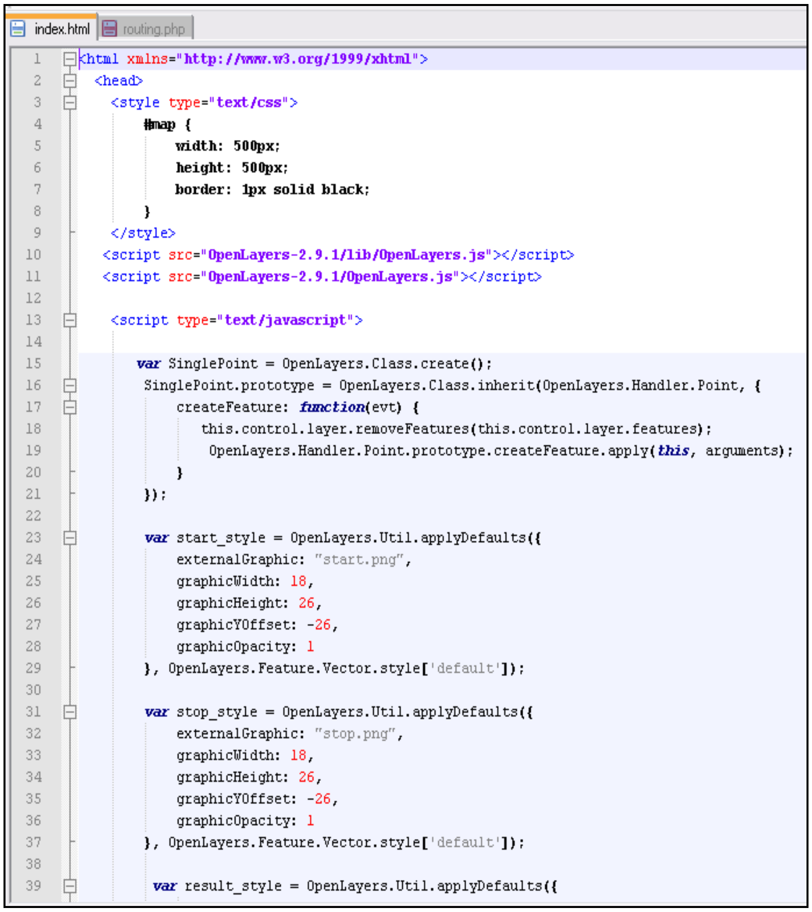


Figure 11.3 – Workflow of routing service

11.4 Implementation of network algorithms

One way to understand network algorithms is to look at a road network. Before implementing a network algorithm, several steps should be taken into consideration. One such step is the implementation of a graph, represented with nodes and edges, with generic terms assigned to the distance for each edge. Shortest path algorithms are defined in term of distances. Shortest path's distance estimation is also categorized as any specific unit such as weight cost, cost, distance cost, length, link cost, and edge cost [27, pp 30-54]. Therefore, one should consider the storage of the data and information concerning the network, among other relevant aspects. A network must be stored in an efficient way in order to assist spatial selection and join of queries. Consequently, the network algorithm is implemented by the use of a routing tool (e.g. pgRouting).

Figure 11.8 - Shows a routing sample performed by the combination of the programming data used for pedestrian route network service.



```
index.html   routing.php
<html xmlns="http://www.w3.org/1999/xhtml">
  <head>
    <style type="text/css">
        #map {
            width: 500px;
            height: 500px;
            border: 1px solid black;
        }
    </style>
   <script src="OpenLayers-2.9.1/lib/OpenLayers.js"></script>
   <script src="OpenLayers-2.9.1/OpenLayers.js"></script>

    <script type="text/javascript">

       var SinglePoint = OpenLayers.Class.create();
        SinglePoint.prototype = OpenLayers.Class.inherit(OpenLayers.Handler.Point, {
            createFeature: function(evt) {
               this.control.layer.removeFeatures(this.control.layer.features);
                OpenLayers.Handler.Point.prototype.createFeature.apply(this, arguments);
            }
        });

        var start_style = OpenLayers.Util.applyDefaults({
            externalGraphic: "start.png",
            graphicWidth: 18,
            graphicHeight: 26,
            graphicYOffset: -26,
            graphicOpacity: 1
        }, OpenLayers.Feature.Vector.style['default']);

        var stop_style = OpenLayers.Util.applyDefaults({
            externalGraphic: "stop.png",
            graphicWidth: 18,
            graphicHeight: 26,
            graphicYOffset: -26,
            graphicOpacity: 1
        }, OpenLayers.Feature.Vector.style['default']);

         var result_style = OpenLayers.Util.applyDefaults({
```

Figure 11.9 – Shows part of the HTML code using OpenLayers library JavaScript.

11.5 Technical solution

Based on the methods used to develop own pedestrian route network (PRNS) the subsequent technical solutions described in this section are fundamental. The use of OpenStreetMap dataset is adequate to fulfill the PRNS objective since it has proved, after the analysis, to have necessary features for the PRNS. The evaluation of the network data between the two datasets OSM and NVDB using application software and tools (Chapter 10) is one of the most important aspects in this study. The ArcGIS was used to generate, view, manipulate, and covert and project the data. It helps to create maps and figures in order to compare the datasets. uDig another GIS application was mostly used to view the data since it can connect directly with the database PostGIS. FME was also utilized to view convert and project the data. The PostgresSQL along with PostGIS, pgRouting and pgAdmin were used to manipulate and organize the data and perform routing service. Results become satisfactory but with its limitation which are described in chapter 13. However, practically the PRNS follows the same pattern as exiting routing service which served as an approach to its development.

11.6 Ensuring completeness of the PRNS

In order to ensure the completeness of the pedestrian route network service (PRNS) the most inspected element over many others were the dataset from both OpenStreetMap OSM and Swedish national road database (NVDB). The OSM dataset is the appropriate to be used for the routing service as mention in section 11.7. Another approach used to guarantee the completeness of the PRNS was to pursue the ideas of other people of routing service. For example, follow accessible work such as pgRouting routing service described in section 11.4, its documentation and discussion among many users worldwide. Suggestions and consultations with supervisors from municipality of Lund, coworker, and Lund's University remarkably assisted on this matter.

11.7 Analysis

This study has revealed that some of the results are not as expected results which are influenced by some of the methods implemented to create own pedestrian route network (PRNS). As shown in Figure 11.10 and Figure 11.11a the performance of the routing process is done throughout the municipality of Lund using the shooting star algorithm described in section 7.4. This is a satisfactory result compared to other existing routing network service described on chapter 12. However, Figure 11.11b is illustrates a problem during the routing process which I denoted as “under shooting” and “over shooting”. Under shooting is when the routing process does not reach its destination or start from its expected position. Over shooting is when the routing process reaches but exceeds or passes its source or destination expected position. These phenomenona can either happen at the start point or end point. One of the reasons for this occurrence is the implementation of the shooting star algorithm which works with edges or link to link approach (cf. section 7.4). Looking at the results in Figure 11.10, Figure 11.11a and Figure 11.11b there are several types of selected path such as highways, cycleway, pedestrian, footway, etc which are collected by the shooting star algorithm. Thus, another weakness of the methods used for the development of own PRNS. Nevertheless, there are relevant solutions for these problems described and discussed in Chapter 13.

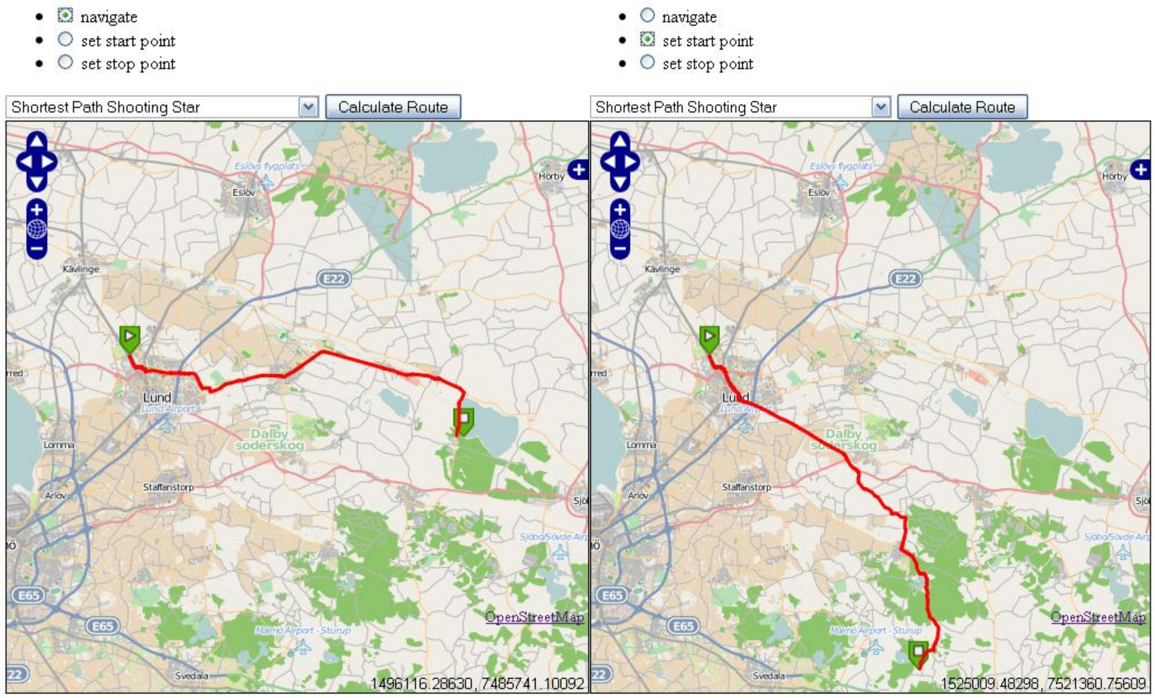


Figure 11.10 – Shows the routing through the boundary of municipality of Lund.

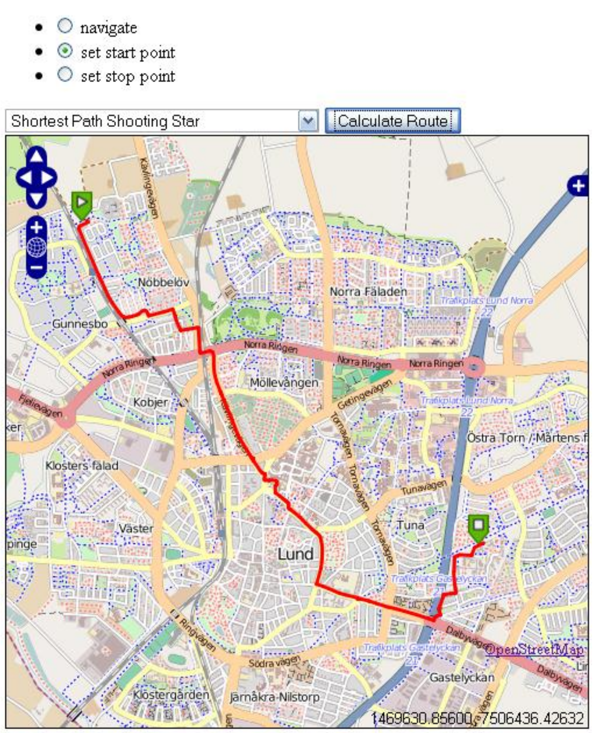


Figure 11.11a – Shows the routing within the city of Lund.

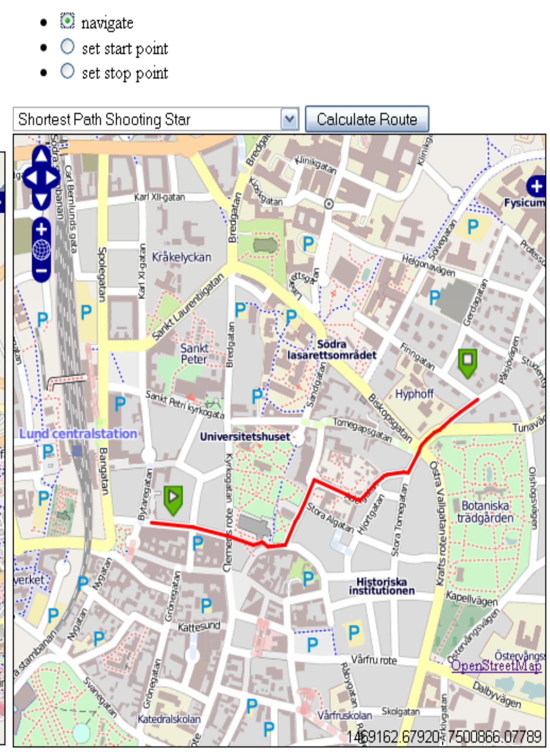


Figure 11.11.b – Shows routing under and over shooting.

12. Case study II: Comparison of pedestrian route network services

The aim of this case study is to compare a number of commercial route services with our own

There are many users of routing service worldwide and they are benefited from the commercial or open source routing service today available on the internet. In fact, there are many contributors to routing services such as pgRouting project in order to its development. Some of these routing services are mentioned in this study. Each of them has a specific purpose and is designed to fulfill the need of its creator.

12.1 iMapTools routing service

The iMaptools and LeadDog Routing DEMO was designed to calculate a route path and give the direction either in English, French or Arabic (cf. Figure 12.1). Another example is the OpenRouteService which also calculates a route for car, bicycle or pedestrian.

Seq	Instruction	Distance meters	Time seconds	Accum. Dist. km	Accum. Time minutes
1	From start point, drive east	615.3	31	0.615	0.5
2	continue straight on to IMAM KHOMEINI BOULEVARD	3472.3	174	4.088	3.4
3	stay right on to an unnamed road	62.1	4	4.15	3.5

Figure 12.1 iMapTool showing routing service providing direction in three different languages.

12.2 OpenRouteService

OpenRoutingService provides travel route and navigation information (cf. Figure 12.2). It is focused on the interoperable Service Oriented Architecture (SOA) and is based on the open standard of Open Geospatial Consortium (OGC). In order to deliver the right information to its users OpenRouteService.org use several parameters such as a route summary, route geometry, route instruction, and route maps. OpenRoutingService also uses OpenStreetMap (OSM) data. It is in the interest of this project to research the OpenGIS Location Service (OLS), implementation of Location Based Service (LBS), OpenGIS Web Services, and OpenStreetMap for accessibility of free geodata [33]. OpenRouteService offers others facilities in this routing service such as the location of a point of interest (POI), GPS eXchange Format GPX, manipulation of display features such as route color distance unit kilometers, meters, yards, and inches, etc.

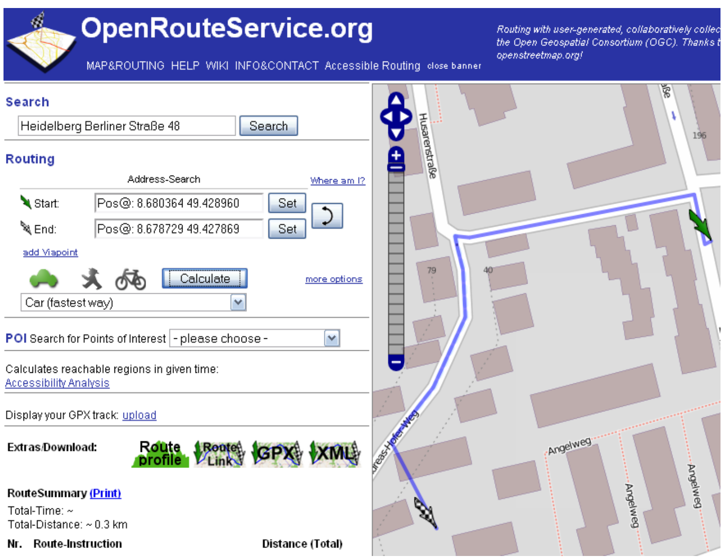


Figure 12.2 Shows some of the characteristics of OpenRouteService routing service.

While analyzing the services a comparison between different services was performed. The OpenRouteService can be used only in certain location and (e.g. Germany) (cf. Figure 12.2) above. If the route is compute in a different location no route is computed or displayed (cf. Figure 12.3). It is because the service is a local based service. There are also other route services which can compute the route in any part of the world such as Bing web map service. However, this is not the focus of this study.

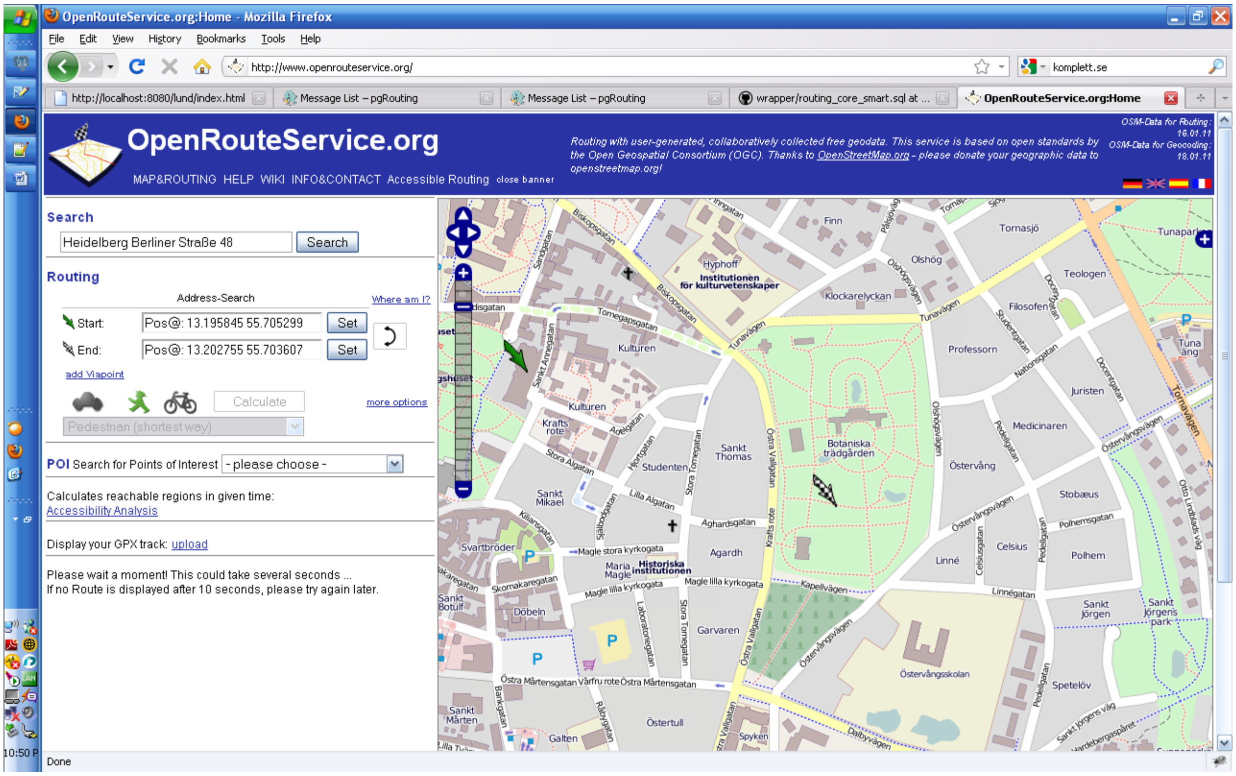


Figure12.3 – Shows no route computation when service is performed in the municipality of Lund

12.3 CloudMade routing

CloudMade is one of several providers of open route services on the web. CloudMade enables developers to build different Location Enabled Applications and Services along with the use of APIs. These APIs enable one to create location based applications (LBA). Today, many people use the CloudMade platform to build their own applications using the APIs. Some of the services offered by CloudMade are StyleEditor, Web Map API for references, Geocoding and Routing for business. CloudMade uses and contributes to OpenStreetMap (OSM) data. [3]. Below, Figure 12.4 shows one of the routing services provided by CloudMade (e.g. get direction request using OSM). In this example, a driver can choose the start point (A) and end point (B) in Roma, Italy.

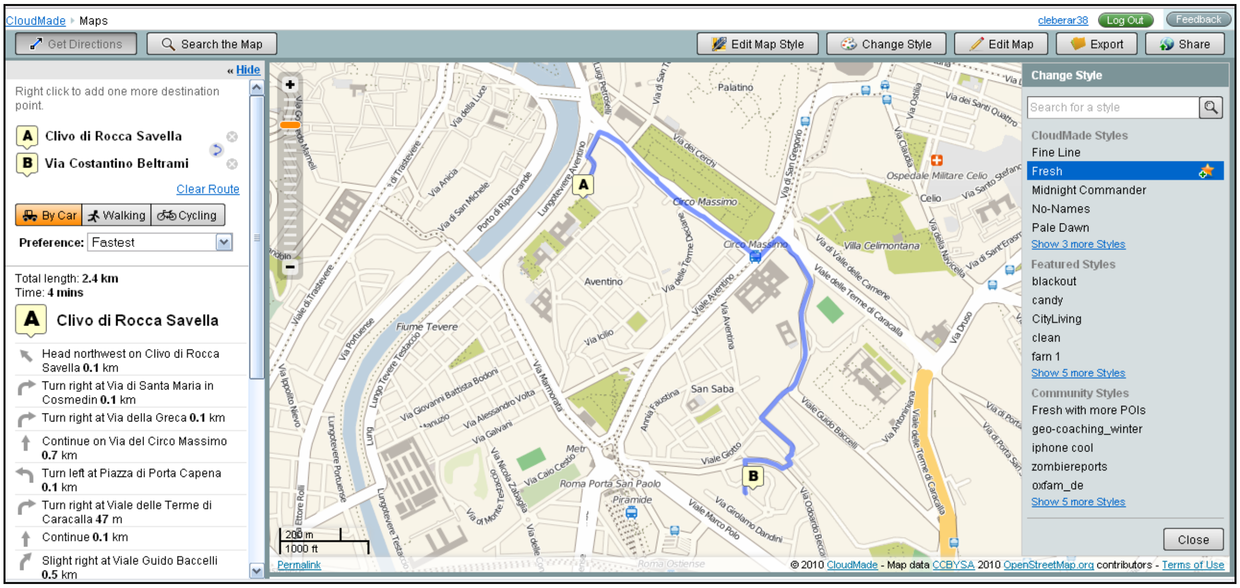


Figure 12.4 – Shows a get direction route service from one point source to the next point target.

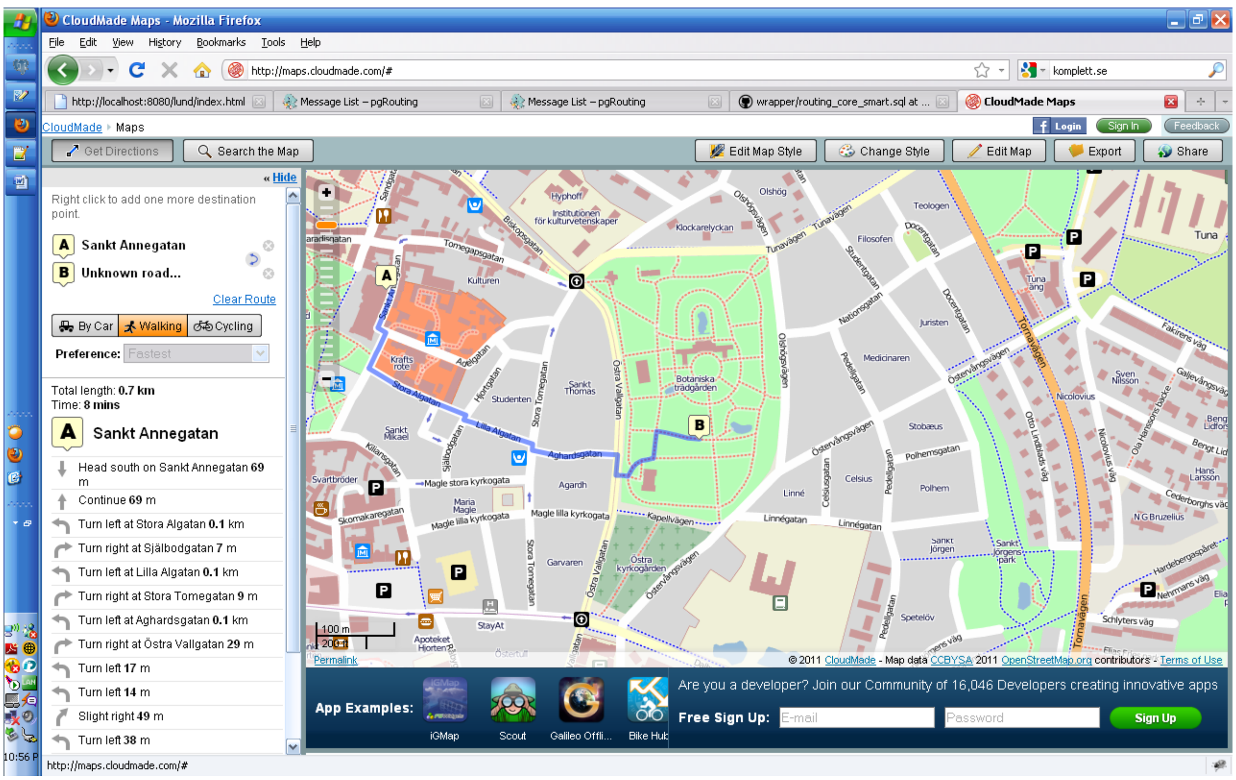


Figure 12.5 Shows routing perform in Lund using CloudeMade.

12.4 Own pedestrian route network service (PRNS)

The pedestrian route network service (PRNS) developed for this thesis is a model of some of the routing service mention above. However, they differ from each other in several aspects. For instance, PRNS is intended to perform the routing only in the municipality of Lund while CloudeMade routing service can be perform route service globally. The routing paths vary as well from service to service. A visual comparison can be seen in Figure 12.5 and Figure 12.6 when the applications and or algorithm perform the calculation in the same area (e.g. municipality of Lund). As both figures as showing they create a completely different route from the same source (start point) to destination (end point). It can be assumed that each routing service has its own objectives for fulfill according to its demand.

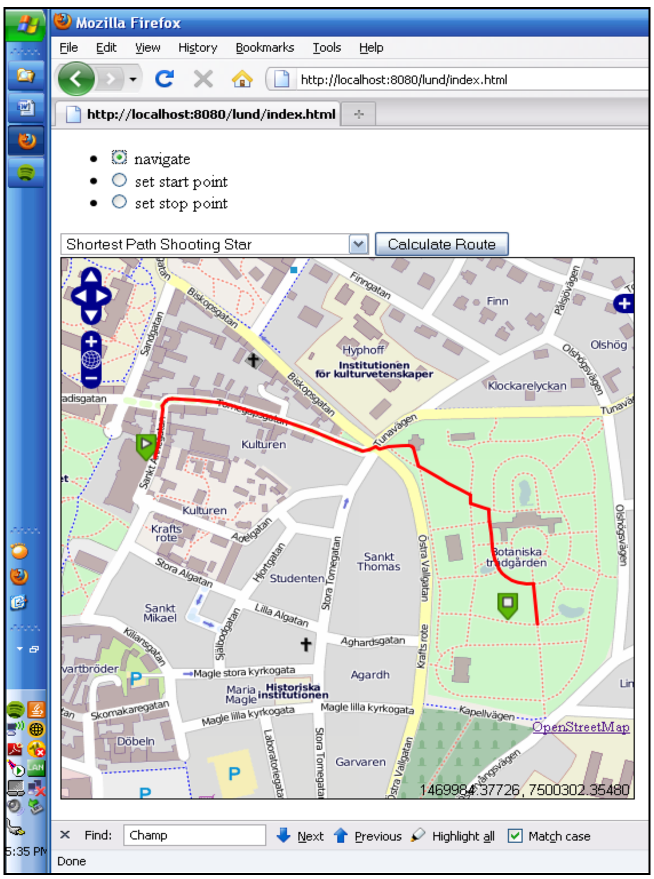


Figure 12.6 Shows route perform in Lund using own pedestrian route network service

12.5 Evaluation

Today, there are several types of route network services used by users in order to facilitate their search for route navigation (wayfinding). Therefore, some comparisons between some of these services were performed based on a number of categories such as services advantage and disadvantage, cartography aspects, map navigation and interaction, actions and mapping route services. Each category was choosen by the general characteristic of all the routing services. Table 12.1 has the names of the providers and the categories that were evaluated. To distinguish the services, a rank of values from 1 to 5 (5 is the highest) and a "yes", "no" or "partially" (some countries) answer to the categories are used when applicable. The analysis and comparison were done visually and during the routing process for each provider and the own PRNS.

In term of cartography the only categories analysed were symbology and visual hierarchy. For this study they are considered the most crucial categories since symbology represents the features on Earth abstractly, and the visual hierarch emphasize the objects according to their importance on each map.

Table 12.1 Comparison between existing and own PRNS.

	CloudeMade	OpenRouteService	iMapTools	Own PRNS
Cartography (*symbology*)	5	3	4	3
visual hierarchy	5	4	4	4
Interactions				
search bar	yes	yes	yes	no
Services				
Global Service	yes	partially	partially	no
text direction	yes	yes	yes	no
Views				
map view	5	5	3	4
satellite view	yes	no	no	no
map+satellite view	yes	no	no	no
Navigations				
pan	yes	yes	yes	yes
draggable	yes	yes	yes	yes
zoom	yes	yes	yes	yes
Actions (share)				
permalink	yes	yes	yes	no
email	yes	no	no	no
printable	no	no	no	no
Mapping route service				
driving direction	yes	yes	yes	no
walking/pedestrians	yes	yes	yes	yes
cycle route	yes	yes	yes	no

13. Discussion

The nature of developing a pedestrian routing service obviously implicates the application of factors which are somehow limited. In response to the results from the analysis in section 11.7 I suggest that some of the methods implemented be enhanced with newer context. The own pedestrian network service (PRNS) is intended for Lund city and for the purpose of guiding pedestrian and allows them to choose the best path to navigate within this municipality. However, it can be expanded in the future, based on its demand. The PRNS has started already and the analysis made on it has revealed that changes have to be made. Data quality analyses are very complex and both dataset OpenStreetMap and Swedish national road database (NVDB) have it pros and cons.

Despite the changes to be made and the limitation that the PRNS has the overall tendency of its objective is clear. It is a little bit problematic and contradictory the fact that there are many contributions and facilities available for routing network service there are also restrictions. Some of the results found in this study and analyzes such as the “under shooting” and “overshooting” problem can be solved. However, there are solution that one can implement such as the manipulation of the database, creation of new functions, modification of the algorithm and existing functions in PostGIS, etc. The shooting star algorithm does what it is intended to do, but one wants it to perform something else such as find the exact position of the start and end point in the routing. There are several examples which is a counter to the problem but it is personalized by developer who possible intention are privatize, commercialize, or even share, but portion of it source.

For the future a further solution to the expansion of PRNS is to design methodically a newer version of the PRNS with the recommendation and suggestion state in this study such as comments and discussion in section 11.7 and section 11.8. It would be an experimental and relevant idea in order to continue the comparison of PRNS between other routing services. I would say that it is impossible, therefore, to identify every single issue in a so complex and dynamic application.

14. Conclusions

14.1 Evaluation of geographic data for PRNS

Considering the idea of estimation we have evaluated two different datasets for the development of PRNS. The first dataset being evaluated is the OpenStreetMap (OSM). OSM is mostly likely to be used for the implementation of the PRNS. The second dataset is the Swedish national road database (NVDB). Although NVDB is considered to have better positional accuracy and quality than OSM in terms of road network, NVDB has a lack of pedestrian route network data. Another advantage of OSM over NVDB, is that OSM is a open source available for the public (cf.10.3). NVDB restrict the use of its data under certain rules and conditions (section 4.3).

Finally, the most appropriated dataset for the PRNS is the OpenStreetMap (OSM). OSM dataset includes pedestrian path and it is updated almost every day (section 4.2) and it is more accessible than NVDB.

14.2 Develop an own PRNS

The creation of the pedestrian route network service (PRNS) had several steps. It started with the evaluation of some aspects of two different datasets, OSM and NVDB (Chapter 10). After the evaluation a decision was made based in some criterias discussed in section 10.7. The decision is to use the OSM dataset, which is considered to be the appropriate dataset for the PRNS. The OSM dataset had to be tailored, converted, projected and manipulated (Chapter 9). The dataset were stored in PostGIS database (section 6.5.1) to facilitate its retrival, access and process. The routing was performed using pgRouting (section 6.5.3). It assited in creating the routing results described in sections 11.4, 11.7, 12.4. To assit the interaction between the map, routing table and database, the implementation of OpenLayers (section6.6), JavaScript (7.1), and PHP (section 7.2) were essential (cf. Figure 11.3).

14.3 Compare the own PRNS with other available pedestrian route network services.

In this study there are three existing routing network services used to compare their services an aspects with the own pedestrian network service (PRNS). They are the CloudMade, OpenRouteService and iMapTools. While comparing their services and characteriscts, significant results were found based on a number of categories (Table .12.1). Those categories were evaluated visually when maps and results of routing process were displayed on the web browsers, and technically by executing the routing services of each routing service provider in differents areas and or location on the Earth (globally and locally). More categories can be added to the table depending on the interests, objective and analysis performed. There are limitation in order to get better results. However, it is possible to develop a better applicaton but it requires further studies and implementation of new or improved methods to generate routing services such as "shooting_star_smart" function (section 5.4). After the routing process the results were produced using the images from "Print Screen" function (cf. Chapter 12).

References

[1] Brisaboa, N.R and Viqueira, J. R., et al., 2005, Improving Accessibility of Web-Based GIS Applications. IEEE Comput. Soc. pp. 490-494.

[2] Chen, Y., and Yao, M., 2008, Dynamic Topology Construction for Road Network, pp. 359-364, Fourth International Conference on Networked Computing and Advanced Information Management

[3] CloudMade Makes Maps Differently. http://cloudmade.com [visited Nov/2010]

[4] e-Adept - Planning: Navigation: Travelling: Report. http://www.eadept.se. [visited Aug/2010].

[5] E-mail (Personal Contact) – Kastl, D. 2011, WORKSHOP - FOSS4G, E-mail to Arruda, C. [Online], 15 Feb., Available E-mail: daniel.kastl@georepublic.de [accessed 2011, Feb. 15]

[6] Elias B., 2007, Pedestrian Navigation - Creating a tailored geodatabase for routing. 4th Workshop On Positioning, Navigation And Communication 2007 (Wpnc'07), Hannover, Germany.

[7] Engberg, L. E., Andersson, B., Jivall L., et al. A New Co-ordinate System for Sweden – Lantmäteriet, pp 1-6.

[8] Environmental Systems Research Institute, Inc (ESRI). www.esri.com [visited 8/Oct/2010] .

[9] Federal Geographic Data Committee (FGDC). http://www.fgdc.gov [visited-Sept/2010

[10] File System - http://www.cs.miami.edu. [visited 16/Sept/2010]

[11] FME. http://www.safe.com/technology/FMEplatform/overview.php. [visited-Aug-Sept/2010]

[12] Forum (Personal Contact) – Patrusher, A. 2011, Searching for Shooting-Star Algorithm, [Online], Available E-mail: anton@orkney.co.jp [accessed 2011, Mar.]

[13] Geofabrik. http://www.geofabrik.de/ [visited Sept-Dec/2010]

[14] HaptiMap project. http://www.haptimap.org. [visited Aug/2010].

[15] Harrie, L., 2008, Spatial Database Lecture Notes 2 in GIS Algorithms. GIS Centre and Department of Physical Geography and Ecosystem Analysis pp. 13-21.

[16] Hentschel, M., and Wagner, B., 2010, Autonomous Robot Navigation Based on OpenStreetMap Geodata, pp. 1645-1650, 13th International IEEE Conference on Intelligent Transportation Systems

[17] Holmstedt, H. A., 2007, Experience Augmented Pedestrian Wayfinding. Halden, Norway.

[18] Howland, J. E., 2008, Functional Programming and the J Programming Language, pp. 1-4.

[19] Information Free Way. http://www.informationfreeway.org/ [visited Sept- Oct/2010].

[20] Introduction to PostGIS - http://workshops.opengeo.org/postgis-intro/introduction.html [visited 21/Sept/2010].

[21] Kastl, D. and Junod F. *Workshop – FOSS4G routing with pgRouting tools, OpenStreetMap road data and GeoExt Manual*. 2010 – Release 1.

[22] Kobayashi, K., Lee, R. and Sumiya, K., 2010, *Pros and Cons of Utilizing Mobile Navigation Systems in Pedestrian Way Finding*. 2010 IEEE 24th International Conference on Advanced Information Networking and Applications Workshops. Hyogo, Japan, pp. 395-400.

[23] Kresse, W. and Fadaie, K., 2004, ISO Standard for Geographic Information, Springer-Verlag Berlin, Germany, pp. 93-96

[24] Licensing Geographic Data, 2004, Board on Earth Sciences and Resources pp. 19-30, pp.31-38, pp. 157-176.

[25] Longley, P. A. , et. al. 2nd edn., 2005, Geographical Information System and Science, Jonh Wiley & Sons Ltd, Atrium, Southern Gate, Chichester, West Sussex, England.

[26] May, A. J., Ross, T., Bayer, S. H. and Tarkiainen, M. J, 2003, Pedestrian navigation aids: information requirements and design implications. Pers Ubiquit Comput, Springer-Verlag London Limited edn. 7: pp. 331–338

[27] Medhi, D. and Ramasamy K., 2007, Network routing Algorithms, protocols, and architectures, pp. 30-54

[28] Millonig, A. and Gartner, G., 2009, Monitoring Pedestrian Spatio-Temporal Behaviour, citeseer, pp. 1-14

[29] Millonig, A. and Schechtner, K. 2007, Part 2, Decision Loads and Route Qualities for Pedestrians — Key Requirements for the Design of Pedestrian Navigation Services, pp. 109-118.

[30] Möser, S., Degener, P., Wahl, R. and Klein R.,2008, *Context Aware Terrain Visualization for Wayfinding and Navigation*. The Eurographics Association and Blackwell Publishing Ltd. Vol. 27, Number 7 pp. 1853-1861

[31] Official Ubuntu Documentation. https://help.ubuntu.com/ [visited-Sept/2010]

[32] Open Geospatial Consortium. http://www.opengeospatial.org. [visited Oct/2010]

[33] OpenRouteService. http://www.openrouteservice.org/index.php [visited Nov/2010]

[34] OpenStreeMap. http://wiki.openstreetmap.org/wiki/Main_Page [visited-Aug/2010- Feb/2011

[35] Peuquet, D. J., 1983, A Hybrid Structure for the Storage and Manipulation of Very Large Spatial Datasets, pp. 11-12

[36] pgAdmin. http://www.pgadmin.org/ [visited Jul, 2010 – Jan, 2011]

[37] pgRouting. http://pgrouting.postlbs.org. [visited 24/Jul/2010]

[38] PHP – Hypertext Preprocessor - http://php.net/ [visited Nov, 2010 – Feb, 2011]

[39] PostGIS. http://postgis.refractions.net/. [visited 24/Jul/2010]

[40] Rigaux, P., Scholl, M., and Voidard, A., 2002, Spatial Database with Application to GIS pp. 100–107.

[41] Seo, W-J, et. al. 2009, An Efficient Hardware Architecture of the A-star Algorithm for the Shortest Path Search Engine, pp. 1499-1502.

[42] Shekhar, S. and Chawla, S. 2003, *Spatial Database: A Tour*. Pearson Education, Inc. New Jersey.

[43] Spatial Database, File System, Relational Database. Wikipedia, the free encyclopedia http://en.wikipedia.org/wiki/. [visited 22/Jul/2010]

[44] Sturgis, E, 1970, Natural and Programming Languages, pp. 1-12

[45] Swedish Association of Local Authorities and Regions (SALAR) - http://english.skl.se/ visited [Mar., 2011]

[46] Swedish Road Administration, Svd department and Mattsson, M.-O., 2008a, 'Content Overview', Swedish NVDB Nationell Vågdatabas version 5.3 pp. 1-14.

[47] Swedish Road Administration, Svd department and Mattsson, M.-O., 2008b, 'Road Network', Swedish NVDB Nationell Vågdatabas version 5.3 pp. 1-37.

[48] Swedish Road Administration, Svd department and Mattsson, M.-O., 2008c, 'Feature Types', Swedish NVDB Nationell Vågdatabas version 5.3 pp. 1-111

[49] Swedish Road Administration, Svd department and Mattsson, M.-O., 2008c, 'Products and Services Catalogue', Swedish NVDB Nationell Vågdatabas pp. 1-24

[50] What are Geographic Data? http://www.spatial.maine.edu/~max/EDBT1.html. [visited 22/Jul/2010]

[51] Wiltschko, T, and Kaufmann, 2002, A quality model for assurance of road information pp. 1-3.

[52] Worboys, M. and Duckham, M. (2nd Edition) 2004, *GIS A Computing Perspective.* CRC Press LLC U.S.A.

[53] Xia, K., Wei, C., 2008, Study on Real-time Navigation Data Model Based on ESRI Shapefile, pp. 174-178.

Lunds Universitets Naturgeografiska institution. Seminarieuppsatser. Uppsatserna finns tillgängliga på Naturgeografiska institutionens bibliotek, Sölvegatan 12, 223 62 LUND. Serien startade 1985. Hela listan och själva uppsatserna är även tillgängliga på http://www.geobib.lu.se/

The reports are available at the Geo-Library, Department of Physical Geography, University of Lund, Sölvegatan 12, S-223 62 Lund, Sweden.
Report series started 1985. The whole complete list and electronic versions are available at http://www.geobib.lu.se/

170 Mats Carlsson (2009): Aerosolers påverkan på klimatet.
171 Elise Palm (2009): Övervakning av gåsbete av vass – en metodutveckling
172 Sophie Rychlik (2009): Relating interannual variability of atmospheric CH_4 growth rate to large-scale CH_4 emissions from northern wetlands
173 Per-Olof Seiron and Hanna Friman (2009): The Effects of Climate Induced Sea Level Rise on the Coastal Areas in the Hambantota District, Sri Lanka - A geographical study of Hambantota and an identification of vulnerable ecosystems and land use along the coast.
174 Norbert Pirk (2009): Methane Emission Peaks from Permafrost Environments: Using Ultra–Wideband Spectroscopy, Sub-Surface Pressure Sensing and Finite Element Solving as Means of their Exploration
175 Hongxiao Jin (2010): Drivers of Global Wildfires — Statistical analyses
176 Emma Cederlund (2010): Dalby Söderskog – Den historiska utvecklingen
177 Lina Glad (2010): En förändringsstudie av Ivösjöns strandlinje
178 Erika Filppa (2010): Utsläpp till luft från ballastproduktionen år 2008
179 Karolina Jacobsson (2010):Havsisens avsmältning i Arktis och dess effekter
180 Mattias Spångmyr (2010): Global of effects of albedo change due to urbanization
181 Emmelie Johansson & Towe Andersson (2010): Ekologiskt jordbruk - ett sätt att minska övergödningen och bevara den biologiska mångfalden?
182 Åsa Cornander (2010): Stigande havsnivåer och dess effect på känsligt belägna bosättningar
183 Linda Adamsson (2010): Landskapsekologisk undersökning av ädellövskogen i Östra Vätterbranterna
184 Ylva Persson (2010): Markfuktighetens påverkan på granens tillväxt i Guvarp
185 Boel Hedgren (2010): Den arktiska permafrostens degradering och metangasutsläpp
186 Joakim Lindblad & Johan Lindenbaum (2010): GIS-baserad kartläggning av sambandet mellan pesticidförekomster i grundvatten och markegenskaper
187 Oscar Dagerskog (2010): Baösberggrottan – Historiska tillbakablickar och en lokalklimatologisk undersökning
188 Mikael Månsson (2010): Webbaserad GIS-klient för hantering av geologisk information
189 Lina Eklund (2010): Accessibility to health services in the West Bank, occupied Palestinian Territory.
190 Edvin Eriksson (2010): Kvalitet och osäkerhet i geografisk analys - En studie om kvalitetsaspekter med fokus på osäkerhetsanalys av rumslig prognosmodell för trafikolyckor
191 Elsa Tessaire (2010): Impacts of stressful weather events on forest ecosystems in south Sweden.

192 Xuejing Lei (2010): Assessment of Climate Change Impacts on Cork Oak in Western Mediterranean Regions: A Comparative Analysis of Extreme Indices
193 Radoslaw Guzinski (2010): Comparison of vegetation indices to determine their accuracy in predicting spring phenology of Swedish ecosystems
194 Yasar Arfat (2010): Land Use / Land Cover Change Detection and Quantification — A Case study in Eastern Sudan
195 Ling Bai (2010): Comparison and Validation of Five Global Land Cover Products Over African Continent
196 Raunaq Jahan (2010): Vegetation indices, FAPAR and spatial seasonality analysis of crops in southern Sweden
197 Masoumeh Ghadiri (2010): Potential of Hyperion imagery for simulation of MODIS NDVI and AVHRR-consistent NDVI time series in a semi-arid region
198 Maoela A. Malebajoa (2010): Climate change impacts on crop yields and adaptive measures for agricultural sector in the lowlands of Lesotho
199 Herbert Mbufong Njuabe (2011): Subarctic Peatlands in a Changing Climate: Greenhouse gas response to experimentally increased snow cover
200 Naemi Gunlycke & Anja Tuomaala (2011): Detecting forest degradation in Marakwet district, Kenya, using remote sensing and GIS
201 Nzung Seraphine Ebang (2011): How was the carbon balance of Europe affected by the summer 2003 heat wave? A study based on the use of a Dynamic Global Vegetation Model; LPJ-GUESS
202 Per-Ola Olsson (2011): Cartography in Internet-based view services – methods to improve cartography when geographic data from several sources are combined
203 Kristoffer Mattisson (2011): Modelling noise exposure from roads – a case study in Burlövs municipality
204 Erik Ahlberg (2011): BVOC emissions from a subarctic Mountain birch: Analysis of short-term chamber measurements.
205 Wilbert Timiza (2011): Climate variability and satellite – observed vegetation responses in Tanzania.
206 Louise Svensson (2011): The ethanol industry - impact on land use and biodiversity. A case study of São Paulo State in Brazil.
207 Fredrik Fredén (2011): Impacts of dams on lowland agriculture in the Mekong river catchment.
208 Johanna Hjärpe (2011): Kartläggning av kväve i vatten i LKAB:s verksamhet i Malmberget år 2011 och kvävets betydelse i akvatiska ekosystem ur ett lokalt och ett globalt perspektiv
209 Oskar Löfgren (2011): Increase of tree abundance between 1960 and 2009 in the treeline of Luongastunturi in the northern Swedish Scandes
210 Izabella Rosengren (2011): Land degradation in the Ovitoto region of Namibia: what are the local causes and consequences and how do we avoid them?
211 Irina Popova (2011): Agroforestry och dess påverkan på den biofysiska miljön i Afrika.
212 Emilie Walsund (2011): Food Security and Food Sufficiency in Ethiopia and Eastern Africa.
213 Martin Bernhardson (2011): Jökulhlaups: Their Associated Landforms and Landscape Impacts.
214 Michel Tholin (2011): Weather induced variations in raptor migration; A study of raptor migration during one autumn season in Kazbegi, Georgia, 2010

215 Amelie Lindgren (2011) The Effect of Natural Disturbances on the Carbon Balance of Boreal Forests.

216 Klara Århem (2011): Environmental consequences of the palm oil industry in Malaysia.

217 Ana Maria Yáñez Serrano (2011) Within-Canopy Sesquiterpene Ozonolysis in Amazonia

218 Edward Kashava Kuliwoye (2011) Flood Hazard Assessment by means of Remote Sensing and Spatial analyses in the Cuvelai Basin Case Study Ohangwena Region –Northern Namibia

219 Julia Olsson (2011) GIS-baserad metod för etablering av centraliserade biogasanläggningar baserad på husdjursgödsel.

220 Florian Sallaba (2011) The potential of support vector machine classification of land use and land cover using seasonality from MODIS satellite data

221 Salem Beyene Ghezahai (2011) Assessing vegetation changes for parts of the Sudan and Chad during 2000-2010 using time series analysis of MODIS-NDVI

222 Bahzad Khaled (2011) Spatial heterogeneity of soil CO2 efflux at ADVEX site Norunda in Sweden

223 Emmy Axelsson (2011) Spatiotemporal variation of carbon stocks and fluxes at a clear-cut area in central Sweden

224 Eduard Mikayelyan (2011) Developing Android Mobile Map Application with Standard Navigation Tools for Pedestrians

225 Johanna Engström (2011) The effect of Northern Hemisphere teleconnections on the hydropower production in southern Sweden

226 Kosemani Bosede Adenike (2011) Deforestation and carbon stocks in Africa

227 Ouattara Adama (2011) Mauritania and Senegal coastal area urbanization, ground water flood risk in Nouakchott and land use/land cover change in Mbour area

228 Andrea Johansson (2011) Fire in Boreal forests

229 Arna Björk Þorsteinsdóttir (2011) Mapping *Lupinus nootkatensis* in Iceland using SPOT 5 images

230 Cléber Domingos Arruda (2011) Developing a Pedestrian Route Network Service (PRNS)