Uses of the OpenStreetMap network for route and travel cost calculation

Mireia Farré (Spain), Begoña Admirable (Spain), Cristina Rovira (Spain)

Abstract

Statistical data contribute to decision-making, research and improving public policies, but their reuse also allows urgent societal demands to be addressed.

The aim of this presentation is to demonstrate a reuse of statistical data concerning topics where mobility and travel costs over time are key factors, such as school planning or accessibility to health facilities.

For this purpose, official geolocated population data were combined with the information derived from the OpenStreetMap (OSM) application, a collaborative project to create editable and free maps. The main output of the OSM project consists of data stored in the form of raster images and vector data, creating a database (planet.osm) in the PBF binary file format.

Using OSM data (streets and roads accessible on foot and by car, as well as public transport network data), a graph was created to model networks with the objective of calculating routes from an origin point to a destination point.

The pgRouting extension of postGIS was used to create the graph and also to evaluate the data quality of the OSM project for each type of network.

Keywords: OpenStreetMap, Mobility and travel time, Geolocated population data

1 Introduction

The aim of the project is to facilitate the reuse of statistical data to quickly respond to demands where mobility and travel costs are key factors.

The Territorial Statistical Register (RET) is an administrative register comprised of georeferenced postal addresses from all units of other administrative registers that form the basis of the statistical production system of Idescat: the Population Statistical Register (REP) and the Statistical Register of Companies and Entities (REE).

Despite producing this basic information, Idescat did not have the tools to address the increasingly frequent requests for route calculations in the territory and their travel costs.

2 Uses and quality of the OpenStreetMap network for route and travel cost calculation.

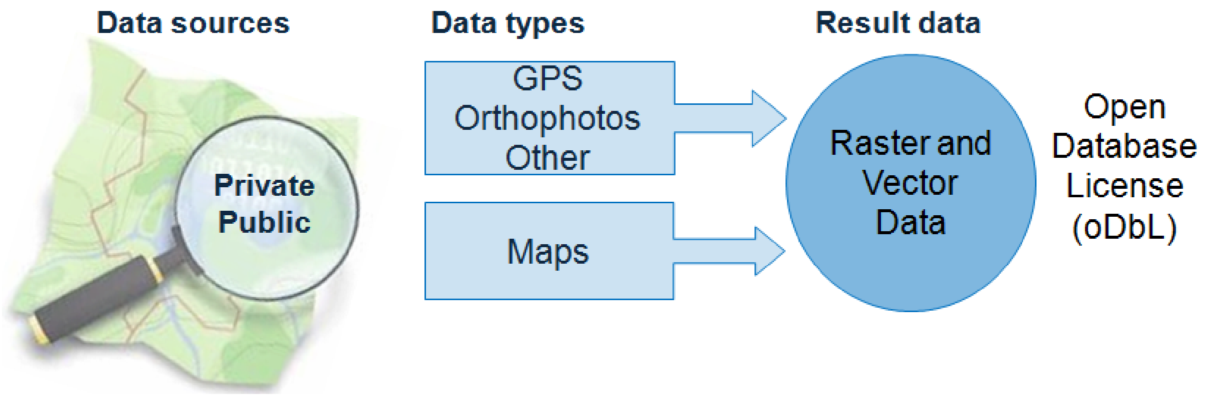
OpenStreetMap (OSM) is a global collaborative project that creates and provides free and open geospatial data obtained through mobile GPS devices, orthophotos and other sources. It draws on user collaboration to collect, update, and maintain information about roads, footpaths, cafes, train stations, rivers, and various geographical entities.

The map data and images obtained are distributed under the Open Database License (ODbL), allowing any user to freely use and redistribute them.

2.1 Data downloading

All OSM information (planet.osm) is updated weekly and can be downloaded in a single XML file or in a more compact binary format with the PBF extension, the latter being smaller and faster to process (~75GB, CAT~390MB). Additionally, tools like OSM2PO or QuickOSM (a QGIS plugin) can be used to select elements of interest such as streets and roads or public transportation networks. Final files contain the complete road and transportation networks represented by nodes, ways and relations.

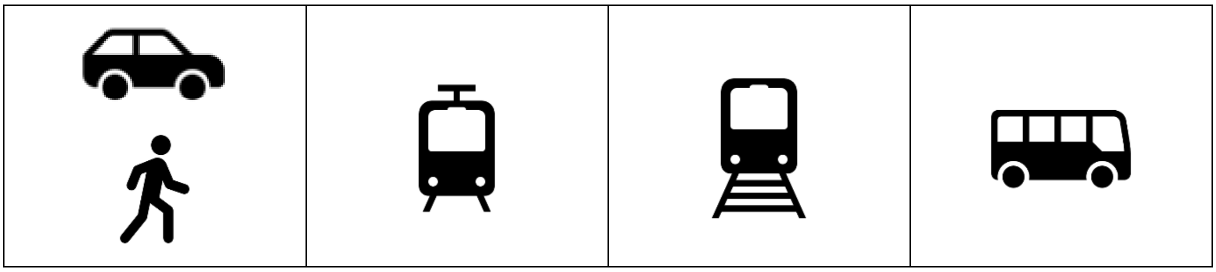
Figure 1: OSM data scheme



2.2 Modelling transport networks: directed graphs

The OSM file is made up of a collection of geographical segments that, in addition to these segment geometries, also contain attributes about allowed directions and maximum travel speeds. With all this information, four directed graphs were constructed to model the typology of each of the four transportation networks (road, metro, railway and buses).

Figure 2: Target transportation networks



The four graphs were generated using the pgRouting program, an extension of PostGIS/PostgreSQL that provides geospatial routing functionalities and network analysis capabilities.

Figure 3: Directed graph representation.

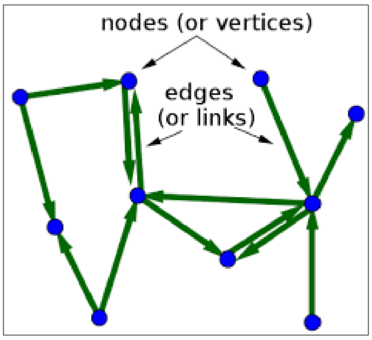
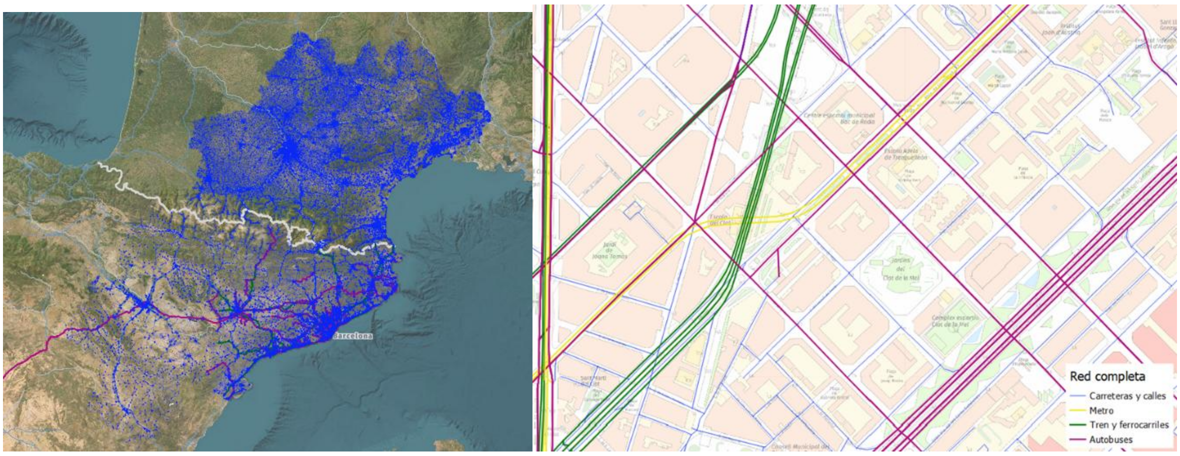


Figure 4: Generated graph and expansion in Barcelona



2.3 Main functions of pgRouting for graph creation and analysis

The main functions used for the directed graphs creation process are:

- *pgr_createTopology and pgr_createVerticesTable* that generate the directed graph and the vertex table.
- *pgr_connectedcomponents* is used to group segments that are connected.

The resulting directed graphs will have acceptable quality as long as they exhibit a minimum number of anomalies such as:

- Isolated or unreachable segments
- Dead-end streets or roads
- Intersections
- Ring geometries

There are pgRouting functions available to analyse the quality of each graph:

- *pgr_analyzeGraph* analyses the graph and reports if there are loops or other anomalous structures.
- *pgr_analyzeOneway* detects flipped segments.

These are the results obtained from the analysis of the 4 networks:

- ✓ The quality of the street and road network is acceptable.
- ✓ Similarly, the quality of the metro network is also acceptable.
- ✓ The train network is acceptable after correcting some erroneous directions.
- × Unfortunately, the bus network exhibits an excess of ring geometries and incorrect directions and thus its use has been discarded.

2.4 Routing and cost calculation

Based on the constructed graphs and using the pgr_bddijkstra function of pgRouting, it is possible to calculate routes and associated travel times between any origin and destination points within each directed graph. The Dijkstra algorithm is a graph search algorithm that resolves the shortest path problem with non-negative edge path costs. The function returns an itinerary as a set of records and their associated costs.

It should be noted that computational costs can be very high when iteratively executing this procedure over large sets of population or households, in our case exceeding 1,000,000 cases.

To make the process more efficient, Idescat has implemented the following optimisation measures:

I. Indexing of tables with cluster-type indexes.
II. Maintenance of the segment and vertex table, such as through periodic vacuum processes.
III. Limiting route calculation to segments and vertices contained within a fixed window that includes the origin and destination points and a certain expansion of it.

Figure 5: Window containing an origin and a destination with different extension values

Table 1: Reduction in execution times based on the extension

Extension	Time (sec)	Time for 1,000,000 records (days)
250	7,476	86
100	3,929	45
40	2,596	30
20	0,826	9
10	0,374	4
5	0,203	2
Without Extension	**8,715**	**92**

IV. Parallel execution with a minimum of 10 threads, managed with Apache NetBeans.

2.5 Uses

The results obtained enable precise distributions of population mobility phenomena. Here are some examples:

2.5.1 School mobility

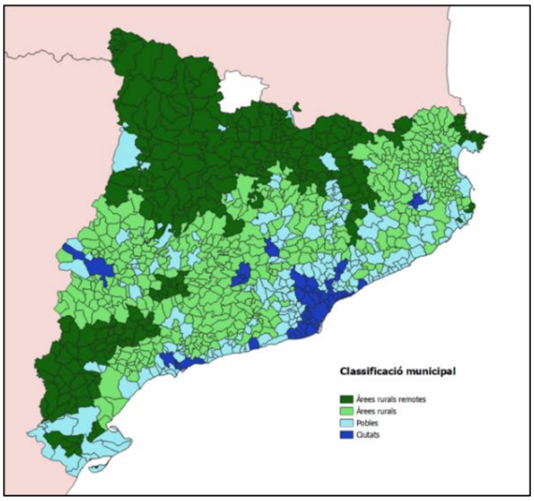
We have collaborated with the Department of Education, providing information on the mobility of non-university students to plan schools and areas. Using the locations of the postal addresses of the students and the schools, indicators were calculated such as the distances and times of the journeys associated with the different means of transport (on foot, by car, by metro/train).

Figure 6: Calculated journeys for a student on foot, by car and by metro/train respectively

2.5.2 Remote Rural Municipalities

Rural municipalities according to the DEGURBA (Degree of Urbanisation) classification are considered remote if the time spent driving to the nearest DEGURBA city is more than 45 minutes. Identifying these municipalities is crucial for promoting public policies to revitalise rural areas.

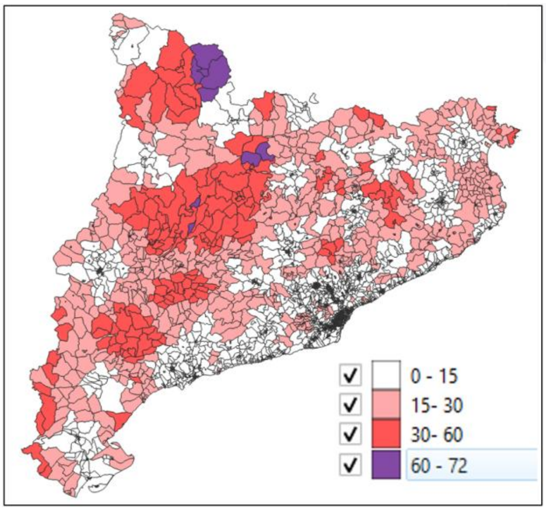
Figure 7: Map with the classification of remote rural municipalities.



2.5.3 Travel times to the nearest hospital

The median travel times to the nearest hospital for the entire georeferenced population of Catalonia have been calculated by municipalities so that we can use the results to get an indicator of accessibility to elements of interest such as hospitals.

Figure 8. Map of median travel times to the nearest hospital



2.6 Conclusions

The calculation of minimum transport routes is a fundamental tool for organising the territory and planning services, and the results obtained allow for greater precision in the distribution of mobility phenomena.

The georeferencing of the population enables low-budget development of applications of great utility and quality when combined with powerful open programming solutions such as OSM.

Despite the high computational costs associated with these processes, today we have greater computing capacity and adequate infrastructures and there are also several techniques to optimise the performance of the intensive use of optimal pathfinding algorithms.

We have new objectives to achieve in the future in terms of the network of cycle lanes and the incorporation of the valid bus network to extend mobility calculations to these means of transport as well.

Acknowledgments

Our thanks to Eduard Suñé, one of the sponsors and mentors of this project.

References

Georeferenced population Idescat:
https://milleniumbeg.cultura.gencat.cat/record=b1701178~S72*cat

OpenStreetMap: https://www.openstreetmap.org

OSM2PO: https://osm2po.de/

QGIS: https://www.qgis.org

QuickOSM: https://plugins.qgis.org/plugins/QuickOSM/

pgRouting: https://pgrouting.org/

PostGIS: https://postgis.net/

PostgreSQL: https://www.postgresql.org/

Apache NetBeans: https://netbeans.apache.org/front/main/index.html