# COST ESTIMATES ASSOCIATED WITH DISTRIBUTION ROUTES USING DIJKSTRA'S ALGORITHM IN GEOGRAPHIC INFORMATION SYSTEMS (GIS)

*Luciana Pereira Barbosa[1, a,b], Morgana Cristhya Silva dos Santos[2,a,b], Allan Kardec Duailibe Barros Filho[3, a,b]*

[a] *Federal University of Maranhão, Brazil,*

[b] *Postgraduate Program in Electrical Engineering.*

**Abstract:** This work presents a simulation for obtaining distribution costs of products using Dijkstra's algorithm executed in a GIS environment. There are many considerations in these design problems, involving various constraints and decisions, and the costs associated with the construction, operation, and maintenance of a system are among the most critical factors for logistics. The present approach proposes a systematic search for optimal and near-optimal route solutions in the state of Maranhão. The investigation assumes that the use of geospatial data in the search for optimal routes can contribute to the reduction of transportation costs of goods. As a result, the simulations provided satisfactory results, generating costs associated with each generated route.

**Keywords:** Shortest Path; Dijkstra's Algorithm; Logistics; Costs.

# ESTIMATIVAS DE CUSTOS ASSOCIADOS A ROTAS DE DISTRIBUIÇÃO USANDO ALGORITMO DE DIJKSTRA EM SISTEMAS DE INFORMAÇÃO GEOGRÁFICA (SIG)

**Resumo:** Este trabalho apresenta uma simulação para obtenção de custos de distribuição de produtos usando algoritmo de Dijkstra executado em ambiente SIG. Há muitas considerações nesses problemas de projeto, envolvendo várias restrições e decisões, e os custos associados à construção, operação, manutenção de um sistema é um dos fatores mais importantes para a logística. A presente abordagem propõe uma busca sistemática de soluções ótimas e quase ótimas de rotas no estado do Maranhão. A investigação assumi que o uso de dados geoespaciais na busca de rotas ótimas pode contribuir para redução dos custos de transporte de mercadorias. Com isso, as simulações ofereceram resultados satisfatórios gerando custos associados a cada rota gerada.

**Palavras-chave:** Caminho Mínimo; Algoritmo de Dijkstra; Logística; Custos.

## 1. INTRODUCTION

The existence of a logistics system is one of the main factors for the development of a specific region. In routing, one of the methodologies used is the Shortest Path problem. The Shortest Path problem involves finding the shortest route or path from a starting point to a final destination. Typically, a graph is used to solve this problem, where a graph is an abstract mathematical object that contains sets of vertices and edges. The edges connect pairs of vertices.

In the real world, graph theory can be applied to different scenarios. For instance, to represent a map, we can use a graph where vertices represent cities, and edges represent the routes connecting the cities. If the routes are one-way, the graph is directed; otherwise, it is undirected.

Simulations in Geographic Information Systems (GIS) environments are a valuable strategy for analysis and decision support in product distribution, as the data used for pathfinding is geospatial and represents the existing real infrastructure.

This work is part of a larger project, and for this reason, the simulations consider the infrastructure of the state of Maranhão to obtain routes with the lowest cost. The objective of this work is to demonstrate that executing the minimum path algorithm in a GIS environment to obtain costs associated with routes is an efficient tool in decision support.

## 2. METHODOLOGY

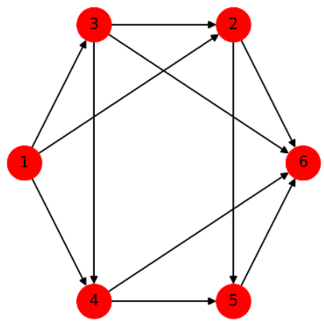
The research methodology of the work is of an applied (technological) nature. The methodology of this research comprised the following stages: literature review, data collection, selection of tools and programming language, implementation in the GIS environment, and testing.

### 2.1. Graph Representation

A graph G is an ordered set G = (V, A), where V represents a non-empty set of vertices, and A is a set of ordered pairs of vertices called arcs [1]. These arcs represent the connections between the vertices and may have a specific weight assigned if the graph is weighted. The weight of an arc represents the cost required to traverse it, as in the case of a weighted graph that may represent a road network connecting multiple cities.

In Figure 1, we illustrate the representation of a graph with six vertices and nine edges.

Figure 1: Graph exemple.




Dijkstra's algorithm solves the shortest path problem for a directed graph with non-negative weighted edges [2].

## 2.2. Dijkstra's Algorithm

The Dijkstra's algorithm consists of n iterations. If all vertices are visited, the algorithm terminates; otherwise, from the list of unvisited vertices, we need to choose the vertex with the minimum value in its label [1] [2]. After that, we consider all neighbors of this vertex (Neighbors of a vertex are those vertices that share common edges with the initial vertex). For each unvisited neighbor, we consider a new length, which is equal to the sum of the label value at the initial vertex v (d[v]) and the length of the edge I that connects them. If the resulting value is smaller than the value in the label, we need to update the label with the newly obtained value.

According to [4], the computational version of the Dijkstra's algorithm can be described as follows:

Figure 2: computational version of the Dijkstra's algorithm.

Given a directed graph $G(V, A)$ and a vertex **a** in $G$:
1. Initialize the minimum cost estimate of vertex **a** (root of the search) as zero and all other estimates as infinity.
2. Assign an arbitrary value to the predecessors (the predecessor of **a** vertex ***t*** is the vertex that comes before **t** in the minimum cost path from ***a*** to ***t***).
3. While there are open vertices:
- Select the still open vertex ***k*** with the smallest estimate among all open vertices.
- Close vertex ***k***.
- For each still open vertex j that is **a** successor of ***k***:
- Add the estimate of vertex ***k*** to the cost of the arc that connects ***k*** to ***j***.
- If this sum is better than the previous estimate for vertex ***j***, replace it and record ***k*** as the predecessor of ***j***.

The computational implementation of this algorithm has a complexity of $O(n^2)$, while, for example, the Floyd's Algorithm has a higher complexity of $O(n^3)$ [4]. This becomes a significant disadvantage as the number of vertices and edges (i.e., cities and roads) increases. Therefore, using Dijkstra's Algorithm allows the computer to perform the necessary calculations in a shorter amount of time.

### 2.3. Geographic Information Systems (GIS)

The technology of Geographic Information Systems (GIS) is becoming increasingly essential as a routine tool for visualizing and analyzing spatial information. It is a tool that enables the manipulation of geospatial and alphanumeric data, allowing for spatial analysis to assist in making decisions related to space, such as determining the most efficient delivery route to follow.

This tool is widely used in various applications, such as land use cartography (urban planning), transportation analysis and planning (traffic and emergency networks), geodemographic analysis (service location), mapping of infrastructure networks (gas, water, and electricity), and various natural resource management applications [5] [6].

The use of GIS offers numerous remote analysis tools, reducing costs. Simulations with these software packages provide benefits and cost reductions in analyzing real-life situations. Other related advantages include the extensive availability of georeferenced data in national open databases, enabling analysis using real-world routes.

## 3. RESULTS AND DISCUSSION

In general, Brazil presents a deficient logistics infrastructure. This problem is even more evident in the region above the sixteenth parallel. The state of Maranhão is one of the states in the Northeast that stands out, both for agribusiness and its availability of natural resources. Currently, the state has 217 municipalities and a population of 6,775,152 people, as shown in Figure 3(a).

To conduct the simulation, the cartographic database was used to obtain the entire transportation network and the data of all existing cities in the state provided by the Brazilian Institute of Geography and Statistics (IBGE). Each city was considered as a vertex, and the roads as arcs connecting one vertex to another. The state of Maranhão has a prominent port, especially for grain and other product shipments, and optimizing logistics systems can significantly contribute to the state's development.

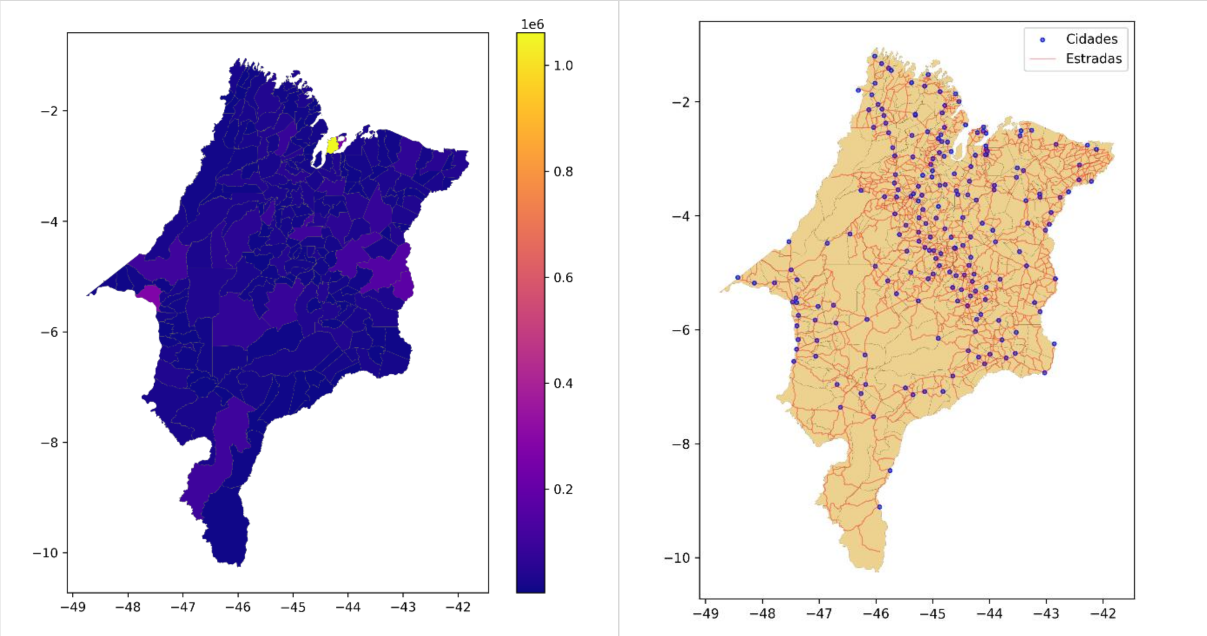
The ongoing work is an analysis of the network aimed at obtaining routes with the lowest cost in the distribution of natural gas, with the objective of promoting and contributing to the internalization of natural gas as a transitional energy source, especially in mitigating $CO_2$ emissions from the transportation sector.

The relevance of this study lies mainly in the fact that one of the sectors that emit the most $CO_2$ is the transportation sector, and replacing diesel with natural gas can reduce these emissions.

In a Geographic Information System (GIS), the Dijkstra's algorithm uses vertices corresponding to coordinates connected by polylines, representing roads or transportation routes. These polylines have lengths associated with weights that indicate the speed achieved by vehicles [7]. In networks with many vertices, the edges between nodes can have different attributes, such as time or distance, making it essential to have reasonable values for each possible scenario.

Therefore, the choice of the path should aim to obtain the shortest route among the available options. Figure 3 (b) shows all the points considered and the road network of the state.

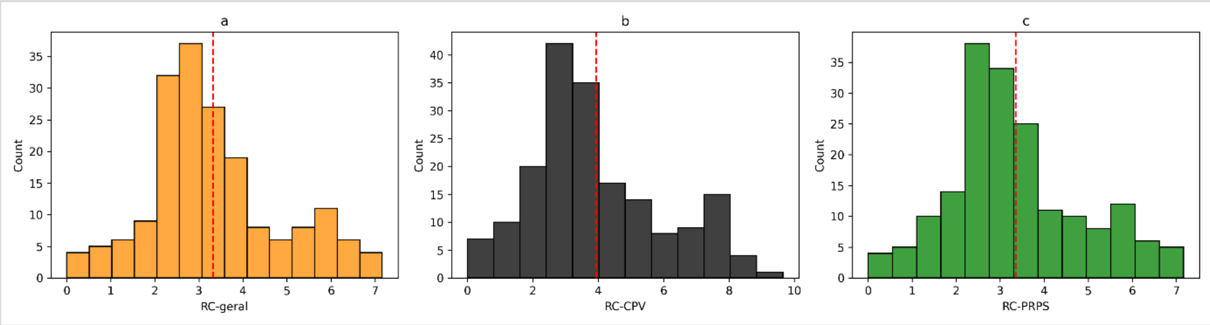
Figure 3: a) Maranhão Population Concentration; (b) Considered nodes and edges.




The deployment of distribution infrastructure holds a strategic nature, demanding studies and simulations that indicate the most viable directions, especially in terms of costs. Armed with this data, three simulations were conducted: one considering the entire existing network, another considering only roads classified as paved (RC-PV), and finally paved roads with loose surface (RC-PRPS). These imposed restrictions mainly arose because, in certain cases, the shortest path may not be the fastest, which could increase these costs.

After that, the simulations were carried out, where it was possible to obtain the costs associated with each vertex, establishing the shortest path for each route starting from the capital, São Luís, and analyzing 203 municipalities. The costs of the routes are presented graphically through histograms in Figure 4.

Figure 3: Costs obtained for each simulation.




Out of the 203 selected municipalities, in all cases, there were cities not reached by the simulation, such as the municipality of Matinha/MA. In the first case (Figure 4a), using the entire road network of Maranhão, there were two municipalities for which the costs of the routes were not obtained, Matinha and Tasso Fragoso. One possible explanation for this case is a potential distortion in the geometry of the vector layer of roads.

In the second case (Figure 4b), there were 20 municipalities not reached, as the restriction imposed was the selection of paved roads only. And in the third case (Figure 4c), considering both paved roads and those with loose primary surfacing, only the municipality of Matinha was not reached. Table 1 presents the main statistics of the costs associated with the routes in the simulations.

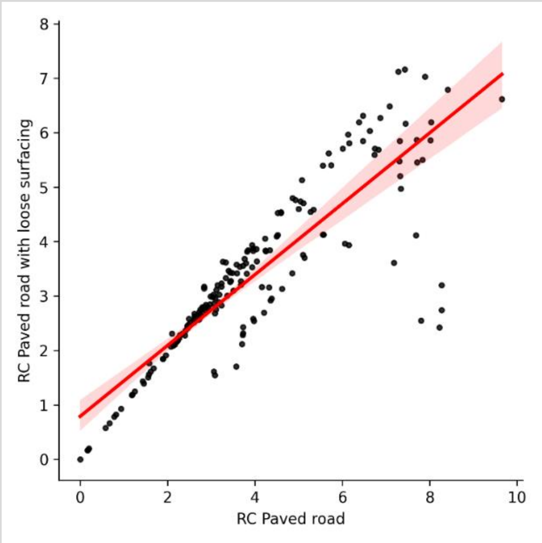
Table 1: Statistical comparison.

| Statistics | | | |
|---|---|---|---|
| Simulation | General | RC-CPV | RC-PRPS |
| **Unique values** | 201 | 180 | 202 |
| **Null (missing) values** | 2 | 20 | 1 |
| **Minimum value** | 0,00068572 | 0,00002637 | 0,00024176 |
| **Maximum value** | 10,18619068 | 9,65278566 | 10,18711816 |
| **Sum** | 688,95200479 | 717,68078375 | 703,31200255 |
| **Mean value** | 3,42762191 | 3,92175292 | 3,48174259 |
| **Median value** | 3,13784218 | 3,47703821 | 3,17399135 |
| **Standard deviation** | 1,58391706 | 2,00625951 | 1,64638359 |
| **Coefficient of variation** | 0,46210378 | 0,51157214 | 0,47286195 |
| **First quartile** | 2,45161395 | 2,60462876 | 2,45254142 |
| **Third quartile** | 4,24100009 | 5,05963522 | 4,37166765 |
| **Interquartile range (IQR)** | 1,78938614 | 2,45500646 | 1,91912623 |

In all cases, the route costs fall within the range of 0 to 10.18711816. With means between 3 and 4, it indicates that the majority of the costs acquired in this problem are concentrated within this interval.

Thus, it becomes clear that the use of geographic information systems for obtaining optimal or near-optimal routes is effective tools, and they can contribute to decision-making in product distribution. These types of simulations also yielded good results in obtaining costs associated with each vertex of the generated tree. By using simulations with constraints, we can observe in the scatter plot the cost relationship for the Shortest Route (RC) simulations for the two cases, as shown in Figure [5].

Figure 4: Cost behavior with the restrictions.




In summary, the costs generated by applying road restrictions are mostly within a certain range; however, when using other roads in addition to the paved ones, there is an increase in these costs. Moreover, there are free software solutions available for this purpose, offering reduced costs both in financial and computational terms for logistics problems.

## 4. CONCLUSION

Simulations in GIS environments offer numerous advantages, primarily by providing conditions for remote analysis of specific issues. The graph search algorithm is a highly effective optimization tool for solving practical problems, such as routing. The Dijkstra's algorithm, also known as the shortest path algorithm executed in a GIS environment, yielded the expected results, such as minimum costs associated with simulated routes and their respective distances. In logistics, this is crucial for optimizing transportation route planning, such as choosing the most efficient path for goods delivery. Thus, it can be perceived that the Dijkstra's algorithm plays a pivotal role in optimizing logistic operations, contributing to efficiency, cost reduction, and enhanced customer experience in various logistics and transportation domains.

However, these costs are primarily associated with the distances traveled. Therefore, the use of additional information and constraints is necessary to make this simulation as close to reality as possible.

### Acknowledgments

My thanks for the financial support from the National Agency of Petroleum, Natural Gas, and Biofuels and the Financier of Studies and Projects (FINEP), through the Human Resources Program of ANP for the Oil and Gas Sector - PRH-ANP (PRH 54.1- UFMA).

## 5. REFERENCES

[1] BAEZA, Daniel; IHLE, Christian F.; ORTIZ, Julián M. A comparison between ACO and Dijkstra algorithms for optimal ore concentrate pipeline routing. Journal of Cleaner Production, v. 144, p. 149-160, 2017.

[2] BRUGNOLLI, Mateus Mussi; Lopes, Guilherme Cano; Lemos, M. A. A hybrid solution for navigation and obstacle avoidance based on the Dijkstra's algorithm and fuzzy logic. Olimpo. Sorocaba, 2013.

[3] DA CUNHA, Angélica et al. Route Tracing Using Dijkstra's Algorithm.

[4] DE CARVALHO, B. M. P. S. Dijkstra's Algorithm. Universidade de Coimbra, Coimbra, Portugal, 2008.

[5] TOMLINSON, Roger. Thinking about GIS: Geographic Information System Planning for Managers. Esri Press, 2013.

[6] RIOS, Marcel L.; NETO, Francisco SS; NETTO, José F. Magalhaes. Analysis and Comparison of Dijkstra and A-Star Algorithms in Discovering Shortest Paths on Grid Maps. In: Anais do I Encontro de Teoria da Computação. SBC, 2016. p. 887-890.

[7] BONETT, João. Graph theory and the Dijkstra algorithm applied to the analysis of transportation networks and the construction of network displacement matrices. 21º Congresso Brasileiro de Transporte e Trânsito, São Paulo, jun. 2017.

[8] SOUSA, Nícolas Gordiano Barbosa de. Aplicação do algoritmo de dijkstra para otimização de rotas em percursos urbanos utilizando transporte público. 2019.

[9] PRIETO, Pedro S. An application of graph theory for optimizing costs of international aeronautical routes.. 2022.

[10] LAURO, Adriano; DE BEAUCLAIR SEIXAS, Roberto; GUARDIA, Luiz Ernesto Torres. Shortest Path in Dynamic Cost Matrices.. IV Simpósio de Pesquisa Operacional e V Simpósio de Logística da Marinha–SPOLM 2001, 2001.