Routing on Geospatial Reference Information for Transport Networks of Spain

Cristina Calvo [a,*], Alicia González, Ángel Expósito

a Spanish National Geographic Institute, Department of Transport Networks, mccalvo@mitma.es; agjimenez@mitma.es; aexposito@fomento.es

Abstract: The Spanish National Geographic Institute (IGN) published the first release of the Geographic Reference Information on Transport Networks (GRI-TN) in March 2017. Its main goal was to fulfil INSPIRE Directive requirements, as well as to become the main data source for other products developed by the IGN regarding this theme. During the years following that first release, the focus has been on updating and improving the data. This fact has encouraged new data use cases, which differ from the initially planned ones, have arisen, allowing to detect problems in the data, and highlighting the need to evolve the data model, as well as the way in which they are provided to users (not only formats, but also update frequency). One of those use cases is finding the shortest paths between different points in the area covered by the Road Transport Network. In this article, the methodology used to do it is exposed; likewise, the setbacks that have come up during the process and the current limitations of the GRI-TN datasets in order to get most accurate results.

Keywords: GRI-TN, Dijkstra, Floyd-Warshall, routing, roads.

1. Introduction

In July 2020, personnel from the Ministry of Ecological Transition and Demographic Challenge (MITECO, hereinafter) contacted IGN to consider the possibility of calculating the cost in time and the distance in kilometres from the capital of each one from the municipalities of Spain to the nearest hospital, the nearest municipality with more than 20.000 inhabitants and the closest municipality with more than 50.000 inhabitants. The reason for this request was based on a study that sought to assess the risk of depopulation in Spanish municipalities, and which took the distance from the various socio-cultural and health facilities and services as one of the factors that a priori could contribute to increasing said risk.

1.1 Approach

In order to identify the data set and the most appropriate tools for resolving the query from among those available at the IGN, we proceeded to analyse in detail the main particularities of the request:

- It was necessary to carry out this route calculation in a massive way (8,131 municipalities, according to Statistics National Institute (INE) data from December 2019).
- In addition, it was necessary to identify, for each starting point (centroid of each municipal capital), the point of interest to which it should go for each of the three types of routes requested.
- It was not only necessary to obtain the cost of each route in kilometres but also in minutes. To do this, it was necessary to estimate the time it takes to cross each of the sections that make up the road network since speed data is not available for each section.

After evaluating the available tools and the state of the data, it was decided to use the sections of the road network collected in the GRI-TN database, located on a PostgreSQL server, using the attributes of said sections such as the type of section (trunk, interchange, service road, etc.), the type of road (highway, highway, conventional road, etc.), the type of road (double or single), etc., to estimate the average speed in them, and the functions of the pgRouting extension to calculate the routes and their associated costs.

2. Methodology

2.1 Input Data

The process described below was carried out based on the following data:

- Hospitals: National Catalogue of Hospitals, provided by MITECO. It is an alphanumeric catalogue that was provided with a specific geometry to perform spatial analysis.
- Municipal capitals:
  o Location: point identified as the capital of the municipality in the Geographic Nomenclature for Municipalities and Population Entities.
  o Population data: official population figures from the Municipal Register as of January 1, 2019, provided on the INE website.
- Routes (linear elements): roads, urban roads, and paths collected in the GRI-TN database.

2.2 Process

In the figure below (which is described in detail in the following sections), are shown schematically the stages followed to obtain the requested routes.
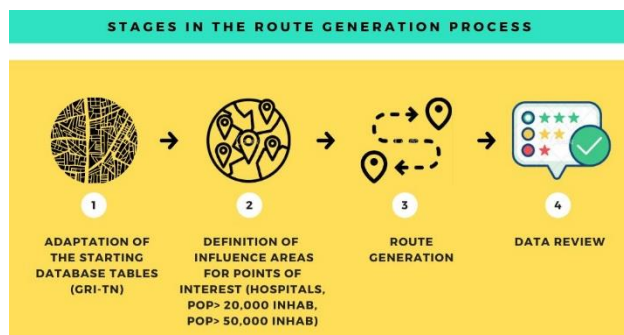
Figure 1: Stages in the route generation process

2.2.1 *Adaptation of the starting database tables (GRI-TN): Node generation.*

The linear elements (segments) of the GRI-TN road network have an implicit topology: they are connected, but these connections are not described or materialized in any way in the database.

To generate the routes from the data collected in the table of road network segments, it is necessary to convert this topology into an explicit topology, i.e. to create nodes at all points where two or more segments connect. This process is done using the pgr_createtopology function on pgRouting, which not only generates these nodes but also identifies the start and end nodes for each section.
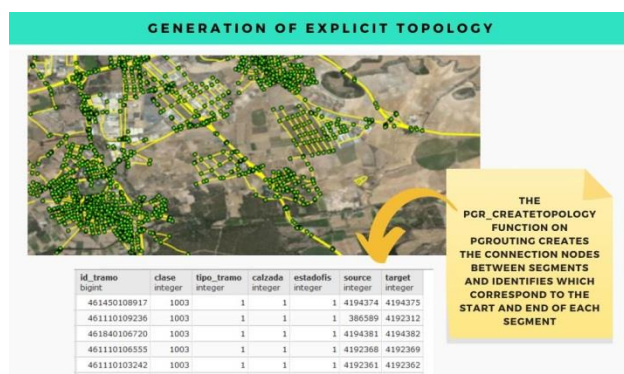

Figure 2: Production of explicit topology

Once the information that describes how the network is connected had been obtained, the next step was to calculate the time required to travel each of the segments, based on their length and the estimated average speed on each type of road. The approximation to this data was obtained in two phases:

- In the first one, a preliminary calculation supported by several studies that propose methodologies and average speeds for the different types of road. It was based on the speeds proposed for each type of road (distinguishing between motorway/highway, conventional road, crossing, urban road, and road or

path), and they were qualified according to whether each section in question is a trunk or not (link, roundabout or service road).

- In the second phase, an evaluation of the speeds obtained was carried out, calculating the cost in kilometres and minutes of various routes from the data of the GRI-TN and comparing them with the results obtained for the same routes in commercial navigation programs. From the results of this comparison, the preliminary speeds were adjusted, finally using the estimates shown in the table below[1],[2].

Type of Road	km/h	Meters/min
Crossing (trunk)	40	666,666667
Crossing (not trunk)	15	250
Urban (trunk)	10	166,666667
Urban (not trunk)	10	166,666667
Highway / motorway (trunk)	100	1666,66667
Highway / motorway (not trunk)	30	500
Named road (trunk)	75	1250
Named road (not trunk)	25	416,666667
Unnamed road (trunk)	60	1000
Unnamed road (not trunk)	25	416,666667
Path	15	250

Table 1: Travel speed estimation by type of road

Besides, a subsequent estimation was carried out for the roads to evaluate the speed based on their slope and the sinuosity of their route, although this data was not finally used to calculate the requested routes.

2.2.2 *Definition of influence areas*

Once the spatial data necessary for the calculation of routes had been obtained (sections of the network, cost of each one, and connection nodes), we proceeded to identify the start and endpoints of each route).

The origin points for all routes, as indicated above, were those identified as a descriptive point of the capital of each municipality in the Geographical Nomenclature for Municipalities and Population Entities published by the IGN.

Since these points are generally not located on segments of the GRI-TN, and therefore they do not coincide with any of the nodes created in the process described above, it was necessary to assign them a node.

Firstly, points were assigned to the closest node. Subsequently, it was found that in many cases the closest node had poor connectivity (it was associated with roads with a very high cost of passage or without connection to other roads in the network), which led to clearly excessive route costs. To avoid this type of errors, a second version

---

[1] Non-trunk segment: junctions, entry or exit roads, roundabouts, and service roads.

[2] The distinction between named roads and unnamed roads is due to the fact that the latter, usually, correspond to roads with

no known title. These are usually forest roads or tracks that have been paved to enable the passage of vehicles, but hardly any maintenance is carried out and they also have a very winding route, both factors that limit the speed greatly.

of the table of linear geometries and nodes was made, in which the shortest connectivity routes and high travel times were eliminated (mainly roads and paths, to force the algorithm to identify only routes in roads and urban roads). From this new table of segments, the nodes with optimal connectivity were identified in an environment of 100-200 meters from the point identified by nomenclature as the centre of the municipality capital. To get this was used the Floyd-Warshall algorithm (implemented in the function pgr_floydWarshall on pgRouting), which calculates all possible routes between a set of segments of a network.

This same process was carried out, for the same reasons, with the points identified by the hospitals.

Any minimum path algorithm requires the identification of the start and end nodes. Consequently, for each of the starting points, it was necessary to find the closest points of interest, that is, according to the objectives of the study, the hospital and the closest municipalities with more than 20,000 and 50,000 inhabitants.

The problem was that this data was precisely part of the missing information. To obtain this, territorial models were generated that show the cost in minutes of going from each point of the territory to the closest point of interest. These models were generated by the interpolation of the nodes that are at a maximum time of 55 minutes from any hospital or 100 minutes from any municipality with more than 20,000 or 50,000 inhabitants. Consequently, the algorithm pgr_drivingdistance was used, which allows calculating the cost from a given node to all nodes that are at less than a given maximum time.

The result was three raster files, one for each set of items of interest. The following figures show part of these cost models.
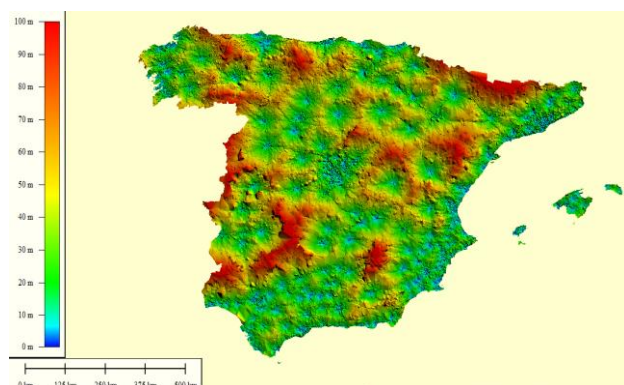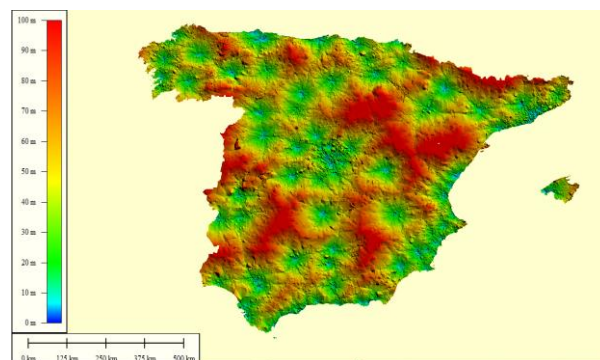

Figure 4: Cost (in minutes) from each point of the territory to the nearest municipality with more than 50,000 inhabitants
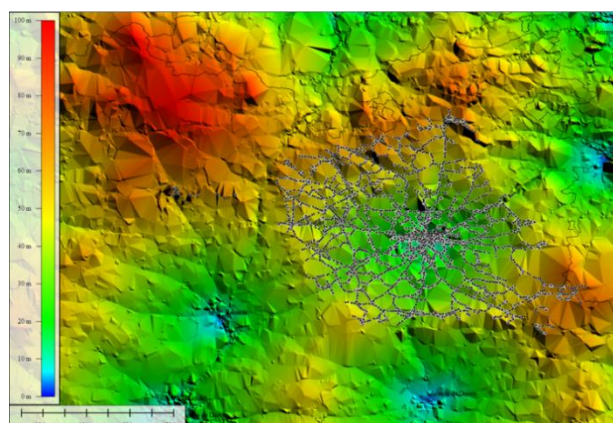

Figure 5: Detail of the cost map of municipalities with more than 20,000 inhabitants (area of influence for the city of Burgos)
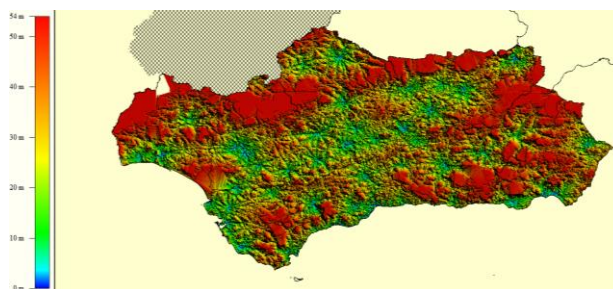

Figure 3: Cost (in minutes) from each point of the territory to the nearest municipality with more than 20,000 inhabitant

Figure 6: Territorial model of costs (min) of distance to the nearest hospital in the Region of Andalusia

2.2.3 Generation of Routes

When the necessary data was available, the Dijkstra algorithm (pgr_dijkstra on pgRouting) was executed to obtain the final routes.

In this stage, the main problem to be faced was the performance of the database, since the tables of segments and nodes had, respectively, 7 and 5 million records and the calculation of each of the routes was too slow to make the study feasible. This aspect was solved by working on views of the section and node tables, dynamically generated for each of the routes from a variable size buffer

depending on the distance between the start and end points of the route.

Once this is completed, each calculation time of each route has been less than 5 seconds in more than 90% of the cases, approaching 30-35 seconds in the cases in which a greater number of segments came into play.

2.2.4 *Review of Routes obtained*

After the routes were calculated, those with the highest cost were reviewed, as well as the municipalities where it was not possible to calculate any of the routes.

Regarding the higher-cost routes, errors were sometimes detected in the network or in the definition of the starting or ending point of the route, and consequently, they were corrected. In the same way, and since the table that did not make use of the roads and paths of the network was finally used, the results were compared with those that would have been obtained using the paths, to identify the cases where this change could have meant choosing a significantly more expensive path (more than 5 minutes slower). These cases were analysed and the route that could be considered the most reliable was chosen (usually, after comparing it with route calculation results from other commercial programs).

In the cases of the municipalities for which it was not possible to calculate the route, these were analysed based on two reasons:

a) Island municipalities located on islands that do not have populations of 20,000/50,000 inhabitants (therefore, it is not possible to calculate the route from the road network; it would be possible to calculate the route to the nearest port or heliport, and from there, estimate the time to heliports or to ports on other islands that do have populations of more than 20,000 or 50,000 inhabitants). As an example, in the figure below it can be seen that the island of Fuerteventura does not have a population higher than 50,000.

Figure 7: Population centres of more than 20,000 (green) / 50,000 (red) inhabitants the islands of Tenerife, Gran Canaria, Fuerteventura, and Lanzarote

b) The singular case of the municipality of Llivia (Gerona), which is surrounded by French territory. Given that the IGN only has data on the road network at the Spanish level, it has not been possible to obtain the cost of the routes from the GRI-TN database.

Figure 8: Situation and limits of the Gerona municipality of Llivia and the nearest municipalities

3. Results

The research made it possible to obtain, mainly, two sets of data:

- On one hand, the cost maps described in section 2.2.2, from which other derivative products have been obtained, such as isochrone maps (lines that join points of territory at the same distance as the nearest point of interest).
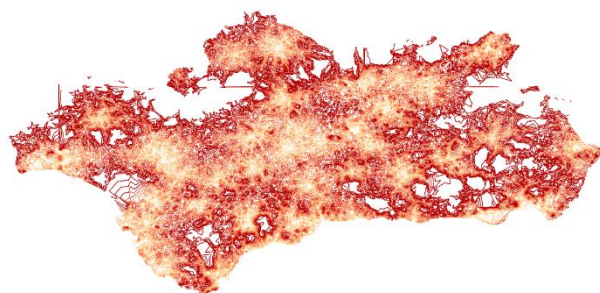

Figure 9: Overall picture of isochrone map of distances to hospitals in the Region of Andalusia
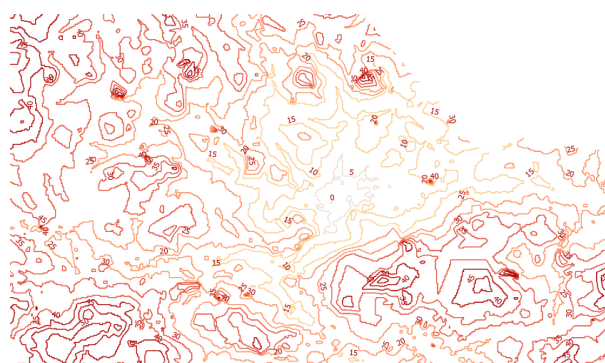

Figure 10: Detail of isochrones for distances to hospitals in the Region of Andalusia

- On the other hand, routes of minimum cost. These tables generated in PostGIS (one for each route) after applying the Dijkstra algorithm do not have geometry but rather refer to the segments through which the route runs (which do have geometry in the source table). To facilitate the review and visualization of the results by the information requestor, they were

exported to shapefile, excel, and kmz files with the following structure:

- o INE code, name, and town of the municipality from which the route is calculated.
- o Name and code of the destination hospital (according to the data of the National Hospital Catalogue) or INE code, name, and population of the nearest municipality with more than 20,000/50,000 inhabitants.
- o Cost data of the route between both municipalities, in minutes and in kilometres.

ine_orig	pob19_orig	nom_orig	codcnh_des	nom_dest	coste_min	coste_km
1001	2876	Alegría-Dulantzi	10066	HOSPITAL DE CUIDADOS SAN ONOFRE SL	23,979	17,667
1002	10350	Amurrio	480078	HOSPITAL UNIVERSITARIO BASURTO	29,156	33,512
1003	1467	Aramaio	200198	HOSPITAL AITA MENNI	3,768	3,966
1004	1821	Artziniega	480078	HOSPITAL UNIVERSITARIO BASURTO	24,770	27,388
1011	291	Baños de Ebro/Mañueta	10088	HOSPITAL DE LEZA	7,577	8,931

Table 2: Sample of the routing table to the nearest hospital

4. Conclusions

This research shows a particular case of use of GRI-TN road network data. With its dissemination, the aim is to show one of the possible use cases of the information collected in GRI-TN, as well as the issues detected and how they were solved.

However, it is necessary to take into account the limitations of the starting data, which may have an impact on the results obtained:

- The segments of the GRI-TN do not have information regarding the direction or turning restrictions. This makes the results obtained (in terms of both of cost in km and cost in time) are, in general, slightly optimistic.
- Currently the updating phase of the GRI-TN road data is being completed, so that soon there could be new data that would change some of the times obtained. Although in practically all the cases it would not affect significantly, it is necessary to point out that during the period in which this study was carried out, the road network in the regions of Catalonia and Andalusia was being updated, also with the incorporation of several new high-speed roads to the GRI-TN.

In general terms, results have been positive, insofar as it has allowed the road network data collected in the IGR-RT to be tested, and above all, it opens the door to the possibility of offering new services to citizens. And not so much those related to the calculation of the optimal route (for which there are numerous commercial programs, both free and commercial use), but mainly concerning the calculation of isochrones, which can be useful for many users and institutions.

5. References

Spanish National Geographic Institute (IGN) (2021). Geographic Nomenclature for Municipalities and Population Entities in Spain. Retrieved from https://www.ign.es/web/ign/portal/rcc-nomenclator-nacional

PgRouting Contributors - Version v3.0.5 (2021). Pgr_createTopology. Retrieved from https://docs.pgrouting.org/3.0/es/pgr_createTopology.html.

PgRouting Contributors - Version v2.4.2 (2017). Pgr_floydWarshall. Retrieved from https://docs.pgrouting.org/2.4/en/pgr_floydWarshall.html.

PgRouting Contributors - Version v3.1.3 (2021). Pgr_drivingDistance. Retrieved from https://docs.pgrouting.org/3.1/en/pgr_drivingDistance.html.

PgRouting Contributors - Version 2.0.0 (0a53b89 master) (2013). Pgr_dijkstra - Camino más corto de Dijkstra. https://docs.pgrouting.org/2.0/es/src/dijkstra/doc/index.html.

González Alicia, Calvo Cristina: "Información geográfica de referencia de redes de transporte: soporte para la implementación de referenciación lineal» (RT) del IGN". X Jornadas Ibéricas de Infraestructuras de Datos Espaciales (JIIDE2019). ["Mapping" Magazine, Vol. 29, Nº 199, 2020, ISSN: 1131-9100]

Mas, Sebastián; Núñez, Eduardo; González Alicia; Sevilla, Celia <<La Información Geográfica de Referencia como base de la cartografía temática y de la Base Topográfica Nacional 1:25.000>>. Los mapas de la Naturaleza. Memorias de la Real Sociedad Española de Historia Natural. Segunda época. Tomo XIV. 2017. ISBN: 978-84-697-4765-0

González A., Velasco A., Camón L., Andrés G., Trigo P.: "Generación de la Información Geográfica de Referencia de «Infraestructuras, Instalaciones y Redes de Transporte» (RT) del IGN". VI Jornadas Ibéricas de Infraestructuras de Datos Espaciales (JIIDE2015). ["Mapping" Magazine Vol. 25, Nº 176, 2016, ISSN: 1131-9100)]