POLITECNICO MILANO 1863

SCUOLA DI INGEGNERIA INDUSTRIALE
E DELL'INFORMAZIONE

Esselunga Logistics Network

PROJECT WORK NETWORK ANALYSIS FOR MOBILITY AND TRANSPORTATION

Beretta Chiara (10615536); Cattaneo Luca (10521219); Fappanni Filippo (10571644)

Index

1. Introduction

The aim of this analysis is to provide an assessment of the logistics network to focus on the freight transportation, instead of classical analysis of passenger network.

The case study is about the logistics network of Esselunga, in particular the area served by the distribution center of Pioltello. The main stakeholder that can be identified from this analysis is Esselunga. Another stakeholder is represented by all the customers affected by the management and the resilience of the supply network.

The report is structured with firstly the assessment of the route assignment problem and then the classical network analysis.

Lastly a What-if approach has been adopted about the availability of more data and the possibility of a second distribution center.

2. Construction of the network

2.1 The Data

The data come from Esselunga and Open Data Regione Lombardia. In particular:

- Esselunga
  - The position of the distribution center (Pioltello)
  - The point of sales supplied by the distribution center
  - The flow and frequency of resupply
    - From 2 to 4 trucks per day for the general merchandising
    - From 2 to 3 trucks per day for fresh food
- Open Data Regione Lombardia [1], [2]
  - More detailed data of the point of sales, i.e., surface and geographical position.

After collecting the data, a cleaning and a merge of them has been made. This part took time because the reported address and the geographical coordinates provided by Regione Lombardia were different with respect to the actual position of the point sales.

After this part, a dataset with the different point of sales has been build, also containing information of the distribution center in the last row.

2.1.1 Assigning the flow

As said before, a supply frequency has been given by Esselunga, but without a direct correlation with each point of sale. To assign that the previous interval has been divided, respectively, in twenty and ten part and assigned proportional to each point sales based on the surface.

2.2 The assignment of the route

The start of the process was finding a way for assigning the route from the distribution center to each point of sales. Two ways has been found:

1. Using QGIS and the shapefile from Open Street Map
2. Using the stplanr R package

2.2.1 QGIS

Firstly, the shapefile downloaded from geofrabrik [3] has been imported with the dataset of the point of sales. Before starting the algorithm of QGIS, that find the shortest path, a check of the shapefiles has been made. A lot of uncategorized roads has been discovered. This is a problem because if these routes are removed (as the route not accessible for truck) the algorithm does not produce a result. Firstly, the shapefile downloaded from geofrabrik [3] has been imported with the dataset of the point of sales. Before starting the algorithm of QGIS, that find the shortest path, a check of the shapefiles has been made. A lot of uncategorized roads has been discovered. This is a problem because if these routes are removed (as the route not accessible for truck) the algorithm does not produce a result. If those routes are added, the truck will pass also in pedestrian area that are not recognized as that.

Therefore, this way cannot be used because the grade of approximation cannot be estimated.

2.2.2 With R

The other way has been the use of stplanr. This package has been cited by Lovelace [4] as useful tool for assign the route.

With *route()* [5] the problem can be easily solved. The input asked by the function are the geographical coordinates of the starting point and of the ending point. It also asks a routing function and in this case the choice go to the free osrmRoute[6] function that compute the quickest path [7]between the points and it return a sf data.

Due to the existence of only three standard profiles (car, foot, bike), an approximation has been made by considering the truck as a car. In Figure 1 the results with a focus on the zone near the distribution center.

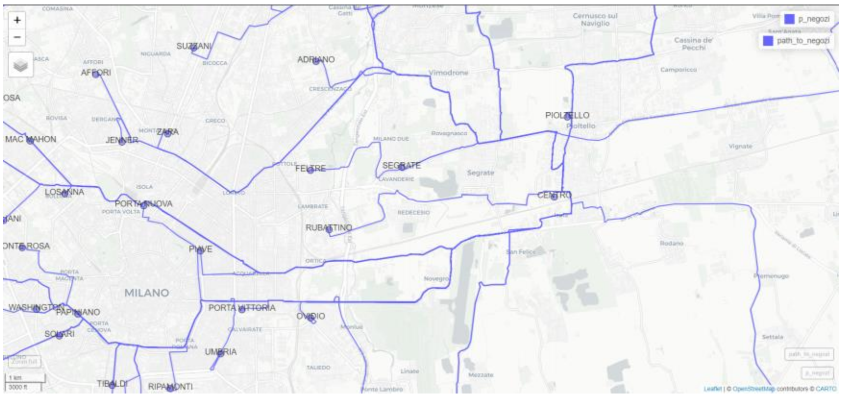


Figure 1 Result from stplanr

2.3 From the physical to the logical network

Given the previous results and with the help of QGIS the intersection between different paths has been found. Those intersection are represented by the "fictitious node". In Figure 2 the result.

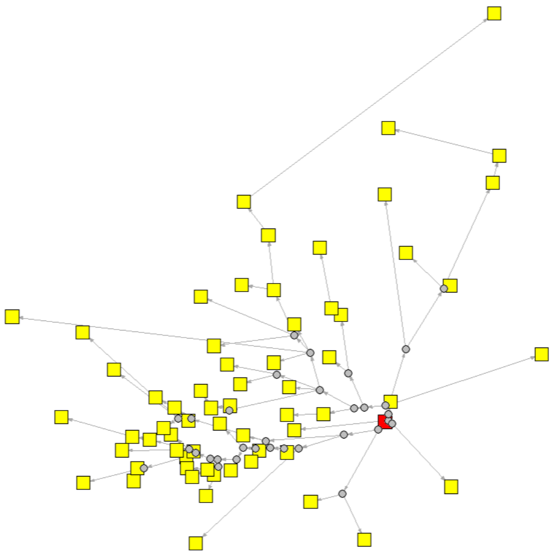


Figure 2 Network (in red the distribution center, in yellow the point of sale, in gray the fictitious node)

3. Analysis

3.1 Macro

Macroanalysis studies the structure of the network, using five indicators to characterized it:

- The **size** of the network is equal to **96** nodes. Precisely the nodes can be distinguished between **32 fictitious, 63 point of sales and 1 distribution center**, as shown in Figure 2.
- The **number of components** is equal to **1**, that means there is only one connected subnetwork

- To measure the maximum distance between pairs of vertices the parameter to consider is the **diameter** of the network. In the case study the diameter intended as spatial distance is equal to **60,5 km**, while the temporal distance is equal to **64,98 minutes**.
- The **density** is equal to **0.011**. The parameter is a value between 0 and 1 and it gives a sign of the frequency of realized edges relative to potential edges. The resulting value is very low, and it indicates a sparse network, due to how the network was built. The network is made up of connections between the distribution center and each supermarket and therefore it has only one possible route for each journey.
- The **clustering coefficient** is equal to **0.019** and it gives an information about the redundancy of the connections that exists in the network, so an information like density parameter. The resulting clustering coefficient is very low, indicating a very fragile network.

3.2 Micro

Microanalysis investigates the importance of single nodes, and it is expectable that results will reflect the fact that nodes in this network are strongly hierarchical (distribution center and shops).

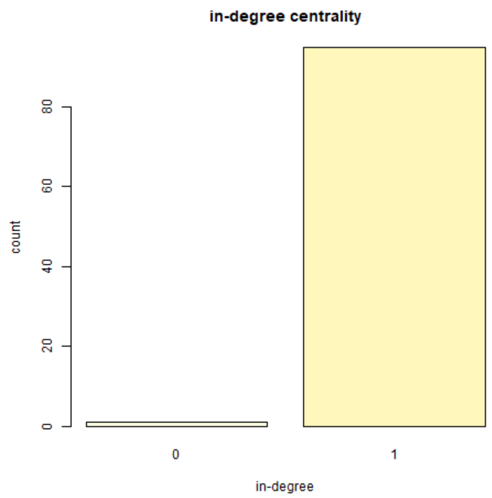


Figure 4: in-degree centrality

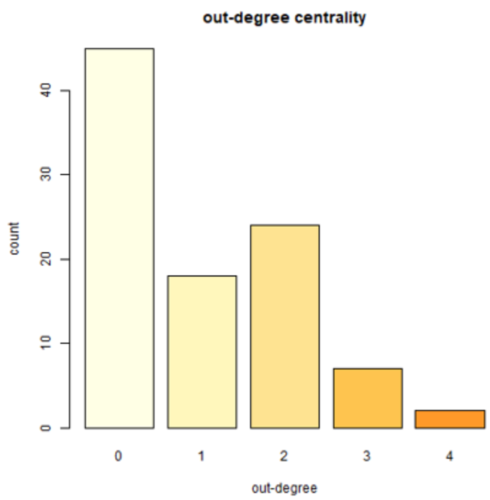
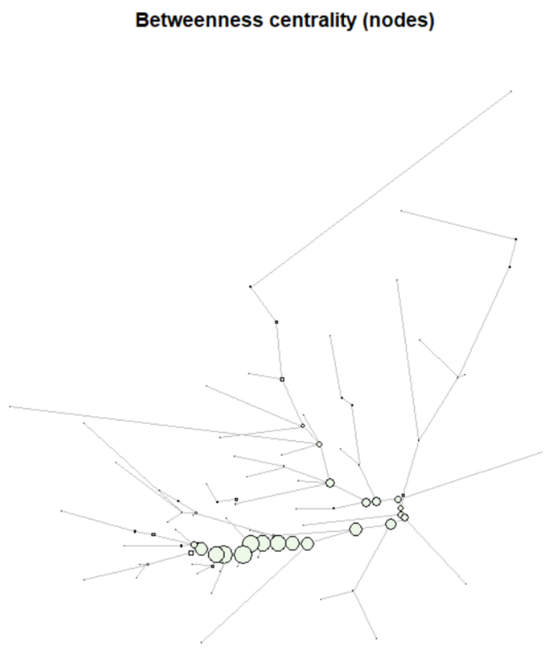
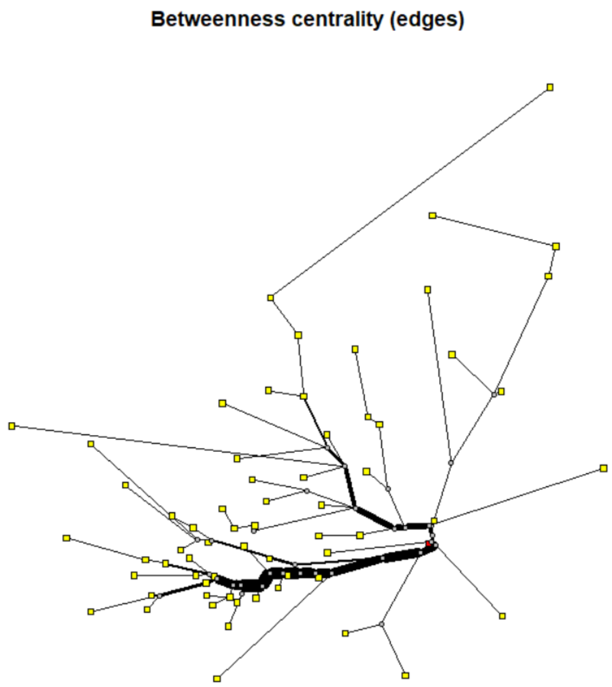


Figure 3: out-degree centrality

- **in-degree centrality** (Figure 4) is equal to 1 for all nodes representing supermarkets and it is equal to 0 for the distribution center. This is coherent with the structure of the network, as edges are directed from the distribution center to shops with only one in-coming edge per supermarket.

- **out-degree centrality** (Figure 3) is equal to 0 for almost all shops, while the other values is for the distribution center or the deviation points. However, few nodes have an out-degree centrality value of 2 and 3, as some shops are located on the same route, meaning that they are connected to the distribution center with a path going through other shops-nodes.
- **Betweenness centrality**: nodes and edges with the highest betweenness value corresponds to the path most used by trucks, as there is the highest concentration of shops. On the map, it corresponds to Viale Forlanini and the so called *Circonvallazione*.





3.3 Meso

With the aim to find the most similar point of sales a focus on the community structure has been made. In particular, the following analysis have been made:

- Community structure detection based on edge betweenness
- Community structure detecting based on the leading eigenvector of the matrix

Both will introduce a small recap on the definition of the used method.

3.3.1 Community structure detection based on edge betweenness

This algorithm computes the edge betweenness (the number of shortest paths through it) of the graph, remove the one with the highest score and then remake the same process iteratively.[8]

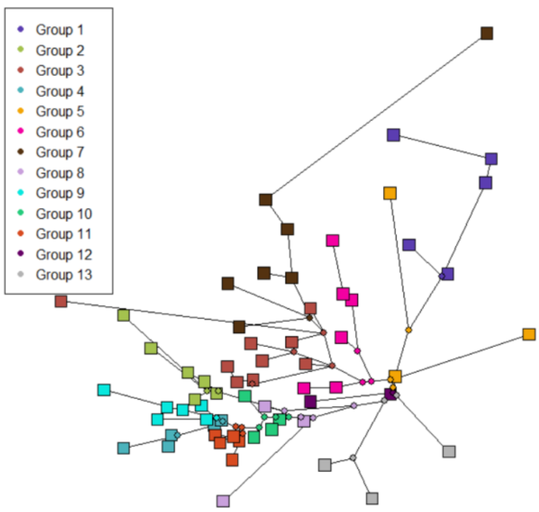


Figure 5 Community structure detection based on edge betweenness

In the Figure 5 the result of the algorithm has been presented. The result is weighted with respect to the distance. 13 groups can be identified.

The first think that can be noted is that the communities funded by the algorithm are compliant with their geographical position.

3.3.2 Community structure detecting based on the leading eigenvector of the matrix

To check if the result change by changing the algorithm the implement leading eigenvector' method developed by Mark Newman, cited in the book as iterative approach for spectral partitioning[9], [10].

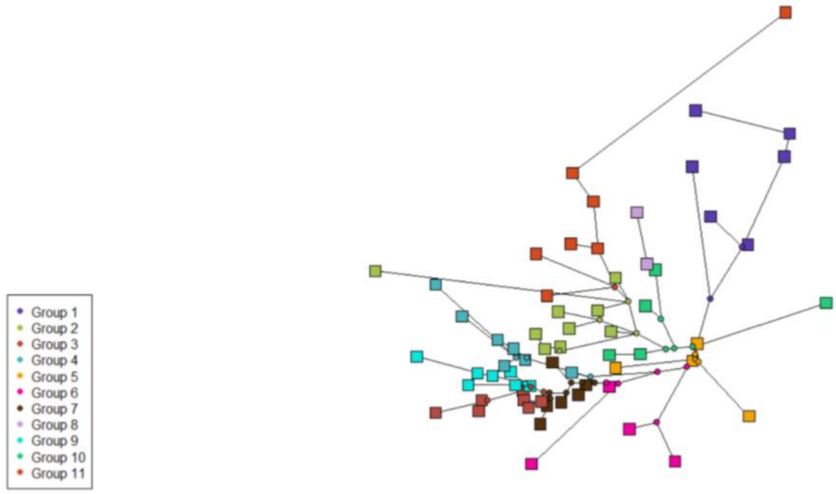


Figure 6 Community detection with Newman method

In this case the number of groups reduce at 11. A more interesting thing here is provided by the dendrogram that recognize the deviation points and at the first level of the tree separate the fictitious nodes and the real nodes (with the previous algorithm this does not happen).

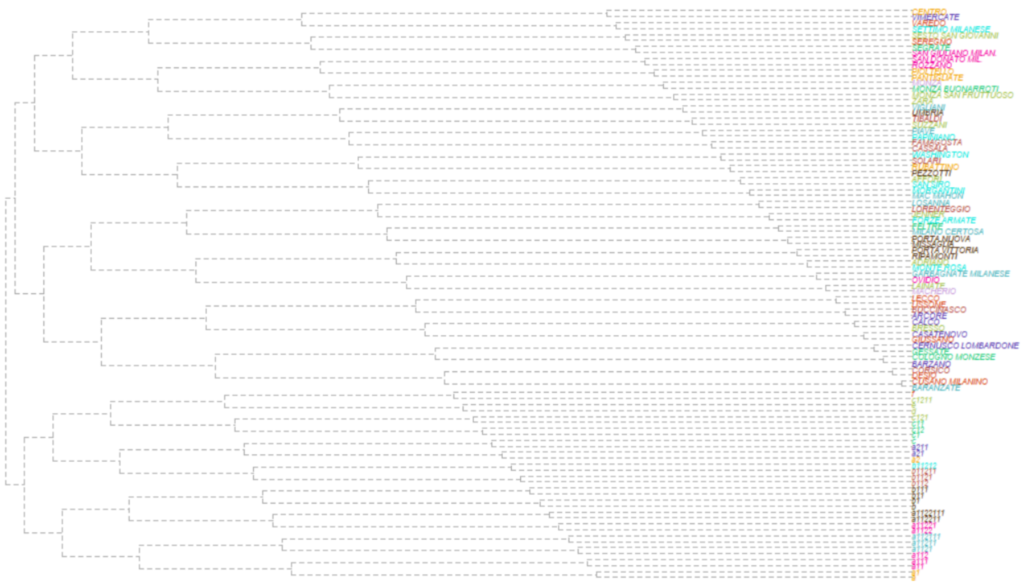


Figure 7 Dendrogram

The two methods presented before can be useful for Esselunga to decide the shops that in case of emergency can supply other one (Example: Corsico and Lorenteggio are in the same group if Corsico goes out of source it can be supplied by Lorenteggio).

4. What If analysis

In this paragraph some scenarios are presented to broaden the vision and the analysis of the network with a what-if approach.

The first scenario concerns the presence of a second distribution center, which is a plausible scenario if the company decided to divide the network to improve its management and the efficiency in distribution.

The other scenarios are about the possibility to have more detailed data on schedules of supplies and on economic value of the goods transported to each supermarket. Lastly, in paragraph 4.2.3 presented the scenario in which the analysis performed in the project considers only the use of roads for trucks.

4.1 Are there two distribution centers?

Thanks to the result in 3.3.2 a part of the network has been selected to be served by a new distribution center. The points of sales are Zara; Suzzani; Affori; Jenner; Adriano; Lainate; Bresso; Varedo; Seregno; Lissone; Lecco; Giussano; Desio.

Now the position of the new distribution center will be found thank to the Centre of Gravity method presented in the Logistics and Freight Transportation course[11], [12].

4.1.1 Centre of Gravity

This method is useful to find an individual location by considering the nodes (the point of sales in this case), the distances, the volume of products to be delivered.

The objective function in this case is the minimization of the transportation costs:

$$\min(Total\ Cost) = \min_{x,y} \sum_i [F_i \cdot R_i \cdot d_i(x,y)]$$

Where:

- (X_i, Y_i) = coordinates of both the points of origin and destination;
- F_i = Inbound and outbound flows. For this case study the alimentary flown assigned as in 2.1.1 has been considered;
- R_i = transportation rate per unit $[€/(km * t)]$ (it depends on the weight and the distance). Due to the absence of data, it has been set at 1.
- d_i = distance between nodes. The Euclidean distance corrected by the circuit factor (for Italy is 1,3 [11]). In this case the formula for the distance became:

$$d_i = \sqrt{\left(lon_{x_i} - lon_X\right)^2 + \left(lat_{x_i} - lat_X\right)^2} \cdot 69 \cdot 1{,}609 \cdot 1{,}3$$

The method consists in the following steps:

1. First approximation of the center of gravity
$$X^* = \frac{\sum_i F_i \cdot R_i \cdot X_i}{\sum_i F_i \cdot R_i} \qquad Y^* = \frac{\sum_i F_i \cdot R_i \cdot Y_i}{\sum_i F_i \cdot R_i}$$
2. Compute the center of gravity
$$X^{**} = \frac{\sum_i \frac{F_i \cdot R_i \cdot X_i}{d_i}}{\sum_i \frac{F_i \cdot R_i}{d_i}} \qquad Y^{**} = \frac{\sum_i \frac{F_i \cdot R_i \cdot Y_i}{d_i}}{\sum_i \frac{F_i \cdot R_i}{d_i}}$$

3. Compute again the distance e continues in an iterative approach

4.1.2 Implementation in R

The method is implemented in R and in Figure 8 can be seen the result. The choice of the number of steps is 30.

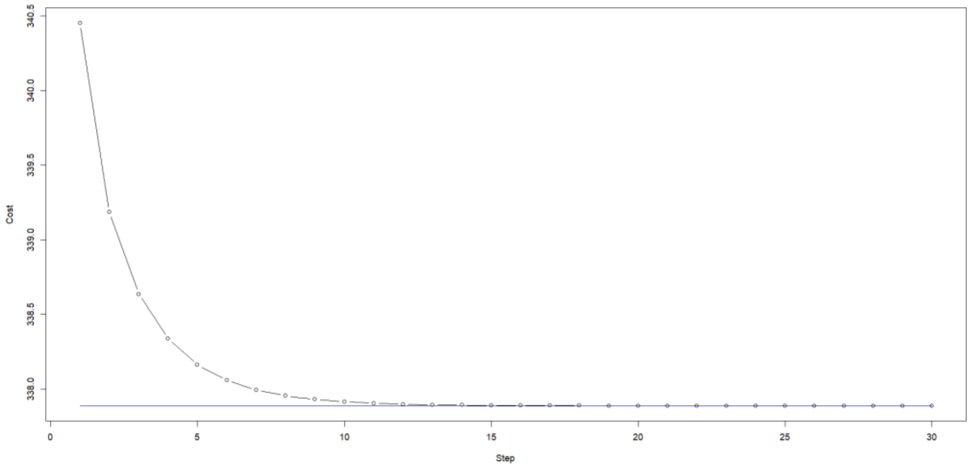


Figure 8 Iterative step of the method

As expected, there is a decrease of the cost in the first step and then it stabilizes reaching the minimum value. This mean that the location of the distribution center has been founded.

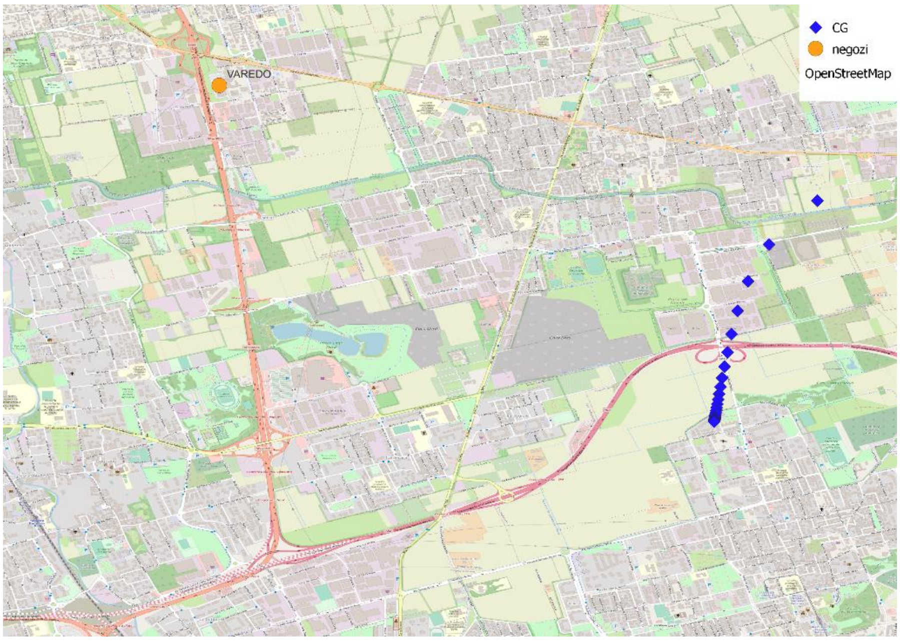


Figure 9 Center of Gravity location after each step

In Figure 9 the location of the Center of gravity after each step. As in the previous figure the location changes significantly in the first step and then stabilize in a point. Obviously, the result depends by the effective feasibility of the distribution center.

After this the entire analysis can be repeated by reassign the route, find the deviation point and so on.

4.2 What if we have more data?

4.2.1 What if there were more detailed data of the supplies' schedule?

In large-scale retail, logistics from the central warehouse to the final point of sale is a complex process, which has been simplified in the previous analyses. One element that has not been considered is the hour of the day at which supplies are carried out. The road network Lombardy is a very busy road and the trip time to each supermarket has, as a first consequence, a different travel time on the road network.

Assuming to have available the data of the hour of the day of supplying, it is possible to estimate the travel time on the road network in the different hour of the day and so the total time needed for transport from Pioltello to the shop and the return trip.

In the Table 1, as an example, it is reported the different hour of departure of the truck from Pioltello for the supply of each supermarket.

Table 1 – Schedule of supplies from Pioltello

		HOUR OF DEPARTURE FROM PIOLTELLO							
		1	4	7	10	13	16	19	22
POINT OF SALES	Lecco	x			x	x	x		
	Zara	x	x			x	x		x
	Monza		x	x		x	x	x	
	Seregno	x	x				x	x	

Those data could be represented on the map of the network to check which stores are restocked at a particular time. In addition, having this data for all the supermarket it is possible to visualize the number of trucks necessary in each time slot during the day, as shown in Figure 10.

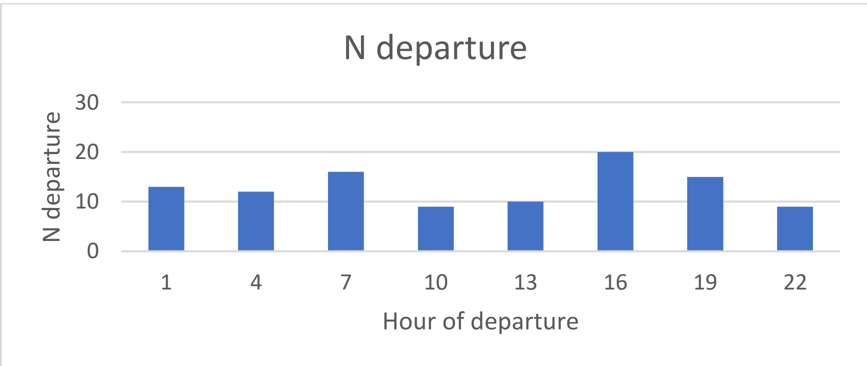


Figure 10 - Number of departures during the day

In this example, the number of trucks needed are variable during the day and the maximum value is at 16 and equal to 20.

Those data could be useful to optimize the use of the truck, the shift of the drivers and so a **minimization of the total costs of transportation**.

The company could adopt different strategies:

- Minimize the number of trucks and drivers necessary during the day, to have the same number of trucks needed during the day.
- Minimize the number of trips during the night, to minimize cost of drivers.
- Minimize the trips in the peak hour and redistribute in the other hour the day to minimize travel time and consumption. In this way there is also a benefit for the whole community because, choosing to travel in less busy hours, allows to improve the problem of daily traffic.
- Concentrate supplies to shops that can be reached from heavily trafficked roads during off-peak hours. This strategy could be the most critical to adopt, given that the area considered is all subject to heavy traffic.
- A combination of the earlier strategies.

4.2.2 What if we had more detailed data about sales?

More specific datasets about daily orders from shops, sales by product and shelf price of products, could have been matched with available to obtain:

- **Estimation of sales-loss risk**: knowing the items transported by each truck and their price, it could have been estimated the potential sales loss in case the truck does not arrive or arrive late. Nodes could be added of this attributed and ranked by their risk.
- **Prioritization strategy**: trucks could be prioritized according to minimize risk of sales loss or to optimize customer satisfaction (e.g., having on-offer products always available).

4.2.3 What if the function route() uses a routing function for truck and not for car?

The approximation cited in 2.2.2 consists in the equivalence of truck with car due to the absence of a truck routing function. As can be seen in Figure 11 , and comparing it with Figure 1 the difference in assigning the path is evident because instead of passing through Vimodrone the truck pass through the Milan's Tangenziale Est.

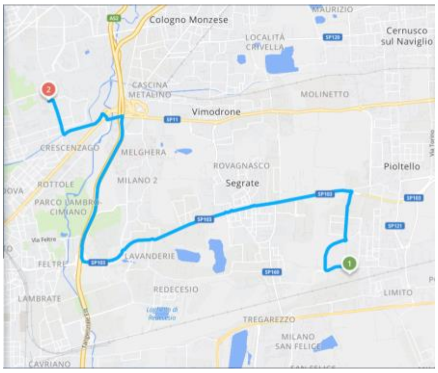


Figure 11 Route Distribution center - Adriano[13]

To solve this problem can address different ways:

- route() asks for a routing function. So instead of using osrmRoute another function designed for truck can be used.
- Changing the osrm.profile for car using it for truck. But to do that some data are needed such as height, width, length and weight as specified in [14]. The same source specific different weight forcing the truck to go on primary highway.

Obviously, the solution can be solved using the actual truck route planner used by Esselunga export it as shapefile containing each route and find the deviations point (the grey one in Figure 2). Then the analysis can be repeated.

5. Conclusions

The main problem of the project was the assignment of the route and some useful tool to address it has been found (is the case of the stplanr package). Then the classical network analysis has been made. The analysis obviously lacks due to the approximation and the low quantity of data available, but the potentiality of the analysis has been presented in the last part.

References

[1] Regione Lombardia, "Medie Strutture di Vendita," *dati.lombardia.it*, Jan. 05, 2022. https://www.dati.lombardia.it/Commercio/Medie-Strutture-di-Vendita/2si5-tumt (accessed Apr. 04, 2022).

[2] Regione Lombardia, "Grandi Strutture di Vendita," *dati.lombardia.it*, Jan. 05, 2022. https://www.dati.lombardia.it/Commercio/Grandi-Strutture-di-Vendita/haxp-4dkd (accessed Apr. 04, 2022).

[3] Geofabrik and OpenStreetMap Contributors, "Geofabrik Download Server," *http://download.geofabrik.de/*, 2022. http://download.geofabrik.de/europe/italy/nord-ovest.html (accessed Apr. 04, 2022).

[4] R. Lovelace and R. Ellison, "stplanr: A Package for Transport Planning," *The R Journal*, vol. 10, no. 2, p. 7, 2019, doi: 10.32614/RJ-2018-053.

[5] R. Lovelace *et al.*, "Documentation stplanr: Sustainable Transport Planning," *https://cran.r-project.org/*, Jun. 10, 2022. https://cran.r-project.org/package=stplanr (accessed Apr. 02, 2022).

[6] T. Giraud, R. Cura, M. Viry, and R. Lovelace, "Documentation osrm: Interface Between R and the OpenStreetMap-Based Routing Service OSRM," *https://cran.r-project.org/*, Jan. 24, 2022. https://cran.r-project.org/package=osrm (accessed Apr. 01, 2022).

[7] T. Giraud, "Issue #65: Does osrmRoute define shortest by time or distance?," *github.com*, Oct. 06, 2020. https://github.com/riatelab/osrm/issues/65 (accessed Apr. 28, 2022).

[8] G. Csardi, "Documentation cluster_edge_betweenness," *igraph.org*, 2022. https://igraph.org/r/doc/cluster_edge_betweenness.html (accessed Jun. 15, 2022).

[9] E. D. Kolaczyk and G. Csárdi, "Spectral Partitioning," in *Statistical Analysis of Network Data with R*, Second edition., Springer, 2020, pp. 62–64.

[10] G. Csardi, "Documentation cluster_leading_eigen," *igraph.org*, 2022.

[11] T. Modica, "Session 5 - Facility Location and Site Selection," *Politecnico di Milano*. 2021.

[12] C. Colicchia and T. Modica, "LOGISTICS AND FREIGHT TRANSPORTATION Lectures materials," *Politecnico di Milano*. 2021.

[13] unspecified, "Route planner," *dispatcher.roadlords.com*. https://dispatcher.roadlords.com/route-planner (accessed Jun. 22, 2022).

[14] Project OSRM, "Truck profile," *github*, 2019.