Road network: Modeling and vulnerability analysis

Cristian Jimenez, Jaider Pinto, Jimmy Prieto

*Abstract*— In this paper, we conducted a literature review and collected data from the INVÍAS website on the main roads in Colombia. We organized and processed this data using Networkx, and applied measures of centrality to generate a network visualization. Our analysis revealed that Bogotá is the principal city in the country. By using the processed data, we can provide recommendations, such as identifying areas where public investments should be focused or identifying critical points in the network.

*Index Terms*— Network science, Road network, Road system, network centrality, CNA, .

# Introduction

The study also explores Colombia's challenges in developing transportation infrastructure and the vital role of transportation systems in economic growth. Additionally, complex network analysis is utilized to identify vulnerable nodes, connectivity issues, and clusters of related entities. The dataset used includes significant points and various sections of national roads represented as nodes and road segments as edges, forming a network model to evaluate the vulnerability of the road system. The research concludes that Bogotá is the country's principal city and highlights the necessity of investing in transport infrastructure to enhance economic development.

# Goal And Scope

Vulnerability analysis and possible effects on a logistics network In Colombia.

# Literature Review

## Colombian Road Network

The road network in Colombia is composed of the Primary Network (major highways managed by the national government), the Secondary Network (managed by departments), and the Tertiary Network (composed of inter-veredal roads or paths managed by municipalities). Colombia has a road network of 206,102 km, of which 6.9%, or 16,983 km, correspond to the Primary Network, 21% or 44,400 km correspond to the Secondary Network, and 69.46%, or 142,284 km correspond to the Tertiary Network.[1]

The Fourth Generation (4G) concession projects are the most ambitious road projects in the history of Colombia. Under a cost of approximately $18 billion in the Public-Private Partnership modality, 8,000 kilometers will be built, including 1,370 km of double-lane highways, and 160 tunnels in more than 40 new concessions. This will improve relevant aspects such as travel times, social and economic benefits, benefits for producers, road safety, connectivity, maintenance of the concession section, especially for freight, from manufacturing points to export ports, and accessibility.

It is projected that the works will be executed within a maximum of 6 years from the date of their award. When the 4G road construction is completed, one of the benefits of the project is expected to be a 30% reduction in travel distances.

For Colombia, this is a national point of interest, considering that land transport has high costs and long travel times. Therefore, strategies must be generated that contribute to achieving international competitiveness and infrastructure standards. Regarding travel times, approximately 80% of cargo transport in Colombia is carried out by land, through the country's six main logistics corridors: Bogotá-Buenaventura, Medellín-Villavicencio, Bogotá-Cúcuta-Caribe, Bogotá-Caribe, Rumichaca-Caribe, and Medellín-Cúcuta.

The road structure in Colombia presents permanent challenges for several reasons, including the country's structural and topographical conditions not being the best, as Colombia has three mountain ranges that generate gaps in connectivity for the country.[2]

## Transport infrastructure

The development of transportation infrastructure is crucial for the economic growth of a nation. Good transportation systems contribute to the redistribution of economic activities and the development of prosperous regions, requiring continual transportation improvements. Many studies have found a positive correlation between transportation and economic development. Accessibility to cities is particularly significant for their competitive advantage and tourism industry. Ground transportation is viewed as critical infrastructure, essential for basic societal operations, and has environmental benefits over road-only transportation.

Various studies have investigated the importance of the transportation system for the economy, society, and the environment using different methods. One such method is Complex Network Analysis (CNA), which represents the interactions between system entities using links (represented by nodes in the network model, such as cities or stations in transportation studies). A topological analysis is conducted on the network structure to identify vulnerable nodes, connectivity issues, or clusters of related entities. CNA has been used in many countries to study different modes of transportation, including air transportation, subways, buses, railways, and maritime transportation.[3]–[5]

# Case Study

In our project we want to review the different metrics and measures that we can apply to our dataset on the different nodes and road routes in our country and review the feasibility of performing vulnerability analysis on the resulting network.

# Data Set

The data set used in this article was found in the invias website, they showed enter to finish city road, all the data is bidirectional and is distributed throughout the country. The dataset was an excel so we had to make some process of the data to get our dataset for the network.

# Implemented Network Science Approach

To carry out the desired approach and development, we began with obtaining data as detailed below with the data life cycle, and then, with the data obtained, we transformed this data into a network (model explain as a complex network) in which we primarily obtained and analyzed centrality measures.

## Data Life Cycle

### Generation

The data generation was carried out by the National Institute of roadways (INVIAS) through a data collection over the years, mapping the different national routes in the country. This data contains each of the important points and different sections of each of the national roads in the country.

### Collection

This stage of the cycle was carried out by us. Using the various open data provided by INVIAS, we collected the most important data from the national routes, taking the origin and destination points of each section as the basis for the data we needed.

### Processing

The data processing consisted of 2 important parts. The first was the cleaning of the dataset, in which irrelevant points as well as incomplete or erroneous data were eliminated. Then, the obtaining of origin and destination points was carried out so that pairs of nodes were available. This had to be done because there were sections that included more than 2 cities, so these cases had to be divided to obtain a clean and structured dataset.

### Storage

With the cleaned dataset, we saved only the necessary data, which was stored in a table with two columns (currently): the origin and destination. Each column contains the name of the node (city, town, point of interest, etc.).

### Management

This data was used to generate a network in which the origin and destination points are the nodes of the network, and the edges represent the route between 2 nodes. This network is our model with which we will start the analysis of centrality metrics and propose a vulnerability analysis of the road network.

### Analysis

Currently, we are at this point in the data life cycle. So far, we have obtained the different centrality measures of the road network to perform an analysis of the most important points of the network, as well as the most vulnerable ones. The next step is to make a prediction of possible critical points and predict viable options for preventing network problems.

## Model of complex network

Complex networks are simplified representations of complex real-world systems using graph theory, which capture the intricate interactions and connections within these systems. A network is made up of nodes and edges, where each node represents an individual entity in the system and each edge represents the relationship or interaction between two entities. As the road system is a constantly evolving complex system, it can also be analyzed using complex network theory. In this study, intersections are considered as nodes, and road segments are represented as edges to construct a network model for analyzing the vulnerability of road system. Therefore, the road network can be represented as a graph , where is the set of nodes, is the set of edges, is the set of edge weights, and is the number of nodes.[3], [5]

## Centrality Measures

Freeman's significant contributions to structural sociology involved consolidating and reviewing previous research dating back to the early fifties, including works by Bavelas, Leavitt, Shimbel, and Shaw. Freeman's work established the first set of centrality indices, including degree, closeness, and betweenness centrality. The fundamental concept in structural sociology is to represent a social or organizational group as a network with nodes representing individuals and edges representing their relationships. Bavelas was the first to recognize that central individuals in a social network often hold a prominent role in the group, with a good location in the network structure correlating with independence, influence, and control over others. The same idea can be used for another kind of networks like in this case the road network, these measures can help to understand the importance of a node (city or main place) in the network and the isolation of each node making see the importance of each node and how the existence/inexistence can affect the network.[5]

### Degree Centrality

The concept of degree centrality is based on the notion that important nodes within a graph have the most connections to other nodes. The degree of a node corresponds to the number of edges associated with the node, which is equivalent to the number of its immediate neighbors. The normalization used for degree centrality produces values between 0 and 1, with a value of 1 indicating that a node is connected to every other node within the graph. However, degree centrality may not be particularly relevant in primal urban networks where a node's degree (the number of roads connected to that node) is constrained by geographic factors.[5]The degree of node is defined in terms of the adjacency matrix as:

The degree centrality () of is defined as:

### Closeness Centrality

The most basic interpretation of closeness is founded on the notion of minimum distance or geodesic . This represents the shortest possible summation of edge lengths over all feasible paths within a weighted graph between i and j, or in a topological graph, the minimum number of edges traveled. , or closeness centrality, is best utilized when evaluating measures that rely on independence. This centrality index is only applicable to connected graphs unless one assumes a finite value for when there is no available path between two nodes and .[5] For non-valued graphs, the centrality index ranges from 0 to 1. The closeness centrality () of is defined as:

### Betweenness Centrality

The relationships between two points that are not directly connected may be influenced by other actors, particularly those along the paths between them. Therefore, actors in the middle may exert strategic control and influence over the others. The basic concept of betweenness centrality is that an actor is central if it lies on many of the shortest paths connecting other actors. Namely, if is the number of geodesics linking the two actors and , and is the number of geodesics linking the two actors and that contain point , the betweenness centrality of actor i can be defined as:

is a measure of betweenness centrality that ranges between 0 and 1 and is highest when actor lies on all the shortest paths. Freeman has proposed various extensions to the betweenness index. When communication does not necessarily follow the shortest path, a more realistic betweenness measure should consider non-shortest paths as well. Two such measures are flow betweenness and random path betweenness, but in our study, we focus only on the simplest case, which is the shortest path betweenness described in formula (4)[5].

# Preliminary results

First, we must process the data and then do several steps to get the information required to create a network using Networkx. The preliminary network that we obtain is shown in the next figure:

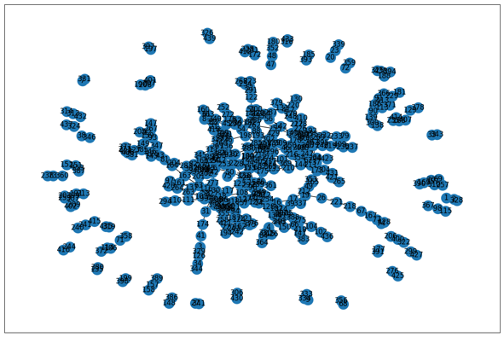


Fig 1 Preliminary Network

From the obtained network we can see that it is not connected, there are many nodes that are isolated and connected with a few other nodes, but without being connected to our large component.

We apply some of the centrality measures for a network and see that the node 6, has the largest value in degree centrality:

6: 0.02968036529680365, the largest value in closeness centrality 6: 0.07820045333596137 and the largest value in betweenness centrality 6: 0.23994774508238034, this is expected since node 6 corresponds to Bogota, which is the largest and most important city in the country, so it gives meaning to our network.

We also have gotten the great component from our network of which we are going to carry out the vulnerability validations:

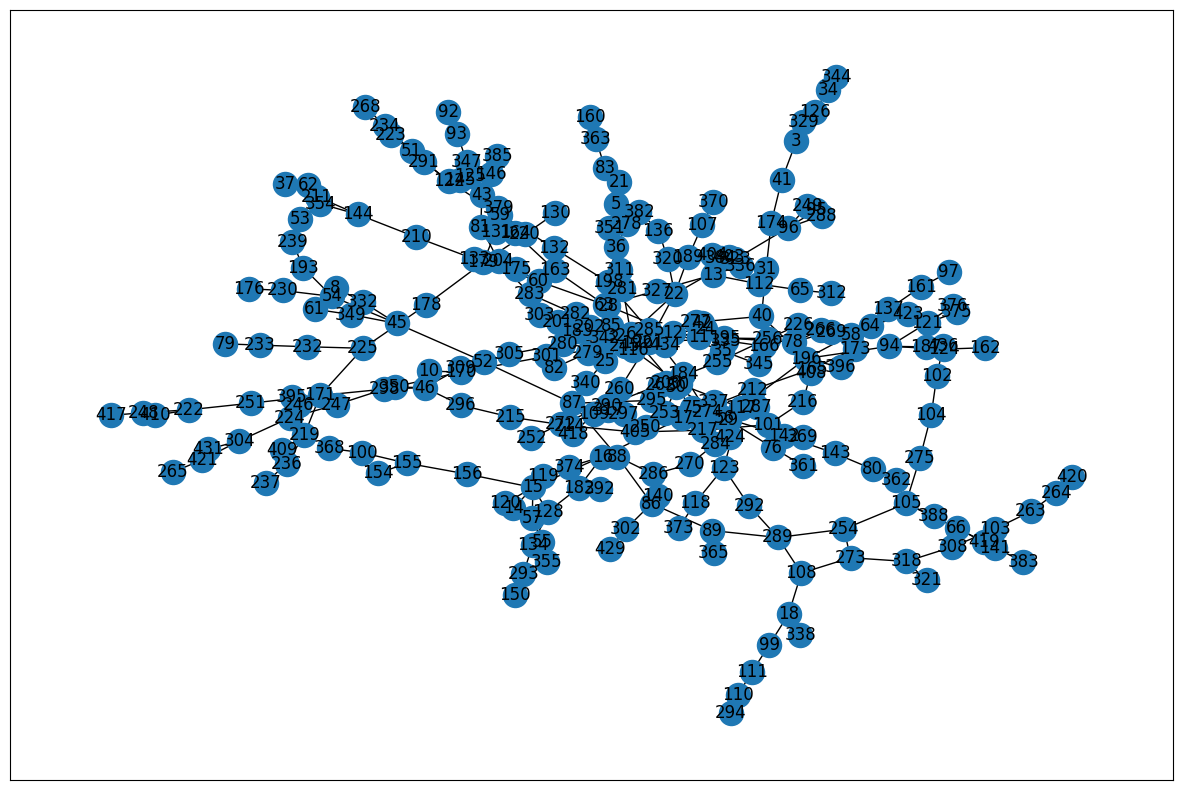


Fig 2 Fixed Network

# Links

## Source code

<https://github.com/Jaider0111/NSFDA_09_Term_Project>

## Explicative video

<https://drive.google.com/drive/folders/1evJOT9Hqp23YWtTrEJaawvOzRXpRtn96?usp=share_link>

# Team Members

|  |  |  |
| --- | --- | --- |
| Team Member | Role | Activities |
| Jaider Pinto | Leader | Guide the team for the goal. |
| Cristian Jimenez | Investigator | Discover |
| Jimmy Prieto | Investigator | Apply |

Table 1. Team members

# Conclusions

We see that the initially obtained network makes sense by placing expected nodes as important, as nodes with the highest metrics

There are many nodes, separated from the large component, we must carry out a major review to see if this is the case, (they are isolated) or it is a matter of the data.

References

[1] “2.3 Colombia Red Carretera - Logistics Capacity Assessment - Digital Logistics Capacity Assessments.” https://dlca.logcluster.org/display/public/DLCA/2.3+Colombia+Red+Carretera (accessed Apr. 17, 2023).

[2] “Portal INVÍAS - Colombia.” https://www.invias.gov.co/ (accessed Apr. 17, 2023).

[3] Z. Wang, Y. Pei, J. Liu, and H. Liu, “Vulnerability analysis of urban road networks based on traffic situation,” *International Journal of Critical Infrastructure Protection*, vol. 41, p. 100590, Jul. 2023, doi: 10.1016/J.IJCIP.2023.100590.

[4] L. Calzada-Infante, B. Adenso-Díaz, and S. García Carbajal, “Analysis of the European international railway network and passenger transfers,” *Chaos Solitons Fractals*, vol. 141, Dec. 2020, doi: 10.1016/J.CHAOS.2020.110357.

[5] S. Porta, P. Crucitti, and V. Latora, “The network analysis of urban streets: A primal approach,” *Environ Plann B Plann Des*, vol. 33, no. 5, pp. 705–725, 2006, doi: 10.1068/B32045.